(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 3 291 679 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
15.12.2021 Bulletin 2021/50

(21) Application number: **16719873.8**

(22) Date of filing: **03.05.2016**

(51) Int Cl.:
A01N 63/00 *(2020.01)* ***A61K 31/7105*** *(2006.01)*
C12N 9/16 *(2006.01)* ***C12N 15/113*** *(2010.01)*

(86) International application number:
PCT/EP2016/059803

(87) International publication number:
WO 2016/177682 (10.11.2016 Gazette 2016/45)

(54) **ALTERING MICROBIAL POPULATIONS & MODIFYING MICROBIOTA**

VERÄNDERUNG MIKROBIELLER POPULATIONEN UND MODIFIZIERUNG DER MIKROBIELLEN FLORA

CHANGEMENT DE POPULATIONS MICROBIENNES ET MODIFICATION DU MICROBIOTE

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **06.05.2015 GB 201507773**
06.05.2015 GB 201507774
06.05.2015 GB 201507775
06.05.2015 GB 201507776
17.05.2015 GB 201508461
31.05.2015 GB 201509366
20.06.2015 GB 201510891
17.10.2015 GB 201518402
10.01.2016 GB 201600418
10.01.2016 GB 201600417

(43) Date of publication of application:
14.03.2018 Bulletin 2018/11

(60) Divisional application:
18205170.6 / 3 461 337
19179004.7 / 3 563 684
20154858.3 / 3 729 967
20155001.9 / 3 711 488
21170379.8 / 3 903 586
21170380.6 / 3 907 285

(73) Proprietor: **SNIPR Technologies Limited**
London EC1V 2NX (GB)

(72) Inventors:
- **CLUBE, Jasper**
London EC1V 2NX (GB)
- **SOMMER, Morten**
2100 Copenhagen (DK)
- **GRØNDAHL, Christian**
2100 Copenhagen (DK)
- **VAN DER HELM, Eric**
2100 Copenhagen (DK)
- **VAZQUEZ-URIBE, Ruben**
2100 Copenhagen (DK)

(74) Representative: **CMS Cameron McKenna Nabarro Olswang LLP**
Cannon Place
78 Cannon Street
London EC4N 6AF (GB)

(56) References cited:
WO-A1-2014/124226

- **A. A. GOMAA ET AL: "Programmable Removal of Bacterial Strains by Use of Genome-Targeting CRISPR-Cas Systems", MBIO, vol. 5, no. 1, 28 January 2014 (2014-01-28), pages e00928-13, XP055227960, DOI: 10.1128/mBio.00928-13**
- **ROBERT J CITORIK ET AL: "Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases", NATURE BIOTECHNOLOGY, NATURE PUBLISHING GROUP, US, vol. 32, no. 11, 1 November 2014 (2014-11-01), pages 1141-1145, XP002742964, ISSN: 1087-0156, DOI: 10.1038/NBT.3011 [retrieved on 2014-09-21]**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Printed by Jouve, 75001 PARIS (FR)

(Cont. next page)

- **M. JINEK ET AL: "A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity", SCIENCE, vol. 337, no. 6096, 17 August 2012 (2012-08-17), pages 816-821, XP055229606, US ISSN: 0036-8075, DOI: 10.1126/science.1225829**
- **KURT SELLE ET AL: "Harnessing CRISPR-Cas systems for bacterial genome editing", TRENDS IN MICROBIOLOGY., vol. 23, no. 4, 1 April 2015 (2015-04-01), pages 225-232, XP055281790, GB ISSN: 0966-842X, DOI: 10.1016/j.tim.2015.01.008**
- **Z. XIE ET AL: "Development of a Tunable Wide-Range Gene Induction System Useful for the Study of Streptococcal Toxin-Antitoxin Systems", APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 79, no. 20, 9 August 2013 (2013-08-09), pages 6375-6384, XP055281880, US ISSN: 0099-2240, DOI: 10.1128/AEM.02320-13**
- **KIMBERLEY D. SEED ET AL: "A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity", NATURE, vol. 494, no. 7438, 27 February 2013 (2013-02-27), pages 489-491, XP055281884, United Kingdom ISSN: 0028-0836, DOI: 10.1038/nature11927 cited in the application**

Remarks:

The file contains technical information submitted after the application was filed and not included in this specification

Description

FIELD OF THE INVENTION

[0001] The invention relates to methods of inhibiting bacterial population growth, altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria, nucleic acid arrays for this purpose and vectors comprising the arrays. The invention relates to engineered systems for modifying host cell nucleic acid, components of such systems and application of these in industry and medicine. The invention is particularly useful, for example, for treatment of microbes such as for environmental, food and beverage use. The invention relates *inter alia* to methods of controlling microbiologically influenced corrosion (MIC) or biofouling of a substrate or fluid in an industrial or domestic system. The invention also relates to treated fluids and vectors for use in the methods. In embodiments, the methods use horizontal transfer of arrays. The invention also provides arrays comprised by mobile genetic elements (MGEs) for this purpose and vectors comprising these arrays.

BACKGROUND OF THE INVENTION

[0002] Inhibiting bacterial population growth and altering the relative ratios of different bacterial species in a mixture finds application in a wide range of industries and settings, for example for treatment of waterways, drinking water or in other environmental settings. Application is also found in altering bacteria in humans and non-human animals, eg, livestock, for reducing pathogenic infections or for re-balancing gut or oral microbiota. Recently, there has been interest in analysing the relative proportions of gut bacteria in humans with differing body mass or obesity profiles, or in investigating possible bacterial influence in disease contexts such as Crohn's disease.

[0003] Although bacterial innate immune mechanisms against phage abound, an extensively documented bacterial adaptive immune system is the CRISPR/Cas system. Engineered CRISPR/Cas systems have been used for precise modification of nucleic acid in various types of prokaryotic and eukaryotic cells, ranging from bacterial to animal and plant cells (eg, see Jiang W *et al* (2013)). Prokaryotes, such as bacteria and archaea, encode adaptive immune systems, called CRISPR/Cas (clustered regularly interspaced short palindromic repeats/CRISPR associated), to provide resistance against mobile invaders, such as viruses (eg, bacteriophage) and plasmids. Reference is made to Seed *et al* (2013), which explains that bacteriophages (or phages) are the most abundant biological entities on earth, and are estimated to outnumber their bacterial prey by tenfold. The constant threat of phage predation has led to the evolution of a broad range of bacterial immunity mechanisms that in turn result in the evolution of diverse phage immune evasion strategies, leading to a dynamic co-evolutionary arms race.

[0004] Host immunity is based on incorporation of invader DNA sequences in a memory locus (CRISPR array), the formation of guide RNAs from this locus, and the degradation of cognate invader DNA (protospacer) situated adjacent a protospacer adjacent motif (PAM). See, for example WO2010/075424. The host CRISPR array comprises various elements: a leader (including a promoter) immediately 5' of one or more repeat-spacer-repeat units where the repeats are identical and the spacers differ. By acquiring spacer sequence from invading virus or plasmid nucleic acid, the host defence system is able to incorporate new spacers into the CRISPR array (each spacer flanked by repeats) to act as a memory to tackle future invasion by the virus or plasmid. It has been observed that recently-acquired spacers tend to be inserted into the host array directly after the leader.

[0005] Reference is made to Heler *et al* (2014), which explains that CRISPR loci and their associated genes (Cas) confer bacteria and archaea with adaptive immunity against phages and other invading genetic elements. A fundamental requirement of any immune system is the ability to build a memory of past infections in order to deal more efficiently with recurrent infections. The adaptive feature of CRISPR-Cas immune systems relies on their ability to memorize DNA sequences of invading molecules and integrate them in between the repetitive sequences of the CRISPR array in the form of 'spacers'. The transcription of a spacer generates a small antisense RNA that is used by RNA-guided Cas nucleases to cleave the invading nucleic acid in order to protect the cell from infection. The acquisition of new spacers allows the CRISPR-Cas immune system to rapidly adapt against new threats and is therefore termed 'adaptation' (ie, vector sequence spacer acquisition).

[0006] Seed *et al* (2013) reported a remarkable turn of events, in which a phage-encoded CRISPR/Cas system was used to counteract a phage inhibitory chromosomal island of the bacterial host. A successful lytic infection by the phage reportedly was dependent on sequence identity between CRISPR spacers and the target chromosomal island. In the absence of such targeting, the phage-encoded CRISPR/Cas system could acquire new spacers to evolve rapidly and ensure effective targeting of the chromosomal island to restore phage replication. Bondy-Denomy *et al* (2012) describe the early observed examples of genes that mediate the inhibition of a CRISPR/Cas system. Five distinct 'anti-CRISPR' genes were found in the genomes of bacteriophages infecting *Pseudomonas aeruginosa.* Mutation of the anti-CRISPR gene of a phage rendered it unable to infect bacteria with a functional CRISPR/Cas system, and the addition of the same gene to the genome of a CRISPR/Cas-targeted phage allowed it to evade the CRISPR/Cas system.

[0007] Immature RNAs are transcribed from CRISPR arrays and are subsequently matured to form crRNAs. Some CRISPR/Cas systems also comprise sequences encoding trans-activating RNAs (tracrRNAs) that are able to hybridise to repeats in the immature crRNAs to form pre-crRNAs, whereby further processing produces mature, or crRNAs. The architecture of cRNAs varies according to the type (Type I, II or III) CRISPR/Cas system involved.

[0008] CRISPR-associated (cas) genes are often associated with CRISPR arrays. Extensive comparative genomics have identified many different cas genes; an initial analysis of 40 bacterial and archaeal genomes suggested that there may be 45 cas gene families, with only two genes, cas1 and cas2, universally present. Cas1 and Cas2 are believed to be essential for new spacer acquisition into arrays, thus are important in mechanisms of developing resistance to invader nucleic acid from phage or plasmids. Nuñez *et al* (2015) reportedly demonstrated the Cas1-Cas2 complex to be the minimal machinery that catalyses spacer DNA acquisition and apparently explain the significance of CRISPR repeats in providing sequence and structural specificity for Cas1-Cas2-mediated adaptive immunity.

[0009] CRISPR/Cas systems also include sequences expressing nucleases (eg, Cas9) for cutting invader nucleic acid adjacent cognate recognition motifs (PAMs) in invader nucleotide sequences. PAM recognition of nucleases is specific to each type of Cas nuclease. The PAMs in the invader sequences may lie immediately 3' of a protospacer sequence, with nucleases typically cutting 3-4 nucleotides upstream of (5' of) the PAM. The conservation of the PAM sequence differs between CRISPR-Cas systems and appears to be evolutionarily linked to cas1 and the leader sequence. Fineran *et al* (2014) observed that Invaders can escape type I-E CRISPR-Cas immunity in Escherichia coli K12 by making point mutations in a region (the "seed region") of the protospacer or its adjacent PAM, but hosts quickly restore immunity by integrating new spacers in a positive-feedback process involving acquisition ("priming"). To date, the PAM has been well characterized in a number of type I and type II systems and the effect of mutations in the protospacer has been documented (see references 5, 14, 23, 46, 47 in Fineran *et al* (2014)). Fineran *et al* (2014) concluded that their results demonstrated the critical role of the PAM and the seed sequence, in agreement with previous work.

[0010] Semenova *et al* (2011) investigated the role of the seed sequence and concluded that that in the case of Escherichia coli subtype CRISPR/Cas system, the requirements for crRNA matching are strict for the seed region immediately following the PAM. They observed that mutations in the seed region abolish CRISPR/Cas mediated immunity by reducing the binding affinity of the crRNA-guided Cascade complex to protospacer DNA.

[0011] The stages of CRISPR immunity for each of the three major types of adaptive immunity are as follows:-

(1) Acquisition begins by recognition of invading DNA by Cas1 and Cas2 and cleavage of a protospacer;
(2) A protospacer sequence is ligated to the direct repeat adjacent to the leader sequence; and
(3) Single strand extension repairs the CRISPR and duplicates the direct repeat.

[0012] The crRNA processing and interference stages occur differently in each of the three major types of CRISPR systems. The primary CRISPR transcript is cleaved by Cas to produce crRNAs. In type I systems Cas6e/Cas6f cleave at the junction of ssRNA and dsRNA formed by hairpin loops in the direct repeat. Type II systems use a trans-activating (tracr) RNA to form dsRNA, which is cleaved by Cas9 and RNaseIII. Type III systems use a Cas6 homolog that does not require hairpin loops in the direct repeat for cleavage. In type II and type III systems secondary trimming is performed at either the 5' or 3' end to produce mature crRNAs. Mature crRNAs associate with Cas proteins to form interference complexes. In type I and type II systems, base-pairing between the crRNA and the PAM causes degradation of invading DNA. Type III systems do not require a PAM for successful degradation and in type III-A systems base-pairing occurs between the crRNA and mRNA rather than the DNA, targeted by type III-B systems.

STATEMENTS OF INVENTION

[0013] The invention is as claimed in claims 1 to 15, including the following statements:

A nucleic acid vector comprising an engineered host modifying (HM) CRISPR array for use in treating or preventing, in a human or animal subject, a condition or disease mediated by bacterial host cells by altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria, the second bacteria comprising host cells, wherein the mixed population is comprised by the subject,
wherein the second sub-population consists of a different bacterial species to that of the first sub-population,

wherein each host cell has a host target sequence and at least one nucleic acid sequence encoding a Cas nuclease endogenous to the host cell,
wherein the engineered host modifying (HM) CRISPR array comprises a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to the host target sequence;

wherein the vector can transform the host cell,
wherein the HM-crRNA guides the Cas nuclease to the target to modify the target sequence in the host cell, whereby the host cell is killed or host cell growth is reduced,
wherein the nucleic acid vector is a plasmid vector comprised by a carrier bacterium, or the vector is a phage vector.

[0014] A nucleic acid vector comprising an engineered host modifying (HM) CRISPR array and a nucleic acid sequence encoding a Cas nuclease for use in treating or preventing in a human or animal subject a condition or disease mediated by bacterial host cells by altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria, the second bacteria comprising host cells, wherein the mixed population is comprised by the subject,

wherein the second sub-population consists of a different bacterial species to that of the first sub-population,
wherein each host cell has a host target sequence,
wherein the vector can transform the host cell,
wherein the engineered host modifying (HM) CRISPR array comprises a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to the host target sequence,
wherein the HM-crRNA guides the Cas nuclease to the target to modify the target sequence in the host cell, whereby the host cell is killed or host cell growth is reduced,
wherein the nucleic acid vector is a plasmid vector comprised by a carrier bacterium, or the vector is a phage vector.

[0015] Use of a nucleic acid vector comprising an engineered host modifying (HM) CRISPR/Cas array for altering the relative ratio of sub-populations of first and second bacteria in an ex vivo mixed population of bacteria, the second bacteria comprising host cells,

wherein the second sub-population consists of a different bacterial species to that of the first sub-population,

wherein each host cell has a host target sequence and at least one nucleic acid sequence encoding a Cas nuclease endogenous to the host cell,
wherein the engineered host modifying (HM) CRISPR array comprises a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to the host target sequence;

wherein the vector can transform the host cell,
wherein the HM-crRNA guides the Cas nuclease to the target to modify the target sequence in the host cell, whereby the host cell is killed or host cell growth is reduced,
wherein the nucleic acid vector is a plasmid vector comprised by a carrier bacterium, or the vector is a phage vector.

[0016] Use of a nucleic acid vector comprising an engineered host modifying (HM) CRISPR/Cas array and a nucleic acid sequence encoding a Cas nuclease for altering the relative ratio of sub-populations of first and second bacteria in an ex vivo mixed population of bacteria, the second bacteria comprising host cells,

wherein the mixed population is comprised by the subject,
wherein the second sub-population consists of a different bacterial species to that of the first sub-population,

wherein each host cell has a host target sequence,
wherein the vector can transform the host cell,
wherein the engineered host modifying (HM) CRISPR array comprises a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to the host target sequence,
wherein the HM-crRNA guides the Cas nuclease to the target to modify the target sequence in the host cell, whereby the host cell is killed or host cell growth is reduced,
wherein the nucleic acid vector is a plasmid vector comprised by a carrier bacterium, or the vector is a phage vector.

[0017] Any part of the description not within the scope of claims 1-15 and these statements does not form a part of the invention but serves in the understanding of the claimed invention.

Disclosed configurations

[0018] Disclosed herein are:

First configuration

[0019] The inventors believe that they have demonstrated for the first time inhibition of population growth of a specific bacterial strain in a mixed consortium of bacteria that naturally occur together in microbiota (human, animal or environmental microbiota) with one or more of the following features:-Population growth inhibition by

- targeting wild-type cells;
- harnessing of wild-type endogenous Cas nuclease activity;
- targeting essential and antibiotic resistance genes;
- wherein the targets are wild-type sequences.

[0020] The inventors have demonstrated this in a mixed bacterial population with the following features:-

- targeting bacterial growth inhibition in a mixed population of human microbiota (such as gut microbiota) species;
- wherein the population comprises three different species;
- comprising selective killing of one of those species and sparing cells of the other species;
- targeting cell growth inhibition in the presence of a phylogenetically-close other species, which is spared such inhbition;
- targeting cell growth inhibition in a mixed population comprising target *Firmicutes* species and non-firmicutes species;
- targeting cell growth inhibition of a specific *Firmicutes* strain whilst sparing a different *Firmicutes* species in a mixed population;
- targeting cell growth inhibition of a specific gram positive bacterial strain whilst sparing a different gram positive bacterial species in a mixed population;
- targeting a pathogenic (in humans) bacterial species whilst sparing a commensul human gut bacterial species;
- targeting a pathogenic bacterial species whilst sparing a priobiotic human gut bacterial species;
- targeting cell growth inhibition in a mixed bacterial population on a surface;
- achieving at least a 10-fold growth inhibition of a specific bacterial species alone or when mixed with a plurality of other bacterial species in a consortium; and
- achieving at least a 10-fold growth inhibition of two different strains of a specific bacterial species.

[0021] The ability to harness endogenous Cas activity in wild-type cells is very useful for *in situ* treatment of host cell infections in organisms (humans and animals, for example) and the environment. Treatment of wild-type (ie, non-engineered or pre-manipulated) bacterial populations, such as human, animal or plant microbiota can also be addressed using the invention. The ability to effect selective growth inhibition in a mixed population is useful for addressing bacterial populations, such as human, animal or plant microbiota, or for addressing environmental microbiomes. This feature is also useful for producing medicaments (eg, bacterial cell transplants for administration to a human or animal subject for any treatment or prevention disclosed herein; or for producing a herbicide or insecticide composition comprising the product bacterial population of the disclosure), wherein the selective killing can be used to selectively alter the ratio of different bacteria in a mixed population to produce an altered bacterial population which is the medicament, herbicide or insecticide; or from which the medicament, herbicide or insecitcide is produced. For example, the medicament can be intranasally transplanted into a human or animal recipient to effect such treatment or prevention.

[0022] In the worked Example below, growth inhibition was addressed in a bacterial population (a gram positive *Firmicutes* population) on a solid surface. A >10-fold population growth inhibition was achieved. Targeting was directed to an antibiotic resistance gene. The invention will be useful in inhibiting the growth of antibiotic-resistant bacteria, wherein the target sequence is a sequence of an antibiotic resistance gene. In an example, co-administration of the engineered nucleotide sequence with the antibiotic may be effective. This may provide more complete treatment or prevention of host cell infection in human or animal subjects and/or enable the reduction of therapeutically-effective antibiotic dose for administration to a human or animal. This is useful in view of the increasing worry regarding over-administration of antibiotics and the development of resistance in human and animal populations. The invention also finds application *ex vivo* and *in vitro* for treating an industrial or medical fluid, surface, apparatus or container (eg, for food, consumer goods, cosmetics, personal healthcare product, petroleum or oil production); or for treating a waterway, water, a beverage, a foodstuff or a cosmetic, wherein the host cell(s) are comprised by or on the fluid, surface, apparatus, container, waterway, water, beverage, foodstuff or cosmetic. The invention finds application also in control of corrosion, biofilms and biofouling. The first configuration thus provides the following concepts:-

[0023] Use of a host modifying (HM) CRISPR/Cas system for altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria, the second bacteria comprising host cells, for each host cell the system comprising components according to (i) to (iv):-

(i) at least one nucleic acid sequence encoding a Cas nuclease;
(ii) a host cell target sequence and an engineered host modifying (HM) CRISPR array comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that hybridises to the host cell target sequence to guide said Cas to the target in the host cell to modify the target sequence;
(iii) an optional tracrRNA sequence or a DNA sequence expressing a tracrRNA sequence;
(iv) wherein said components of the system are split between the host cell and at least one nucleic acid vector that transforms the host cell, whereby the HM-crRNA guides Cas to the target to modify the host CRISPR/Cas system in the host cell; and

wherein the target sequence is modified by the Cas whereby the host cell is killed or host cell growth is reduced.

[0024] A host modifying (HM) CRISPR/Cas system for the use of claim 1 for modifying a target nucleotide sequence of a bacterial host cell, the system comprising components according to (i) to (iv):-

(i) at least one nucleic acid sequence encoding a Cas nuclease;
(ii) a host cell target sequence and an engineered host modifying (HM) CRISPR array comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to the host target sequence to guide said Cas to the target in the host cell to modify the target sequence;
(iii) an optional tracrRNA sequence or a DNA sequence for expressing a tracrRNA sequence;
(iv) wherein said components of the system are split between the host cell and at least one nucleic acid vector that can transform the host cell, whereby the HM-crRNA guides Cas to the target to modify the host CRISPR/Cas system in the host cell.

[0025] This is exemplified by the worked Examples herein where we show selective host cell growth inhibition by at least 10-fold in a mixed and non-mixed cell population. The mixture simulates a combination of species and strains found in human microbiota.

[0026] Use of wild-type endogenous Cas nuclease activity of a bacterial host cell population to inhibit growth of the population, wherein each host cell has an endogenous CRISPR/Cas system having wild-type Cas nuclease activity, the use comprising transforming host cells of the population, wherein each transformed host cell is transformed with an engineered nucleotide sequence for providing host modifying (HM) cRNA or guide RNA (gRNA) in the host cell, the HM-cRNA or gRNA comprising a sequence that is capable of hybridising to a host cell target protospacer sequence for guiding endogenous Cas to the target, wherein the cRNA or gRNA is cognate to an endogenous Cas nuclease of the host cell that has said wild-type nuclease activity and following transformation of the host cells growth of the population is inhibited.

[0027] Use (optionally the use is according to the use of the immediately preceding paragraph above) of a host modifying (HM) CRISPR/Cas system for killing or reducing the growth of bacterial host cells, for each host cell the system comprising components according to (i) to (iv):-

(i) at least one nucleic acid sequence encoding a Cas nuclease;
(ii) an engineered host modifying (HM) CRISPR array comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that hybridises to a host cell target sequence to guide said Cas to the target in the host cell to modify the target sequence;
(iii) an optional tracrRNA sequence or a DNA sequence expressing a tracrRNA sequence;
(iv) wherein said components of the system are split between the host cell and at least one nucleic acid vector that transforms the host cell, whereby the HM-crRNA guides Cas to the target to modify the target sequence in the host cell;

[0028] Wherein the Cas nuclease is endogenous to the host cell; and wherein the target sequence is modified by the Cas whereby the host cell is killed or host cell growth is reduced.

[0029] Thus, the HM-cRNA is capable of hybridising to the host cell target sequence to guide said Cas to the target in the host cell to modify the target sequence.

[0030] In an alternative, HM-crRNA and tracrRNA are comprised by a single guide RNA (gRNA).

[0031] By harnessing endogenous Cas nuclease, embodiments of the invention use endogenous Cas nuclease activity (ie, without the need for prior genetic modification of the host cell to activate or enhance the nuclease activity). Thus, in an example, the Cas nuclease is encoded by a wild-type gene of the host cell. In an example, the nuclease is active to achive the cell killing or growth inhibition without inactivation of an endogenous Cas nuclease (or Cas nuclease gene)

repressor in the host cell. Thus, the invention can address wild-type bacterial populations without the need for prior manipulation to bring about effective Cas-mediated cell killing or growth reduction. Thus, the population can be exposed to the cRNA when the population is in its wild-type environment (such as a waterway or comprised by a human or animal microbiome).

[0032] In an example, the first bacteria are *Bacteroidetes* (eg, *Bacteroides*) cells. In an example, the second bacteria are *Firmicutes* cells. The method is, for example, used to alter the ratios in a gut microbiota population (eg, *ex vivo* or *in vivo),* which is for example for treating or preventing increased body mass or obesity (eg, wherein the first bacteria are *Firmicutes* cells).

[0033] The first configuration also provides: A method of altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria comprising said sub-populations, wherein the first bacteria are host cells (eg, *Bacteroidetes* cells) infected by a phage and the second bacteria are not infected by said phage (or not *Bacteroidetes* bacteria), the method comprising combining the mixed population with a plurality of vectors in one or more steps for introduction of vector nucleic acid into host cells and allowing bacterial growth in the mixed population, wherein the relative ratios of said first and second bacteria is altered; wherein each vector comprises an engineered phage-modifying (PM) CRISPR array for introduction into a phage-infected host cell for modifying a target nucleotide sequence of said phage in the cell, (a) wherein the PM-CRISPR array comprises one or more sequences for expression of a PM-crRNA and a promoter for transcription of the sequence(s) in a phage-infected host cell; and (b) wherein the PM-crRNA is capable of hybridising to the phage target sequence to guide Cas (eg, a Cas nuclease) in the infected host cell to modify the target sequence.

Second configuration:-

[0034] A host modifying (HM) CRISPR/Cas system for modifying a target nucleotide sequence of a host cell (eg, for the use of the first configuration), the system comprising components according to (i) to (iv):

(i) at least one nucleic acid sequence encoding a Cas nuclease;
(ii) an engineered host modifying (HM) CRISPR array comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to a host target sequence to guide said Cas to the target in the host cell to modify the target sequence;
(iii) an optional tracrRNA sequence or a DNA sequence for expressing a tracrRNA sequence;
(iv) wherein said components of the system are split between the host cell and at least one nucleic acid vector that can transform the host cell, whereby the HM-crRNA guides Cas to the target to modify the target sequence in the host cell;

wherein optionally component (i) is endogenous to the host cell.

[0035] The second configuration also provides: An engineered phage-modifying (PM) CRISPR array for use in the method of the first configuration for modifying the genome of said phage,

(a) wherein the PM-CRISPR array comprises one or more sequences for expression of a PM-crRNA and a promoter for transcription of the sequence(s) in a phage-infected host cell; and
(b) wherein the PM-crRNA is capable of hybridising to a phage genome target sequence to guide Cas (eg, a Cas nuclease) in the infected host cell to modify the target sequence.

[0036] In an example, the phage is a *Bacteroidetes* (eg, *Bacteroides)* phage, eg, crAssphage.

[0037] In an example, the array comprises CRISPR repeats that are functional with a host cell CRISPR/Cas system. This is beneficial to increase selectivity of the array for the desired cell in a bacterial mixture. This also simplifies production of the array and vectors containing the array of the invention as it may not be necessary to include bulky nucleotide sequenes encoding one or more Cas proteins (and/or tracrRNA) required for functioning of the array in the host cell. In an alternative, the array is provided with a cognate Cas9-encoding sequence and optionally a cognate tracrRNA-encoding sequence.

Third configuration:-

[0038] An engineered nucleic acid vector for modifying a bacterial host cell comprising an endogenous CRISPR/Cas system, the vector

(a) comprising nucleic acid sequences for expressing a plurality of different crRNAs (eg, single guide RNAs, ie, gRNAs) for use in a use of the invention or a CRISPR/Cas system according to the disclosure; and

(b) lacking a nucleic acid sequence encoding a Cas nuclease, wherein a first of said crRNAs is capable of hybridising to a first nucleic acid sequence in said host cell; and a second of said crRNAs is capable of hybridising to a second nucleic acid sequence in said host cell, wherein said second sequence is different from said first sequence; and
(c) the first sequence is comprised by an antibiotic resistance gene (or RNA thereof) and the second sequence is comprised by an antibiotic resistance gene (or RNA thereof); optionally wherein the genes are different;
(d) the first sequence is comprised by an antibiotic resistance gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof);
(e) the first sequence is comprised by an essential gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof); or
(f) the first sequence is comprised by a virulence gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof).

[0039] The third configuration also provides: A nucleic acid vector for use in the method of the invention, the vector comprising a CRISPR array of the invention.

Fourth configuration:-

[0040] A nucleic acid vector (eg, a plasmid, virus, phage or phagemid) comprising an engineered CRISPR array for modifying a target sequence of the genome of a host bacterial cell (eg, pathogenic bacterial cell, such as described above) or the genome of a virus (eg, phage) in a host cell,

(a) wherein the CRISPR array comprises one or more sequences for expression of a crRNA (eg, provided as a gRNA) and a promoter for transcription of the sequence(s) in the host cell;
(b) wherein the crRNA is capable of hybridising to the target sequence to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence;
(c) wherein the array is comprised by a transposon that is capable of horizontal transfer between first and second bacterial cells of different species.

Fifth configuration:-

[0041] An engineered CRISPR nucleic acid vector comprising or consisting of a mobile genetic element (MGE), wherein the MGE comprises an origin of transfer (*ori*T) and a CRISPR array for modifying a target sequence of the genome of a host cell (eg, pathogenic bacterial cell) or the genome of a virus (eg, prophage) in a host cell,

(a) wherein the CRISPR array comprises one or more sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in the host cell;
(b) wherein the crRNA is capable of hybridising to the target sequence to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence;
(c) wherein the vector is capable of transfer between (i) first and second nucleic acid positions of a first host cell, wherein each position is a position on a chromosome or a plasmid and the target sequence is comprised by the host cell, or (ii) first and second host cells, wherein the target sequence is comprised by the first and/or second host cell.

Sixth configuration:-

[0042] A method of controlling microbiologically influenced corrosion (MIC) or biofouling of a substrate in an industrial or domestic system, wherein a surface of the substrate is in contact with a population of first host cells of a first microbial species that mediates MIC or biofouling of the substrate, the method comprising

(i) contacting the population with a plurality of vectors that are capable of transforming or transducing the cells, each vector comprising a CRISPR array whereby CRISPR arrays are introduced into the host cells, wherein

(a) each CRISPR array comprises one or more nucleotide sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in a host cell; and
(b) each crRNA is capable of hybridising to a target sequence of a host cell to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence (eg, to cut the target sequence); the target sequence being a gene sequence for mediating host cell viability; and

(ii) allowing expression of said cRNAs in the presence of Cas in host cells, thereby modifying target sequences in host cells, resulting in reduction of host cell viability and control of MIC or biofouling of said substrate.

[0043] In another embodiment, there is provided:-

[0044] A method of controlling microbiologically influenced corrosion (MIC) or biofouling of a substrate comprised by a crude oil, gas or petrochemicals recovery, processing, storage or transportation equipment, wherein a surface of the substrate is in contact with a population of first host cells, wherein the first host cells are sulphur- or sulphate-reducing bacteria (SRB), extracellular polymeric substance-producing bacteria (EPSB), acid-producing bacteria (APB), sulphur- or sulphide-oxidizing bacteria (SOB), iron-oxidising bacteria (IOB), manganese-oxidising bacteria (MOB), ammonia producing bacteria (AmPB) or acetate producing bacteria (AcPB) of a first species that mediates MIC or biofouling of the substrate, wherein the surface and cell population are in contact with a liquid selected from sea water, fresh water, a fracking liquid or liquid in a well, the method comprising

(i) contacting the cell population with vectors by mixing the liquid with a plurality of vectors that are capable of transforming or transducing first host cells, each vector comprising a CRISPR array whereby CRISPR arrays are introduced into the host cells, wherein

(a) each CRISPR array comprises one or more sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in a host cell;
(b) each crRNA is capable of hybridising to a target sequence of a host cell to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence (eg, to cut the target sequence); the target sequence being a gene sequence for mediating host cell viability;
(c) wherein each sequence of (a) comprises a sequence R1-S1-R1' for expression and production of the respective crRNA in a first host cell, wherein R1 is a first CRISPR repeat, R1' is a second CRISPR repeat, and R1 or R1' is optional; and S1 is a first CRISPR spacer that comprises or consists of a nucleotide sequence that is 80% or more identical to a target sequence of a said first host cell and

(ii) allowing expression of said cRNAs in the presence of Cas in host cells, thereby modifying target sequences in host cells, resulting in reduction of host cell viability and control of MIC or biofouling of said substrate.

Other embodiments provide:-

[0045] A vector for use in the method, wherein the first cells are sulphate reducing bacteria (SRB) cells, eg, Desulfovibrio or Desulfotomaculum cells, the vector comprising one or more CRISPR arrays for targeting the SRB, wherein each array is as defined in (a)-(c).

[0046] In another embodiment, there is provided: A method of controlling microbial biofouling of a fluid in an industrial or domestic system, wherein the fluid comprises a population of first host cells of a first microbial species that mediates said biofouling, the method comprising

(i) contacting the population with a plurality of vectors that are capable of transforming or transducing the cells, each vector comprising a CRISPR array whereby CRISPR arrays are introduced into the host cells, wherein

(a) each CRISPR array comprises one or more nucleotide sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in a host cell; and
(b) each crRNA is capable of hybridising to a target sequence of a host cell to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence (eg, to cut the target sequence); the target sequence being a gene sequence for mediating host cell viability; and

(ii) allowing expression of said cRNAs in the presence of Cas in host cells, thereby modifying target sequences in host cells, resulting in reduction of host cell viability and control of said biofouling.

[0047] For example, there is provided: A method of controlling bacterial biofouling in ballast water of a ship or boat, wherein the water comprises a population of first host cells of a first microbial species that mediates said biofouling, the method comprising

(i) contacting the population with a plurality of vectors that are capable of transforming or transducing the cells, each vector comprising a CRISPR array whereby CRISPR arrays are introduced into the host cells, wherein

(a) each CRISPR array comprises one or more sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in a host cell; and
(b) each crRNA is capable of hybridising to a target sequence of a host cell to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence (eg, to cut the target sequence); the target sequence being a gene sequence for mediating host cell viability; and

(ii) allowing expression of said cRNAs in the presence of Cas in host cells, thereby modifying target sequences in host cells, resulting in reduction of host cell viability and control of said biofouling.

[0048] Other embodiments provide: Ballast sea water (for example, a sample of sea water or sea water in a container) comprising CRISPR arrays, wherein the ballast water is obtained or obtainable by the method. A ship, boat, sea container or rig comprising the ballast sea water. A vector for use in the method, wherein the first cells are Cholera (eg, vibrio, eg, O1 or 0139), E coli or Enterococci sp cells, the vector comprising one or more CRISPR arrays for targeting the cells, wherein each array is as defined in (a) and (b) of the method.

[0049] The invention also provides vectors and CRISPR arrays suitable for use in this sixth configuration or for other applications, such as for medical use, or for food or beverage treatement. To this end, there is provided: A vector comprising a CRISPR array for introduction into a bacterial host cell, wherein the bacterium is capable of water-borne transmission, wherein

(a) the CRISPR array comprises a sequence for expression of a crRNA and a promoter for transcription of the sequence in a said host cell;
(b) the crRNA is capable of hybridising to a host cell target sequence to guide a Cas (eg, a Cas nuclease) in the host cell to modify the target sequence (eg, to cut the target sequence); the target sequence being a nucleotide sequence for mediating host cell viability;
(c) wherein the sequence of (a) comprises a sequence R1-S1-R1' for expression and production of the crRNA, wherein R1 is a first CRISPR repeat, R1' is a second CRISPR repeat, and R1 or R1' is optional; and S1 is a first CRISPR spacer that comprises or consists of a nucleotide sequence that is 80% or more identical to the host cell target sequence.

[0050] Also provided are: A water or food treatment composition comprising a plurality of such vectors. A medicament for treatment or prevention of a bacterial infection (eg, a *Vibrio cholerae* infection) in a human, the medicament comprising a plurality of such vectors. The disclosure also provides bacterial populations, compositions, foodstuffs and beverages. For example, the foodstuff or beverage is a dairy product.

Seventh configuration:-

[0051] In a first aspect:-
A method of modifying an expressible gene encoding a first Cas, the method comprising

(a) combining a guide RNA (gRNA1) with the Cas gene in the presence of first Cas that is expressed from said gene; and
(b) allowing gRNA1 to hybridise to a sequence of said Cas gene (eg, a promoter or a first Cas-encoding DNA sequence thereof) and to guide first Cas to the gene, whereby the Cas modifies the Cas gene.

[0052] A first nucleic acid vector or combination of vectors, eg, for use in the method, wherein

(a) the first vector or a vector of said combination comprises an expressible nucleotide sequence that encodes a guide RNA (gRNA1, eg, a single gRNA) that is complementary to a predetermined protospacer sequence (PS1) for guiding a first Cas to modify PS1 at a first site (CS1), wherein PS1 is adjacent a PAM (P1) that is cognate to the first Cas; or the expressible sequence encodes a crRNA that forms gRNA1 with a tracrRNA; and

(b) PS1 and P1 are sequences of an expressible first Cas-encoding gene and PS1 is capable of being modified at CS1 by the first Cas.

[0053] These aspects of the invention are useful for regulating Cas activity, eg, in a cell or *in vitro.* The invention involves targeting a Cas-encoding gene to restrict Cas activity, which is advantageous for temporal regulation of Cas. The invention may also be useful in settings where increased stringency of Cas activity is desirable, eg, to reduce the chances for off-target Cas cutting in when modifying the genome of a cell. Applications are, for example, in modifying

human, animal or plant cells where off-target effects should be minimised or avoided, eg, for gene therapy or gene targeting of the cell or a tissue or an organism comprising the cell. For example, very high stringency is required when using Cas modification to make desired changes in a human cell (eg, iPS cell) that is to be administered to a patient for gene therapy or for treating or preventing a disease or condition in the human. The disclosure provides these applications as part of the methods and products of the invention.

[0054] The invention also addresses the problem of restricted insert capacity in vectors, particularly in viral vectors.

Eighth configuration:-

[0055] A nucleic acid vector comprising more than 1.4kb of exogenous DNA sequence encoding components of a CRISPR/Cas system, wherein the sequence comprises an engineered array or engineered sequence (optionally as described herein) for expressing one or more HM- or PM-crRNAs or gRNAs in host cells (any cell herein, eg, human, anial or bacterial or archael host cells), wherein the array or engineered sequence does not comprise a nucleotide sequence encoding a Cas nuclease that is cognate to the cRNA(s) or gRNA(s); optionally wherein at least 2, 3 or 4 cRNAs or gRNAs are encoded by the exogenous DNA.

[0056] A nucleic acid vector comprising more than 1.4kb or more than 4.2kb of exogenous DNA sequence, wherein the exogenous DNA encodes one or more components of a CRISPR/Cas system and comprises an engineered array or sequence (eg, any such one described herein) for expressing one or more HM-crRNAs or gRNAs in host cells, wherein the exogenous sequence is devoid of a nucleotide sequence encoding a Cas nuclease that is cognate to the cRNA(s) or gRNA(s); optionally wherein at least 2 different cRNAs or gRNAs are encoded by the exogenous DNA.

[0057] Herein in any configurations, for example the cRNA(s) are provided by one or more single guide RNAs (gRNAs), and in this case "CRISPR array" may refer to one or more expressible nucleotide sequences that encode said gRNA(s). Thus, the sequences are capable of being expressed in host cell(s) for expressing the gRNA(s) inside the cell(s).

[0058] The disclosure is mainly described in terms of bacteria, but it is also applicable *mutatis mutandis* to archaea.

BRIEF DESCRIPTION OF THE FIGURES

[0059]

FIGURE 1: Xylose inducible system.

FIGURE 2: ST1-CRISPR array.

FIGURE 3: Spot assay on TH-agar of the strains used in this work. All strains were grown on TH-agar at 37°C for 20 hours. Serial dilutions of overnight cultures were done in duplicate for *E.coli, L Lactis* and *S.mutans,* and triplicate for both strains of *S. thermophilus* in order to count individual colonies.

FIGURE 4: Selective growth of *S. thermophilus, S. mutans, L. lactis* and *E. coli* under different culture conditions. Tetracycline cannot be used to selectively grown *S. thermophilus LMD-9. However,* 3g l^{-1} of PEA proved to selectively grow S. *thermophilus* LMD-9 while limiting growth of *E. coli.*

FIGURE 5: Construction of two xylose induction cassettes (middle, right) based on the wild type B. *megaterium* operon (left). (Xie *et al.* 2013)

FIGURE 6: Characterization of the xylose inducible cassette in *Streptoccocus thermophilus* LMD-9 with the plasmid pBAV1KT5-XylR-mCherry-PIdha. A clear response in fluorescence can be observed with increasing amount of xylose.

FIGURE 7: Design of CRISPR array in pBAV1KT5-XylR-mCherry-$P_{Idha+XylA}$. The array contains 2 spacer sequences that target *S. thermophilus* genes under an inducible xylose promoter and a tracrRNA under a strong constitutive promoter P_{3A}.

FIGURE 8: Transformation efficiency of *Streptoccocus thermophilus* LMD-9 with the plasmid pBAV1KT5-XylR-CRISPR-$P_{Idh+XylA}$ (left) and with pBAV1KT5-XylR-CRISPR-P_{XylA} (right).

FIGURE 9: A schematic of the xylose-inducible CRISPR device. Upon induction of xylose the CRISPR array targeting both *polIII* and *tetA* on the S. thermophiles LMD-9 genome are expressed. Together with the constitutively expressed tracrRNA a complex is formed with Cas9. This complex will introduce a double stranded break in the *tetA* and *polIII*

genes in the S. thermophilus LMD-9 genome resulting in limited cell viability.

FIGURE 10: Growth inhibition of Streptoccocus thermophilus DSM 20617(T) with the plasmid pBAV1KT5-XylR-CRISPR-PXylA (left) or pBAV1KT5-XylR-CRISPR-PIdha+XylA (right). Not induced (upper panel) and induced (lower panel). Picture taken after 63H of incubation. Colony counts in bottom left corner (top row: >1000, >1000, bottom row: 336, 113).

FIGURE 11: Maximum-likelihood phylogenetic tree of 16S sequences from *S. thermophilus, L. lactis* and *E. coli.*

FIGURE 12: Shows the selective S thermophilus growth inhibition in a co-culture of E. coli, L. lactis and S. thermophiles harboring either the pBAV1KT5-XylR-CRISPR-PxylA or the pBAV1KT5-XylR-CRISPR-PIdhA+XylA plasmid. No growth difference is observed between E. coli harboring the pBAV1KT5-XylR-CRISPR-PxylA or the pBAV1KT5-XylR-CRISPR-PIdhA+XylA plasmid (middle column). However, S. thermophiles (selectively grown on TH agar supplemented with 2.5 gl-1 PEA, last column) shows a decrease in transformation efficiency between the pBAV1KT5-XylR-CRISPR-PxylA (strong) or the pBAV1KT5-XylR-CRISPR-PldhA +XylA (weak) plasmid as we expected. We thus demonstrated a selective growth inhibition of the target S thermophilus sub-population in the mixed population of cells. Colony counts in bottom left corner (top row: >1000, >1000, 68, bottom row: >1000, >1000, 32).

DETAILED DESCRIPTION

INHIBITING MICROBIAL POPULATION GROWTH & ALTERING MICROBIAL RATIOS

[0060] The disclosure relates to methods, uses, systems, arrays, cRNAs, gRNAs and vectors for inhibiting bacterial population growth or altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria, eg, for altering human or animal microbiomes, such as for the alteration of the proportion of *Bacteroidetes* (eg, *Bacteroides), Firmicutes* and/or gram positive or negative bacteria in microbiota of a human. See, for example, the first to third configurations described herein. The invention involves modifying one or more target nucleotide sequences of a host bacterial cell, eg, a *Bacteroidetes* cell or *Firmicutes* cell.

[0061] There have been a number of studies pointing out that the respective levels of the two main intestinal phyla, the *Bacteroidetes* and the *Firmicutes,* are linked to obesity, both in humans and in germfree mice. The authors of the studies deduce that carbohydrate metabolism is the important factor. They observe that the microbiota of obese individuals are more heavily enriched with bacteria of the phylum *Firmicutes* and less with *Bacteroidetes,* and they surmise that this bacterial mix may be more efficient at extracting energy from a given diet than the microbiota of lean individuals (which have the opposite proportions). In some studies, they found that the relative abundance of *Bacteroidetes* increases as obese individuals lose weight and, further, that when the microbiota of obese mice are transferred to germfree mice, these mice gain more fat than a control group that received microbiota from lean mice. See, eg, Turnbaugh, P. J., R. E. Ley, M. A. Mahowald, V. Magrini, E. R. Mardis, and J. I. Gordon. 2006, "An obesity-associated gut microbiome with increased capacity for energy harvest", Nature 444:1027-1131.

CONCEPTS

[0062] Disclosed herein are the following concepts involving a host cell target:-
In a first concept, there is provided the use of a host modifying (HM) CRISPR/Cas system for killing or reducing the growth of bacterial host cells, for each host cell the system comprising components according to (i) to (iv):-

(i) at least one nucleic acid sequence encoding a Cas nuclease;
(ii) an engineered host modifying (HM) CRISPR array comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that hybridises to a host cell target sequence to guide said Cas to the target in the host cell to modify the target sequence;
(iii) an optional tracrRNA sequence or a DNA sequence expressing a tracrRNA sequence;
(iv) wherein said components of the system are split between the host cell and at least one nucleic acid vector that transforms the host cell, whereby the HM-crRNA guides Cas to the target to modify the target sequence in the host cell;

wherein the Cas nuclease is endogenous to the host cell; and wherein the target sequence is modified by the Cas whereby the host cell is killed or host cell growth is reduced.

[0063] The first concept alternatively provides:
Use of a host modifying (HM) CRISPR/Cas system for altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria, the second bacteria comprising host cells, for each host cell the system

comprising components according to (i) to (iv):-

(i) at least one nucleic acid sequence encoding a Cas nuclease;
(ii) an engineered host modifying (HM) CRISPR array comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that hybridises to a host cell target sequence to guide said Cas to the target in the host cell to modify the target sequence;
(iii) an optional tracrRNA sequence or a DNA sequence expressing a tracrRNA sequence;
(iv) wherein said components of the system are split between the host cell and at least one nucleic acid vector that transforms the host cell, whereby the HM-crRNA guides Cas to the target to modify the target sequence in the host cell;

wherein optionally the Cas nuclease is endogenous to the host cell; and
wherein the target sequence is modified by the Cas whereby the host cell is killed or host cell growth is reduced.

[0064] The first concept also provides: A method of altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria, the second bacteria comprising host cells, and the method comprising combining the mixed population with of a host modifying (HM) CRISPR/Cas system whereby second bacteria host cells are killed or the growth of said cells is reduced thereby altering said ratio,
wherein for each host cell the system comprises components according to (i) to (iv):-

(i) at least one nucleic acid sequence encoding a Cas nuclease;
(ii) an engineered host modifying (HM) CRISPR array comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that hybridises to a host cell target sequence to guide said Cas to the target in the host cell to modify the target sequence;
(iii) an optional tracrRNA sequence or a DNA sequence expressing a tracrRNA sequence;
(iv) wherein said components of the system are split between the host cell and at least one nucleic acid vector that transforms the host cell, whereby the HM-crRNA guides Cas to the target to modify the target sequence in the host cell;

wherein optionally the Cas nuclease is endogenous to the host cell; and
wherein the target sequence is modified by the Cas whereby the host cell is killed or host cell growth is reduced.

[0065] The first concept also provides:-
Use of a host modifying (HM) CRISPR/Cas system for altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria, the second bacteria comprising a plurality of host cells each comprising a target protospacer sequence, for each host cell the system comprising components (ii) and (iii) defined above, the system further comprising at least one nucleic acid sequence encoding a Cas nuclease; wherein said component (ii) and said Cas-encoding sequence are compried by at least one nucleic acid vector that transforms the host cell, whereby the HM-crRNA encoded by (i) guides Cas to the target to modify the target sequence in the host cell;
wherein the Cas nuclease is endogenous to the host cell; and wherein the target sequence is modified by the Cas whereby the host cell is killed or host cell growth is reduced.

[0066] In an embodiment, the growth of first bacteria is not inhibited; or the growth inhibition of said host cells is at least 2x, 3x, 4x, 5x, 6x, 7x, 8x, 9x, 10x, 50x, 100x or 1000x the growth inhibition of the first cells. The growth inhibition can be calculated as a fold-inhibition or as a percentage inhibition (as described herein). In another example, inhibition is measured in a culture sample by a spectrophotometer, wherein light absorbance (eg, at OD_{600}) is determined at the start and end of a predetermined crRNA/gRNA treatment period (see the description of such a period herein when determining inhibition by fold or percentage). In an example, the increase in absorbance (comparing the absorbance at the beginning of the predetermined period with absorbance at the end of that period) for the host cell sampe is less than for the control sample (which has not been exposed to said cRNA or gRNA), eg, the increase for the former is at least 10, 100, 1000, 10000 or 100000 times lower than for the latter (eg, determined as $OD_{600)}$. In an example, the determination of growth inhibition (ie, the end of the predermined period) is made at the mid-exponential growth phase of each sample (eg, 6-7 hours after the start of the predetermined period).

[0067] In an example, the host cells are comprised by a microbiota population comprised by an organism or environment (eg, a waterway microbiota, water microbiota, human or animal gut microbiota, human or animal oral cavity microbiota, human or animal vaginal microbiota, human or animal skin or hair microbiota or human or animal armpit microbiota), the population comprising first bacteria that are symbiotic or commensal with the organism or environment and second bacteria comprising said host cells, wherein the host cells are detrimental (eg, pathogenic)to the organism or environment. In an embodiment, the population is *ex vivo.*

[0068] The ratio of the first bacteria sub-population to the second bacteria sub-population is increased. The first concept also provides a use for inhibiting host cell growth as described further below.

[0069] A second concept relates to a host modifying (HM) CRISPR/Cas system for modifying a target nucleotide sequence of a host cell (eg, for the use of the first concept), the system comprising components according to (i) to (iv):-

(i) at least one nucleic acid sequence encoding a Cas nuclease;
(ii) an engineered host modifying (HM) CRISPR array comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to a host target sequence to guide said Cas to the target in the host cell to modify the target sequence;
(iii) an optional tracrRNA sequence or a DNA sequence for expressing a tracrRNA sequence;
(iv) wherein said components of the system are split between the host cell and at least one nucleic acid vector that can transform the host cell, whereby the HM-crRNA guides Cas to the target to modify the target sequence in the host cell;

wherein optionally component (i) is endogenous to the host cell.

[0070] In an alternative, HM-crRNA and tracrRNA are comprised by a single guide RNA (gRNA).

[0071] By harnessing endogenous Cas nuclease, embodiments of the invention use endogenous Cas nuclease activity (ie, without the need for prior genetic modification of the host cell to activate or enhance the nuclease activity). Thus, in an example, the Cas nuclease is encoded by a wild-type gene of the host cell. In an example, the nuclease is active to achive the cell killing or growth reduction without inhibition of an endogenous Cas nuclease (or Cas nuclease gene) repressor in the host cell. Thus, the invention can address wild-type bacterial populations without the need for prior manipulation to make bring about effective Cas-mediated cell killing or growth reduction. Thus, the population can be exposed to the cRNA when the population is in its wild-type environment (such as a waterway or comprised by a human or animal microbiome).

[0072] In an example, the second bacteria are *Bacteroidetes* (eg, *Bacteroides*) cells. In an example, the second bacteria are *Firmicutes* cells. The use, system or method is, for example, used to alter the ratios in a gut microbiota population (eg, *ex vivo* or *in vivo),* which is for example for treating or preventing increased body mass or obesity (eg, wherein the second bacteria are *Firmicutes* cells).

[0073] In an example, the use, method, system, vector, engineered nucleotide sequence, cRNA or gRNA is for therapeutically or prophylactically rebalancing microbiota of a human or non-human animal comprising the mixed population, eg for treating or preventing obesity, diabetes IBD, a GI tract condition or an oral cavity condition.

[0074] In an example, the microbiota mentioned herein is microbiota of a human or animal microbiome (eg, gut, vaginal, scalp, armpit, skin bloodstream, throat or oral cavity microbiome).

[0075] In an example, the microbiota mentioned herein is an armpit microbiota and the use, method, system, vector, engineered nucleotide sequence, cRNA or gRNA is for preventing or reducing body odour of a human.

[0076] In an example, the host cell population or mixed population is harboured by a beverage or water (eg, a waterway or drinking water) for human consumption.

[0077] In an example, the use, method, system, vector, engineered nucleotide sequence, cRNA or gRNA is for reducing pathogenic infections or for re-balancing gut or oral microbiota eg, for treating or preventing obesity or disease in a human or animal. For example, the use, method, system, vector, engineered nucleotide sequence, cRNA or gRNA is for knocking-down *Clostridium dificile* bacteria in a gut microbiota.

[0078] In an example, the first bacteria are *Bacteroides* bacteria and the second bacteria are *Firmicutes* or pathogenic bacteria, eg, gut bacteria. In an example, the host cells or second bacteria are *Firmicutes* cells, eg, selected from *Streptococcus* (eg, *thermophilus* and/or *pyogenes), Bacillus, Lactobacillus, Listeria, Clostridium, Heliobacterium* and *Staphylococcus* cells. In an example, the mixed populaton contains *Bacteroides* and metronidazole (MTZ)-resistant *C dificile* strain 630 sub-populations, wherein the host cells comprise said *C dificile* cells.

[0079] In an example, the host cell population, mixed population or system is comprised by a composition (eg, a beverage, mouthwash or foodstuff) for administration to a human or non-human animal for populating and rebalancing the gut or oral microbiota thereof.

[0080] In an example, the product of the use or method, or the system, vector, engineered nucleotide sequence, cRNA or gRNA is for administration to a human or non-human animal by mucosal, gut, oral, intranasal, intrarectal, intravaginal, ocular or buccal administration.

[0081] In an example of any configuration herein, the mixed population (prior to combining with the array, gRNA, crRNA or engineered sequence) is a sample of a microiota of a human or animal subject, eg, a gut or any other microbiota disclosed herein or a microbiota of any microbiome disclosed herein. In an example, in this instance the product of the use of the invention is a modified microbiota population that is useful for an treatment or therapy of a human or animal subject, as disclosed herein. A third concept relates to the system of the second concept, wherein the vector or vectors lack a Cas (eg, a Cas9) nuclease-encoding sequence.

[0082] A fourth concept relates to the use, method or system of any preceding concept, wherein each host cell is of a strain or species found in human microbiota, optionally wherein the host cells are mixed with cells of a different strain or species, wherein the different cells are Enterobacteriaceae or bacteria that are probiotic, commensal or symbiotic with humans (eg, in the human gut. In an example, the host cell is a *Firmicutes,* eg, *Streptococcus,* cell.

[0083] A fifth concept relates to the use, method or system of any preceding concept for the alteration of the proportion

of *Bacteroidetes* (eg, *Bacteroides)* bacteria in a mixed bacterial population (eg, in a human, such as in human microbiota).

[0084] A sixth concept relates to the use, method or system of the fifth concept for increasing the relative ratio of *Bacteroidetes versus Firmicutes.*

[0085] A seventh concept relates to the use, method or system of any preceding concept, wherein said Cas nuclease is provided by an endogenous Type II CRISPR/Cas system of the cell.

[0086] An eighth concept relates to the use, method or system of any preceding concept, wherein component (iii) is endogenous to the host cell.

[0087] A ninth concept relates to the use, method or system of any preceding concept, wherein the target sequence is comprised by an antibiotic resistance gene, virulence gene or essential gene of the host cell. A tenth concept relates to the use, method or system of any preceding concept, the array being comprised by an antibiotic composition, wherein the array is in combination with an antibiotic agent. An eleventh concept relates to the use, method or system of any preceding concept, wherein alternatively HM-crRNA and tracrRNA are comprised by a single guide RNA (gRNA), eg provided by the vector.

[0088] A twelfth concept relates to the use, method or system of any preceding concept, wherein the host cell comprises a deoxyribonucleic acid strand with a free end (HM-DNA) encoding a HM-sequence of interest and/or wherein the system comprises a sequence encoding the HM-DNA, wherein the HM-DNA comprises a sequence or sequences that are homologous respectively to a sequence or sequences in or flanking the target sequence for inserting the HM-DNA into the host genome (eg, into a chromosomal or episomal site).

[0089] A thirteenth concept relates to an engineered nucleic acid vector for modifying a bacterial host cell comprising an endogenous CRISPR/Cas system, the vector

(a) comprising nucleic acid sequences for expressing a plurality of different crRNAs (eg, gRNAs) for use in a CRISPR/Cas system, method or use according to any preceding concept; and
(b) optionally lacking a nucleic acid sequence encoding a Cas nuclease,
wherein a first of said crRNAs is capable of hybridising to a first nucleic acid sequence in said host cell; and a second of said crRNAs is capable of hybridising to a second nucleic acid sequence in said host cell, wherein said second sequence is different from said first sequence; and
(c) the first sequence is comprised by an antibiotic resistance gene (or RNA thereof) and the second sequence is comprised by an antibiotic resistance gene (or RNA thereof); optionally wherein the genes are different;
(d) the first sequence is comprised by an antibiotic resistance gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof);
(e) the first sequence is comprised by an essential gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof); or
(f) the first sequence is comprised by a virulence gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof).

[0090] A fourteenth concept relates to the vector of the thirteenth concept inside a host cell comprising one or more Cas that are operable with cRNA (eg, single guide RNA) encoded by the vector.

[0091] A fifthteenth concept relates to the use, method, system or vector of any preceding concept, wherein the HM-CRISPR array comprises multiple copies of the same spacer.

[0092] A sixteenth concept relates to the use, method, system or vector of any preceding concept, wherein the vector(s) comprises a plurality of HM-CRISPR arrays.

[0093] A seventeenth concept relates to the use, method, system or vector of any preceding concept, wherein each vector is a plasmid, cosmid, virus, a virion, phage, phagemid or prophage.

[0094] An eighteenth concept relates to the use, method, system or vector of any preceding concept, wherein the system or vector comprises two, three or more of copies of nucleic acid sequences encoding crRNAs (eg, gRNAs), wherein the copies comprise the same spacer sequence for targeting a host cell sequence (eg, a virulence, resistance or essential gene sequence).

[0095] A nineteenth concept relates to the use, method, system or vector of the eighteenth concept, wherein the copies are split between two or more vector CRISPR arrays.

[0096] A twentieth concept relates to a bacterial host cell comprising a system or vector recited in any preceding concept.

[0097] A twenty-first concept relates to the system, vector or cell of any one of the second to the twentieth concepts in combination with an antibiotic agent (eg, a beta-lactam antibiotic).

[0098] A twenty-second concept relates to the use, method, system, vector or cell of any preceding concept, wherein the or each host cell is a *Staphylococcus, Streptococcus, Pseudomonas, Salmonella, Listeria, E coli, Desulfovibrio* or *Clostridium* host cell. In an example, the or each host cell is a *Firmicutes* cell, eg, a *Staphylococcus, Streptococcus, Listeria* or *Clostridium* cell.

[0099] In an example, each CRISPR array comprises a sequence R1-S1-R1' for expression and production of the

respective crRNA (eg, comprised by a single guide RNA) in the host cell, (i) wherein R1 is a first CRISPR repeat, R1' is a second CRISPR repeat, and R1 or R1' is optional; and (ii) S1 is a first CRISPR spacer that comprises or consists of a nucleotide sequence that is 95% or more identical to said target sequence.

[0100] In an example, R1 and R1' are at least 95% identical respectively to the first and second repeat sequences of a CRISPR array of the second host cell species. In an example, R1 and R1' are at least 95% (eg, 96, 97, 98, 99 or 100%) identical respectively to the first (5'-most) and second (the repeat immediately 3' of the first repeat) repeat sequences of a CRISPR array of said species, eg, of a said host cell of said species. In an example, R1 and R1' are functional with a Type II Cas9 nuclease (eg, a *S thermophilus, S pyogenes* or *S aureus* Cas9) to modify the target in a said host cell.

[0101] An alternative first concept, use of invention provides the following, as demonstrated by the worked experimental Example:

The use of wild-type endogenous Cas nuclease activity of a bacterial host cell population to inhibit growth of the population, wherein each host cell has an endogenous CRISPR/Cas system having wild-type Cas nuclease activity, the use comprising transforming host cells of the population, wherein each transformed host cell is transformed with an engineered nucleotide sequence for providing host modifying (HM) cRNA or guide RNA (gRNA) in the host cell, the HM-cRNA or gRNA comprising a sequence that is capable of hybridising to a host cell target protospacer sequence for guiding endogenous Cas to the target, wherein the cRNA or gRNA is cognate to an endogenous Cas nuclease of the host cell that has said wild-type nuclease activity and following transformation of the host cells growth of the population is inhibited.

[0102] In the worked Example below, inhibition was addressed in a bacterial population (a gram positive *Firmicutes)* on a solid surface. A >10-fold inhibition of host cell population growth was achieved. Targeting was directed to an antibiotic resistance gene and an essential gene. The invention will be useful in inhibiting the growth of antibiotic-resistant bacteria, wherein the target sequence is a sequence of an antibiotic resistance gene. In an example, co-administration of the engineered nucleotide sequence with the antibiotic may be effective. This may provide more complete treatment or prevention of host cell infection in human or animal subjects and/or enable the reduction of therapeutically-effective antibiotic dose for administration to a human or animal. This is useful in view of the increasing worry regarding over-administration of antibiotics and the development of resistance in human and animal populations.

[0103] The demonstration of the invention's ability to inhibit host cell growth on a surface is important and desirable in embodiments where the invention is for treating or preventing diseases or conditions mediated or caused by microbiota as disclosed herein in a human or animal subject. Such microbiota are typically in contact with tissue of the subject (eg, gut, oral cavity, lung, armpit, ocular, vaginal, anal, ear, nose or throat tissue) and thus we believe that the demonstration of activity to inhibit growth of a microbiota bacterial species (exemplified by *Streptococcus)* on a surface supports this utility.

[0104] In an example, wild-type host cell endogenous Cas9 or cfp1 activity is used. The engineered nucleotide sequence may not be in combination with an exogenous Cas nuclease-encoding sequence.

[0105] In an example, the host cells are wild-type (eg, non-engineered) bacterial cells. In another example, the host cells are engineered (such as to introduce an exogenous nucleotide sequence chromosomally or to modify an endogenous nucleotide sequence, eg, on a chromosome or plasmid of the host cell), and wherein the host cells comprise an endogenous CRISPR/Cas system having wild-type Cas nuclease activity that is operable with the crRNA or gRNA. In an example, the formation of bacterial colonies of said host cells is inhibited following said transformation. In an example, proliferation of host cells is inhibited following said transformation. In an example, host cells are killed following said transformation.

[0106] By "cognate to" it is intended that the endogenous Cas is operable with crRNA or gRNA sequence to be guided to the target in the host cell. The skilled addressee will understand that such Cas guiding is generally a feature of CRISPR/Cas activity in bacterial cells, eg, wild-type CRISPR/Cas activity in bacterial cells having endogenous active wild-type CRISPR/Cas systems.

[0107] By "wild-type"Cas activity it is intended, as will be clear to the skilled addressee, that the endogenous Cas is not an engineered Cas or the cell has not been engineered to de-repress the endogenous Cas activity. This is in contrast to certain bacteria where Cas nuclease activity is naturally repressed (ie, there is no wild-type Cas nuclease activity or none that is useful for the present invention, which on the contrary is applicable to addressing wild-type host cells *in situ* for example where the endogenous Cas activity can be harnessed to effect cell population growth inhibition).

[0108] In an example, inhibition of host cell population growth is at least 2, 3, 4, 5, 6, 7, 8, 9 or 10-fold compared to the growth of said host cells not exposed to said engineered nucleotide sequence. For example, growth inhibition is indicated by a lower bacterial colony number of a first sample of host cells (alone or in a mixed bacterial population) by at least 2, 3, 4, 5, 6, 7, 8, 9 or 10-fold compared to the colony number of a second sample of the host cells (alone or in a mixed bacterial population), wherein the first cells have been transformed by said engineered nucleotide sequence but the second sample has not been exposed to said engineered nucleotide sequence. In an embodiment, the colony count is determined 12, 24, 36 or 48 hours after the first sample has been exposed to the engineered sequence. In an embodiment, the colonies are grown on solid agar *in vitro* (eg, in a petri dish). It will be understood, therefore, that growth inhibition can be indicated by a reduction (<100% growth compared to no treatment, ie, control sample growth) in growth

of cells or populations comprising the target sequence, or can be a complete elimination of such growth. In an example, growth of the host cell population is reduced by at least 10, 20, 30, 40, 50, 60, 70, 80, 90 or 95%, ie, over a predetermined time period (eg, 24 hours or 48 hours following combination with the cRNA or gRNA in the host cells), ie, growth of the host cell population is at least such percent lower than growth of a control host cell population that has not been exposed to said cRNA or gRNA but otherwise has been kept in the same conditions for the duration of said predetermined period. In an example, percent reduction of growth is determined by comparing colony number in a sample of each population at the end of said period (eg, at a time of mid-exponential growth phase of the control sample). For example, after exposing the test population to the crRNA or gRNA a time zero, a sample of the test and control populations is taken and each sample is plated on an agar plate and incubated under identical conditions for said predetermined period. At the end of the period, the colony number of each sample is counted and the percentage difference (ie, test colony number divided by control colony number and then times by 100, and then the result is subtracted from 100 to give percentage growth reduction). The fold difference is calculated by dividing the control colony number by the test colony number.

[0109] Inhibition of population growth can be indicated, therefore, by a reduction in proliferation of host cell number in the population. This may be due to cell killing by the nuclease and/or by downregulation of host cell proliferation (division and/or cell growth) by the action of the nuclease on the target protospacer sequence. In an embomdiment of a treatment or prevention as disclosed herein, host cell burden of the human or animal subject is reduced, whereby the disease or condition is treated (eg, reduced or eliminated) or prevented (ie, the risk of the subject developing the disease or condition) is reduced or eliminated.

[0110] The invention is useful for targeting wild-type bacterial populations found naturally in the environment (eg, in water or waterways, cooling or heating equipment), comprised by beverages and foodstuffs (or equipment for manufacturing, processing or storing these) or wild-type bacterial populations comprised by human or animal microbiota. Thus, the invention finds utility in situations when pre-modification of host cells to make them receptive to killing or growth inhibition is not possible or desirable (eg, when treatment *in situ* of microbiota in the gut or other locations of a subject is desired). In another application, the invention finds utility for producing *ex vivo* a medicament for administration to a human or animal subject for treating or preventing a disease or condition caused or mediated by the host cells, wherein the medicament comprises a modified mixed bacterial population (eg, obtained from faeces or gut microbiota of one or more human donors) which is the product of the use of the invention or method of the disclosure, wherein the population comprises a sub-population of bacteria of a species or strain that is different to the species or strain of the host cells. The former sub-population cells do not comprise the target and thus are not modified by the use or method. Thus, for example, the method can be used to reduce the proportion of a specific *Firmicutes* sub-population and spare *Bacteroidetes* in the mixed population, eg, for producing a medicament for treating or preventing a metabolic or GI condition or disease disclosed herein. In this way, the invention can provide a modified bacterial transplant (eg, a modified faecal transplant) medicament for such use or for said treatment or prevention in a human or animal. For example, the method can be used to modify one or more microbiota *in vitro* to produce a modified collection of bacteria for administration to a human or animal for medical use (eg, treatment or prevention of a metabolic condition (such as obesity or diabetes) or a GI tract condition (eg, any such condition mentioned herein) or a cancer (eg, a GI tract cancer)) or for cosmetic or personal hygiene use (eg, for topical use on a human, eg, for reducing armpit or other body odour by topical application to an armpit of a human or other relevant location of a human). In another example, the array, crRNA, gRNA or engineered nucleotide sequence is administered to a human or animal and the host cells are harboured by the human or animal, eg, comprised by a microbiota of the human or animal (such as a gut microbiota or any other type of micriobiota disclosed herein). In this way, a disease or condition mediated or caused by the host cells can be treated or prevented. In an example, the transformation is carried out *in vitro* and optionally the array, crRNA, gRNA or engineered nucleotide sequence is comprised by nucleic acid that is electroporated into host cells. In an example, the nucleic acid are RNA (eg, copies of the gRNA). In another example, the nucleic acid are DNA encoding the crRNA or gRNA for expression thereof in host cells.

[0111] Thus, in an example, the invention provides an engineered nucleotide sequence for providing host cell modifying (HM) cRNA or guide RNA (gRNA) in a population of wild-type bacterial host cells comprised by a microbiota of a human or animal subject for treating or preventing a disease or condition mediated or caused by host cells of the microbiota of the subject, the cRNA or gRNA comprising a sequence that is capable of hybridising to a host cell target protospacer sequence for guiding Cas to the target, wherein the cRNA or gRNA is cognate to an endogenous host cell Cas nuclease that has wild-type nuclease activity, wherein following transformation of host cells growth of the population is inhibited and the disease or condition is treated or prevented.

[0112] In an example, the engineered nucleotide sequence comprises a HM-CRISPR array as defined herein. In an example, the engineered nucleotide sequence encodes a single guide RNA. In an example, the engineered nucleotide sequence is a guide RNA (eg, a singe guide RNA) or crRNA. In an example, the engineered sequence is comprised by a bacteriophage that is capable of infecting the host cells, wherein the transformation comprises transduction of the host cells by the bacteriophage. The bacteriophage can be a bacteriophage as described herein. In an example, the engineered nucleotide sequence is comprised by a plasmid (eg, a conjugative plasmid) that is capable of transforming host cells.

The plasmid can be a plasmid as described herein. In an example, the engineered nucleotide sequence is comprised by a transposon that is capable of transfer into and/or between host cells. The transposon can be a transposon as described herein.

[0113] Any use or method of the invention can comprise transforming host cells with nucleic acid vectors for producing cRNA or gRNA in the cells. For example, the vectors or nucleic acid comprising the engineered nucleotide sequence are administered orally, intravenously, topically, ocularly, intranasally, by inhalation, by rectal administration, in the ear, by vaginal administration or by any other route of administration disclosed herein or otherwise to a human or animal comprising the mixed bacterial population (eg, as part of microbiota of the human or animal), wherein the administration transforms the host cells with the vectors or nucleic acid.

[0114] In an example, the host cell population is *ex vivo.* In an example, the mixed population is comprised by a human or animal subject and a host cell infection in the subject is treated or prevented.

[0115] In an example, the first and second bacteria are comprised by a microbial consortium wherein the bacteria live symbiotically. In an example, the consortium is a human or animal microbiota; in an example the consortium is comprised by a human or animal (eg, wherein the use, system, engineered sequence, vector or cell is for treating infection by host cells of the consortium in the human or animal, eg, wherein the host cells mediate or cause antibiotic resistance or a deleterious disease or condition in the human or animal). The species (*E coli, L lactis* and *5 thermophilus*) used in the worked Example below are strains that co-exist symbiotically in human and animal gut microbiota. The Example also addresses targeting in a mixed gram positive and gram negative bacterial population. Additionally, the Example addresses a population of *Firmicutes* (*S thermophilus*) and a population of *Enterobacteriaceae* (*E coli*), both of which are found in human microbiota. Other examples of *Enterobacteriaceae are Salmonella, Yersinia pestis, Klebsiella, Shigella, Proteus, Enterobacter, Serratia,* and *Citrobacter.*

[0116] In an example, the method, use, engineered nucleotide sequence, array, crRNA, gRNA, vector or system is for treating host cell infection in a human gut microbiota population, optionally the population also comprising first bacteria that are human commensal gut bacteria and/or *Enterobacteriaceae,* eg, wherein the host cells and commensal cells (first and second bacteria) live symbiotically in human gut microbiota.

[0117] In an example the use or system is for the alteration of the proportion of *Bacteroidetes* bacteria in a mixed bacterial population comprising *Bacteroidetes* bacteria and other bacteria. For example, for for increasing the relative ratio of *Bacteroidetes* versus one, more or all *Firmicutes* (eg, versus *Streptococcus*) in the population. In this case, the host cells can be *Firmicutes* cells comprising the target(s). In an example, the population is a bacterial population of a microbiota comprised by a human or animal subject and the method, use, engineered nucleotide sequence, vector or system is for (i) treating an infection in the subject by said host cells comprised (eg, comprised by the mixed population); (ii) treating or preventing in the subject a condition or disease mediated by said host cells; (iii) reducing body odour of the human that is caused or mediated by said host cells; or (iv) personal hygiene treatment of the human. In an example, the engineered nucleotide sequence, array, crRNA, gRNA or vector of the disclosure is for use in such a system or use of the disclosure.

[0118] In an example, the condition or disease is a metabolic or gastrointestinal disease or condition, eg, obesity, IBD, IBS, Crohn's disease or ulcerative colitis. In an example, the condition or disease is a cancer, eg, a solid tumour or a GI cancer (eg, stomach cancer), liver cancer or pancreatic cancer. In an example, the condition is resistance or reduced responsiveness to an antibiotic (eg, any antibiotic disclosed herein).

[0119] In an example, the cell comprises an endogenous RNase III that is operable with component (ii) in the production of said HM-crRNA in the cell. In an alternative, one or more of the vectors comprises a nucleotide sequence encoding such a RNase III for expression of the RNase III in the host cell.

[0120] In an example, the essential gene (comprising the target) encodes a DNA polymerase of the cell. This is exemplified below.

[0121] In an example of the use, system, vector or cell, array, cRNA or gRNA comprises a sequence that is capable of hybridising to a host cell target protospacer sequence that is a adjacent a NGG, NAG, NGA, NGC, NGGNG, NNGRRT or NNAGAAW protospacer adjacent motif (PAM), eg, a AAAGAAA or TAAGAAA PAM (these sequences are written 5' to 3'). In an embodiment, the PAM is immediately adjacent the 3' end of the protospacer sequence. In an example, the Cas is a S aureus, S theromophilus or S pyogenes Cas. In an example, the Cas is Cpf1 and/or the PAM is TTN or CTA.

[0122] In an example the engineered nucleotide sequence, crRNA, gRNA or array is in combination with an antibiotic agent, eg, wherein the target is comprised by an antibiotic resistance gene wherein the antibiotic is said agent. In embodiment, the host cells are sensitive to the antibiotic. For example, there may be insufficient sensitivity to use the antibiotic to eradicate infection of presence of the host cells (eg, in a human or manufacturing vessel/equipment comprising the population), but the antibiotic can dampen down or reduce host cell sub-population size or growth whilst further killing or growth inhibition is effected using Cas modification (eg, target cutting) according to the invention.

[0123] The disclosure provides the use, system, array, crRNA, gRNA, engineered nucleotide sequence, vector or cell for a method of antibiotic (first antibiotic) treatment of an infection of said host cells in a human or animal subject, wherein an antibiotic resistance gene (for resistance to the first antibiotic) is Cas-targeted by the system or vector in host cells,

wherein the method comprises administering the system, array, crRNA, gRNA, engineered nucleotide sequence, vector or cell and the antibiotic to the subject. The gene is downregulated, ie, expression of a protein product encoded by the gene is reduced or eliminated in the host cell, whereby antibiotic resistance is downregulated. The infection is reduced or prevented in the subject. In an example, the antibiotic is administered simultaneously with the system, array, crRNA, gRNA, engineered nucleotide sequence, vector or cell; in another example, the administration is sequential (eg, the antibiotic before the system, array, crRNA, gRNA, engineered nucleotide sequence, vector or cell). This feature of the invention can be useful for enhancing antibiotic treatment in the subject, eg, when antibiotic alone is not fully effective for treating such a host cell infection. The antibiotic can be any antibiotic disclosed herein, eg, tetracycline.

[0124] In an example, each engineered nucleotide sequence or vector comprises a said CRISPR array or a sequence encoding a said crRNA or gRNA and further comprises an antibiotic resistance gene (eg, kanamycin resistance), wherein the HM-crRNA or gRNA does not target the antibiotic resistance gene. In an example, the target sequence is comprised by an antibiotic resistance gene of the host cell, wherein the antibiotic is different from the first antibiotic (eg, kanamycin). In this way, the system, engineered sequence or vector is able to target the host without targeting itself. By exposing the host cells to the first antibiotic, one can promote retention of the engineered sequence or vector therein by positive selection pressure since cells containing the first antibiotic resistance gene will have a survival advantage in the presence of the first antibiotic (when host cells that are not transformed by the engineered sequence or vectors are not resistant to the first antibiotic). Thus, an example provides: The use of the disclosure comprising exposing the host cell or mixed population to said antibiotic (eg, kanamycin) and said engineered sequence or vector(s), for promoting maintenance of cRNA or gRNA-encoding sequences in host cells; or the system, engineered sequence, array or vector of the disclosure is in combination with said antibiotic.

[0125] In an example the sequence encoding the cRNA or gRNA or the component (ii) is under a constitutive promoter (eg, a strong promoter)operable in the host cell species, or an inducible promoter. In an example component (iii) is under a constitutive promoter operable in the host cell species, or an inducible promoter.

[0126] In an example, the or each host cell is a gram positive cell. In another example, the or each host cell is a gram positive cell.

[0127] In an example the method, use, system, engineered sequence or vector is for treating host cell infection in a human gut microbiota population, optionally the population comprising human commensal gut bacteria (ie, gut bacteria that are commensal with humans).

[0128] In an example of the method, use, system, array, crRNA, gRNA,engineered sequence or vector, the host cells are comprised by a mixed bacterial population comprised by a human or animal subject and the method, use, system, array, crRNA, gRNA,engineered sequence or vector is for (i) treating an infection in the subject by said host cells comprised by the mixed population; (ii) treating or preventing in the subject a condition or disease mediated by said host cells; (iii) reducing body odour of the human that is caused or mediated by said host cells; or (iv) personal hygiene treatment of the human.

[0129] In an example of the method, use, system, array, crRNA, gRNA,engineered sequence or vector is for *in vitro* treating an industrial or medical fluid, solid surface, apparatus or container (eg, for food, consumer goods, cosmetics, personal healthcare product, petroleum or oil production); or for treating a waterway, water, a beverage, a foodstuff or a cosmetic, wherein the host cell(s) are comprised by or on the fluid, surface, apparatus, container, waterway, water, beverage, foodstuff or cosmetic.

[0130] The disclosure also provides: An *ex vivo* mixed population of bacteria obtainable by the use or method of any concept herein.

[0131] In an example, the mixed population or the product of the use or method is in a container for medical or nutiritional use. For example, the container is a sterilised container, eg, an inhaler or connected to a syringe or IV needle.

[0132] In an example, the product population of the use or method is useful for administration to a human or animal to populate a microbiome thereof.

[0133] The disclosure provides: A foodstuff or beverage for human or non-human animal consumption comprising the the population product of the use or method.

[0134] Herein, in an example of any configuration, concept or aspect, the *Bacteroides* is a species selected from *caccae, capillosus, cellulosilyticus, coprocola, coprophilus, coprosuis, distasonis, dorei, eggerthii, faecis, finegoldii,fluxus, fragalis, intestinalis, melaninogenicus, nordii, oleiciplenus, oralis, ovatus, pectinophilus, plebeius, stercoris, thetaiotaomicron, uniformis, vulgatus* and *xylanisolvens.* For example, the *Bacteroides* is *thetaiotaomicron,* eg, wherein the host cell or mixed population is a gut microbiota population *ex vivo* or *in vitro.* In an example, the host cells, first or second bacteria sub-population comprises a plurality of different *Bacteroidetes* species, or a plurality of *Bacteroides* species (eg, comprising *B thetaiotaomicron* and *B fragalis),* or *Bacteroides* and *Prevotella* species. Herein, in an example, the *Prevotella* is a species selected from *bergensis, bivia, buccae, buccalis, copri, melaninogenica, oris, ruminicola, tannerae, timonensis* and *veroralis.* In an alternative, the host cells, first or second bacteria are *Firmicutes* cells. In an example, the host cells, first or second sub-population comprises or consists of one or more *Firmicutes* selected from *Anaerotruncus, Acetanaerobacterium, Acetitomaculum, Acetivibrio, Anaerococcus, Anaerofilum, Anaerosinus, Anaerostipes, Anaerovo-*

rax, Butyrivibrio, Clostridium, Capracoccus, Dehalobacter, Dialister, Dorea, Enterococcus, Ethanoligenens, Faecalibacterium, Fusobacterium, Gracilibacter, Guggenheimella, Hespellia, Lachnobacterium, Lachnospira, Lactobacillus, Leuconostoc, Megamonas, Moryella, Mitsuokella, Oribacterium, Oxobacter, Papillibacter, Proprionispira,Pseudobutyrivibrio, Pseudoramibacter, Roseburia, Ruminococcus, Sarcina, Seinonella, Shuttleworthia, Sporobacter, Sporobacterium, Streptococcus, Subdoligranulum, Syntrophococcus, Thermobacillus, Turibacter and *Weisella.* In an example, the host cells, or the first or second sub-population consists of *Clostridium* cells (and optionally the other sub-population consists of *Bacteroides* (eg, *thetaiotaomicron)* cells). In an example, the host cells, or the first or second sub-population consists of *Enterococcus* cells (and optionally the other sub-population consists of *Bacteroides* (eg, *thetaiotaomicron)* cells). In an example, the host cells, or the first or second sub-population consists of *Ruminococcus* cells (and optionally the other sub-population consists of *Bacteroides* (eg, *thetaiotaomicron)* cells). In an example, the host cells, or the first or second sub-population consists of *Streptococcus* cells (and optionally the other sub-population consists of *Bacteroides* (eg, *thetaiotaomicron)* cells). In an example, the host cells, or the first or second sub-population consists of *Faecalibacterium* cells (and optionally the other sub-population consists of *Bacteroides* (eg, *thetaiotaomicron)* cells). For example, the Faecalibacterium is a Faecalibacterium prausnitzii (eg, A2-165, L2-6, M21/2 or SL3/3).

[0135] In an example, the host cells, or the first or second sub-population comprises or consists of one or more *Firmicutes* selected from *Anaerotruncus, Acetanaerobacterium, Acetitomaculum, Acetivibrio, Anaerococcus, Anaerofilum, Anaerosinus, Anaerostipes, Anaerovorax, Butyrivibrio, Clostridium, Capracoccus, Dehalobacter, Dialister, Dorea, Enterococcus, Ethanoligenens, Faecalibacterium, Fusobacterium, Gracilibacter, Guggenheimella, Hespellia, Lachnobacterium, Lachnospira, Lactobacillus, Leuconostoc, Megamonas, Moryella, Mitsuokella, Oribacterium, Oxobacter, Papillibacter, Proprionispira,Pseudobutyrivibrio, Pseudoramibacter, Roseburia, Ruminococcus, Sarcina, Seinonella, Shuttleworthia, Sporobacter, Sporobacterium, Streptococcus, Subdoligranulum, Syntrophococcus, Thermobacillus, Turibacter* and *Weisella.* In an example, the host cells, or the first or second sub-population consists of *Clostridium* (eg, dificile) cells (and optionally the other sub-population consists of *Bacteroides* (eg, *thetaiotaomicron)* cells). In an example, the host cells, or the first or second sub-population consists of *Enterococcus* cells (and optionally the other sub-population consists of *Bacteroides* (eg, *thetaiotaomicron)* cells). In an example, the host cells, or the first or second sub-population consists of *Ruminococcus* cells (and optionally the other sub-population consists of *Bacteroides* (eg, *thetaiotaomicron)* cells). In an example, the host cells, or the first or second sub-population consists of *Streptococcus* cells (and optionally the other sub-population consists of *Bacteroides* (eg, *thetaiotaomicron)* and/or *Enterobacteriaceae* (eg, *E coli)* cells). In an example, the host cells, or the first or second sub-population consists of *Faecalibacterium* cells (and optionally the other sub-population consists of *Bacteroides* (eg, *thetaiotaomicron)* cells). In an example, the host cells, or the first or second sub-population consists of *Streptococcus* cells (optionally *S thermophilus* and/or *pyogenes* cells) and the other sub-population consists of *Bacteroides* (eg, *thetaiotaomicron*) and/or *Enterobacteriaceae* (eg, *E coli)* cells.

[0136] The population product of the use of the invention or method of the disclosure is, in an embodiment, for administration to a human or non-human animal by mucosal, gut, oral, intranasal, intrarectal, intravaginal, ocular or buccal administration.

[0137] Optionally the host cells, or the first or second sub-population bacteria are *B fragalis* bacteria and the population is harboured by water.

[0138] A suitable beverage comprising an array, system, engineered sequence, vector or gRNA of the disclosure is, for example, a probiotic drink, eg, an adapted Yakult (trademark), Actimel (trademark), Kevita (trademark), Activia (trademark), Jarrow (trademark) or similar drink for human consumption.

PHAGE SEQUENCE TARGETS

[0139] In aspects of the disclosure, the target sequence is a sequence of a phage that infects a host bacterial cell. Desired modification of phage genomes, as achieved by the disclosure, not only relates to phage killing or knock-down, but instead can be desired phage gene or regulatory element activation in the host cell (eg, when the phage expresses a desired protein or other product that is associated with increased host cell viability or proliferation). Alternatively, modification may be inducible phage gene expression regulation, eg, by use of an inducible Cas that is targeted according to the disclosure to the phage target site. In an embodiment, the disclosure provides for modifying the phage target site by cutting with a Cas nuclease in the host cell. This may be useful for various reasons, for example:-

[0140] A. to mutate the target site to activate or inactivate it (eg, for gene knock-down or inactivation of an anti-host gene; or for killing the host cell when the phage target is integrated in the host chromosome);

[0141] B. to delete the target sequence or a larger sequence comprising the target sequence (eg, when the disclosure is used with first and second PM-crRNAs that target spaced sites in the phage genome, wherein cuts in each site result in deletion of phage nucleic acid between the cuts); C. to insert a desired PM-DNA sequence into the host cell genome (eg, by providing one or more PM-crNA-guided cuts in a host nucleic acid for homologous recombination insertion of the desired PM-DNA).

[0142] The disclosure provides the following aspects:-

In a first aspect, there is provided a method of altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria comprising said sub-populations, wherein the first bacteria are host cells (eg, *Bacteroidetes* host cells) (wherein the first bacteria are optionally infected by a phage and the second bacteria are not infected by said phage (or not *Bacteroidetes)),* the method comprising combining the mixed population with a plurality of vectors in one or more steps for introduction of vector nucleic acid (eg, a PM-containing transposon thereof) into host cells and allowing bacterial growth in the mixed population, wherein the relative ratios of said first and second bacteria is altered; wherein each vector comprises an engineered phage-modifying (PM) CRISPR array for introduction into host cell for modifying a target nucleotide sequence (eg, of said phage) in the cell, (a) wherein the PM-CRISPR array comprises one or more sequences for expression of a PM-crRNA respectively and a promoter for transcription of the sequence(s) in a host cell; and (b) wherein the PM-crRNA is capable of hybridising to the target sequence to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence.

[0143] By targeting phage sequence(s) to inactivate gene(s) required for phage viability, propogation or infectivity, in one aspect the disclosure provides the array with a positive selective advantage that may promote its uptake and retention by host cells infected with the phage. When host cells are killed or growth is reduced, the relative ratio of first to second bacteria in the population is reduced. The disclosure provides such a product population, eg, for use as a medicament for treatment or prevention (reducing the risk) of a disease or condition in a human or animal subject, wherein the medicament is administered to the subject. The disease or condition can be any disease or condition disclosed herein. In an example, a single guide RNA (gRNA) is expressed in the host cells to provide the crRNA and each vector comprises an exprssible engineered nucleotide sequence encoding such a gRNA.

[0144] In an example using a PM-array, the target sequence is a *Bacteroides thetaiotaomicron* sequence. Optionally the target sequence is not comprised by *B fragalis.* This is useful, for example, where the modifying cuts or otherwise renders the target sequence non-functional, whereby the ratio of *B thetaiotaomicron* host cells is increased without targeting *B fragalis,* eg, where the mixed population is a gut microbiota population as described herein. *B fragalis* is in some settings associated with abscesses and thus this example reduces the risk of this, whilst enabling alteration of ratios (increase of *B thetaiotaomicron* cell proportion) as per the disclosure that is useful for example to rebalance gut microbiota, eg, for treating or preventing obesity or diabetes or IBD.

[0145] The promoter (or a HM- or PM-array) is operable in a host cell. In an example, the promoter is a viral or phage promoter, eg, a T7 promoter. In another example, the promoter is a bacterial promoter (eg, a promoter of the host cell species).

[0146] A second aspect relates to the method of the first aspect, wherein the first bacteria are *Bacteroides* (eg, *thetaiotamicron* or *fragalis),* Alistipes, *Alkaliflexus, Parabacteroides, Tannerella, Xylanibacter* and/or *Prevotella* bacteria.

[0147] A third aspect relates to the method of the first or second aspect, wherein the second bacteria are *Firmicutes* bacteria (eg, when the first bacteria are *Bacteroidetes* or *Bacteroides).*

[0148] A fourth aspect relates to the method of any preceding aspect, wherein the ratio of the first bacteria sub-population to the second bacteria sub-population is increased, ie, is greater after said method has been carried out than before.

[0149] A fifth aspect relates to the method of the fourth aspect, wherein the mixed population is comprised by a composition (eg, a beverage, mouthwash or foodstuff) for administration to a human or non-human animal for populating and rebalancing the gut or oral microbiota thereof, eg, wherein the mixed population is *in vitro,* or *in vivo* in the human or non-human animal. The method of aspect 1, 2 or 3, wherein the ratio of the first bacteria sub-population to the second bacteria sub-population is decreased, ie, is less after said method has been carried out than before.

[0150] A sixth aspect relates to the method of the fifth aspect, wherein the mixed population is harboured by a beverage or water (eg, a waterway or drinking water) for human consumption.

[0151] A seventh aspect realts to the method of any preceding aspect, wherein each vector is a plasmid, phage (eg, a packaged phage) or phagemid.

[0152] An eighth aspect relates to the method of the seventh aspect, wherein each vector is a phage (eg, a packaged phage) and vector nucleic acid is introduced into host cells by phage vector nucleic acid transduction into host cells, ie, by infection of host cells by phage vectors. In an example, the phage comprises one or more transposons as described herein.

[0153] A ninth aspect relates to the method of the eighth aspect, wherein each vector is a plasmid and vector nucleic acid is introduced into host cells by transformation or horizontal plasmid transfer from bacteria harbouring the vectors. In an example, the plasmid comprises one or more transposons as described herein. In an example, the bacteria harbouring the vectors is a *non-Bacteroidetes* or *non-Bacteroides* species.

[0154] Additionally or alternatively, the bacteria harbouring the vectors is a *non-Firmicutes* species. In an example, the bacteria harbouring the vectors are bacteria of one or more species selected from the group consisting of a *Lactobacillus* species (eg, *acidophilus* (eg, La-5, La-14 or NCFM), *brevis, bulgaricus, plantarum, rhammosus, fermentum, caucasicus, helveticus, lactis, reuteri* or *casei eg, casei* Shirota), a *Bifidobacterium* species (eg, *bifidum, breve, longum* or *infantis), Streptococcus thermophilus* and *Enterococcus faecium.* For example, the bacteria are *L acidophilus* or lactis

bacteria.

[0155] A tenth aspect relates to an engineered *Bacteroidetes* phage-modifying (PM) CRISPR array for use in the method of any preceding aspect for modifying the genome of said *Bacteroidetes* phage,

(a) wherein the PM-CRISPR array comprises one or more sequences for expression of a PM-crRNA and a promoter for transcription of the sequence(s) in a *Bacteroidetes* phage-infected host cell; and
(b) wherein the PM-crRNA is capable of hybridising to a *Bacteroidetes* phage genome target sequence to guide Cas (eg, a Cas nuclease) in the infected host cell to modify the target sequence.

[0156] An eleventh aspect relates to a nucleic acid vector (eg, a plasmid, phage or phagemid) for use in the method of any one of the first to tenth aspects, the vector comprising a PM-CRISPR array of the tenth aspect.

In a general embodiment, there is alternatively provided a twelfth aspect:-

[0157] A nucleic acid vector (eg, a plasmid, virus, phage or phagemid) comprising an engineered HM-CRISPR array for modifying a target sequence of the genome of a host bacterial cell (eg, pathogenic bacterial cell, such as described above) or the genome of a virus (eg, phage) in a host cell,

(a) wherein the CRISPR array comprises one or more sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in the host cell; and
(b) wherein the crRNA is capable of hybridising to the target sequence to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence.

[0158] The promoter is operable in a host cell. In an example, the promoter is a viral or phage promoter, eg, a T7 promoter. In another example, the promoter is a bacterial promoter (eg, a promoter of the host cell species).

[0159] In an example, the array is comprised by a transposon described herein. In an example, the array is comprised by a carrier bacterium as described herein. In an example, a plurality of the arrays is provided for targting one or more target nucleotide sequences of the phage or host cell, wherein the plurality of arrays are comprised by bacterial cells, eg, carrier, first recipient or second recipient cells as described herein. In an example, the carrier cells are comprised by a beverage (eg, a probiotic drink for human consumption) or foodstuff as described herein. In an example, the array or carrier bacteria are for administration to a human or non-human animal for treating or preventing an infection of the human or animal, eg wherein the host cell is pathogenic. In an example, the array or carrier bacteria are for administration to the gut of a human or non-human animal for treating or preventing obesity, diabetes or IBD of the human or animal.

[0160] A thirteenth aspect relates to the array or vector of the eleventh or twelfth aspectwherein the array or vector is comprised by a bacterial cell, eg, a probiotic cell for human or non-human animal consumption. A fourteenth aspect relates to the method, array or vector of any preceding aspect, wherein the vectors are comprised by a third bacterial population (eg, carrier bacteria described herein) that is used for said combining with the mixed population or is for combination with the mixed population, whereby vector nucleic acid is introduced into host cells by transformation (eg, by horizontal plasmid vector or transposon transfer from the third bacteria to the first bacteria host cells) or transduction (eg, by phage vector inefection of first bacteria host cells).

[0161] A fifteenth aspect relates to the method, array or vector of any preceding aspect, wherein the or each array or vector is comprised by a human or non-human animal gut commensal or symbiotic bacterial cell (eg, a carrier bacterial cell as described herein). Thus, the cell is of a gut bacterial species that is commensal or symbiotic with the human or non-human animal.

[0162] A sixteenth aspect relates to the method or vector of any one of the twelfth to fifteenth aspects, wherein the or each vector is a plasmid, phage or phagemid comprising an origin of replication that is operable in a *Firmicutes* host cell or in a *Bacteroidetes* phage-infected host cell (eg, a *Bacteroides* cell), and optionally operable in a commensal or symbiotic bacterial cell as defined in aspect 15. In an example, the origin of replication is *ori*T or any other origin of replication described herein.

[0163] A seventeenth aspect relates to the method or vector of any one of the twelfth to sixteenth aspects, wherein the or each vector is a plasmid or phagemid comprising a sequence (eg, a transposon described herein) that is capable of horizontal transfer between (1) a human or non-human animal commensal or symbiotic bacterial cell that is not a *Bacteroides* cell and (2) a said phage-infected cell which is a *Bacteroides* cell; or between (3) a a human or non-human animal commensal or symbiotic bacterial cell that is not a *Firmicutes* cell and (4) a *Firmicutes* cell comprising the target sequence.

[0164] An eighteenth aspect relates to the method or vector of any one of the twelfth to seventeenth aspects, wherein the or each vector is a plasmid or phagemid sequence (eg, a transposon described herein) that is capable of horizontal transfer between (1) a said phage-infected cell which is a *Bacteroides* cell and (2) a bacterial cell that is suitable for

probiotic administration to a human or non-human animal gut; or between (3) a *Firmicutes* cell comprising the target sequence and (4) a bacterial cell that is suitable for probiotic administration to a human or non-human animal gut.

[0165] A nineteenth aspect relates to the method or vector of any one of the fifteenth to eighteenth aspects, wherein the commensal, symbiotic or probiotic species is selected from the group consisting of a *Lactobacillus* species (eg, *acidophilus* (eg, La-5, La-14 or NCFM), *brevis, bulgaricus, plantarum, rhammosus, fermentum, caucasicus, helveticus, lactis, reuteri* or *casei* eg, *casei* Shirota), a *Bifidobacterium* species (eg, *bifidum, breve, longum* or *infantis), Streptococcus thermophilus* and *Enterococcus faecium.*

[0166] The method, array or vector of any preceding aspect, wherein the promoter is operable for transcription of said sequence(s) in a said phage-infected *Bacteroidetes* host cell and in a commensal, symbiotic or probiotic bacterial cell as defined in any one of aspects 15 to 19; or in a *Firmicutes* cell comprising the target sequence and in a commensal, symbiotic or probiotic bacterial cell as defined in any one of aspects 15 to 19. For example, the promoter is a viral or bacterial promoter, eg, a T7 promoter. In an example, the promoter is a host cell promoter, eg, a promoter *of a host* CRISPR/Cas array.

[0167] A twentieth aspect relates to the method, array or vector of any preceding aspect, or any use herein, wherein the modifying is (i) cutting of the target *sequence,* (ii) downregulating transcription of a gene comprising the target sequence, (iii) upregulating transcription of a gene comprising the target sequence, or (iv) adding, deleting or substituting a nucleic acid sequence at the target.

[0168] A twenty-first aspect relates to the method, array or vector of any preceding aspect, wherein the *Bacteroidetes* phage is a *Bacteroides* phage selected from a crAssphage, a GB-124 phage, a GA-17 phage, a HB-13 phage, a H16-10 phage, a B40-8 phage and *B fragalis* phage ATCC51477-B1. Reference is made to Nat Commun. 2014 Jul 24;5:4498. doi: 10.1038/ncomms5498, "A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes", Dutilh BE et al. The crAssphage ~97 kbp genome is six times more abundant in publicly available metagenomes than all other known phages together; it comprises up to 90% and 22% of all reads in virus-like particle (VLP)-derived metagenomes and total community metagenomes, respectively; and it totals 1.68% of all human faecal metagenomic sequencing reads in the public databases. Using a new co-occurrence profiling approach, Dutilh et al predicted a *Bacteroides* host for this phage, consistent with *Bacteroides*-related protein homologues and a unique carbohydrate-binding domain encoded in the phage genome. A twenty-second aspect relates to the method, array or vector of any preceding aspect, , or any use herein, wherein the target sequence is comprised by a phage gene required for host cell infectivity, the phage lysogenic or lytic cycle, or phage viability, eg, an essential gene or coat protein gene.

[0169] A twenty-third aspect relates to the method, array or vector of any preceding aspect, wherein the target sequence is comprised by a BACON (*Bacteroidetes*-associated carbohydrate-binding) domain-encoding sequence (eg, wherein the host is a *Bacteroides* host) or an endolysin-encoding sequence. Reference is made to FEBS Lett. 2010 Jun 3;584(11):2421-6, doi:10.1016/j.febslet.2010.04.045. Epub 2010 Apr 21, "Mining metagenomic data for novel domains: BACON, a new carbohydrate-binding module", Mello L et al. The presence of the BACON domain in a phage-structural protein might be explained by the proposed bacteriophage adherence to mucus model. According to this model, phage adhere to the mucin glycoproteins composing the intestinal mucus layer through capsid-displayed carbohydrate-binding domains (such as the immunoglobulin-like fold or the BACON domain), facilitating more frequent interactions with the bacteria that the phage infects.

[0170] A twenty-fourth aspect relates to the method, array or vector of any preceding aspect, or any use herein, wherein the CRISPR array comprises a sequence R1-S1-R1' for expression and production of the crRNA in the host cell,

(i) wherein R1 is a first CRISPR repeat, R1' is a second CRISPR repeat, and R1 or R1' is optional; and
(ii) S1 is a first CRISPR spacer that comprises or consists of a nucleotide sequence that is 95% or more identical to said target sequence. For example, the target sequence comprises a protospacer or is comprised by a protospacer sequence that is immediately adjacent to a protospacer adjacent motif (PAM) that is cognate to a Cas when the array of the invention is in the host cell, wherein the Cas is also cognate to the crRNA expressed from the array. In an embodiment, the Cas is endogenous to the cell. In another example, the Cas is exogenous to the host cell, eg, provided by a vector of the invention.

[0171] A twenty-fifth aspect relates to the method, array or vector of the twenty-fourth aspect, wherein R1 and R1' are at least 95% (eg, 96, 97, 98, 99 or 100%) identical to repeat sequences of a CRISPR array of a cell of the same species as the host cell.

[0172] A twenty-sixth The method, array or vector of the twenty-fourth aspect, wherein R1 and R1' is each at least 95% (eg, 96, 97, 98, 99 or 100%) identical to a repeat sequence of a CRISPR array (eg, a Type II-C array) of a *Bacteroides* species selected from *thetaiotamicron* and *fragalis* (eg, *Bacteroides fragalis* NCTC 9343), wherein the host cells comprise a CRISPR/Cas system that is functional with the repeat sequence and are *Bacteroides* cells, eg, of said species.

[0173] A twenty-seventh aspect relates to the method, array, use or vector of the twenty-sixth aspect, wherein R1 and R1' are at least 95% (eg, 96, 97, 98, 99 or 100%) identical respectively to the first (5'-most) and second (the repeat

immediately 3' of the first repeat) repeat sequences of a CRISPR array of said species, eg, of a said host cell of said species. In an example, the array is a Type II-C array. In an example, the array or vector further comprises R2-S2-R2', wherein the spacer S2 is the same or different from the spacer S1 (eg, for targeting a different target site in the host cell or phage genome), wherein R2 and R2' are functional in the host cell and are optionally the same as R1. For example, each of R1, R1', R2 and R2' is a *Bfragalis* CRISPR repeat.

[0174] A twenty-eigth aspect relates to the method, array, use or vector of the twenty-fourth aspect, wherein (iii) each of R1 and R1' is identical to a repeat sequence of a CRISPR array (eg, a Type II-C array) of a *Bacteroides* species cell, wherein the species is selected from the group consisting of *caccae, capillosus, cellulosilyticus, coprocola, coprophilus, coprosuis, distasonis, dorei, eggerthii, faecis, finegoldii,fluxus, fragalis* (eg*, fragalis* NCTC 9343), *intestinalis, melaninogenicus, nordii, oleiciplenus, oralis, ovatus, pectinophilus, plebeius, stercoris, thetaiotaomicron, uniformis, vulgatus* and *xylanisolvens,* and (iv) wherein the host cell comprises a CRISPR/Cas system that is funtional with the repeat sequence and is a *Bacteroides* cell of a species selected from said group (eg, the same species as the selected species of (iii)).

[0175] A twenty-ninth aspect relates to the method, array, use or vector of the twenty-fourth aspect, wherein R1 and R1' are functional with a CRISPR/Cas system of a said host *Bacteroidetes* or *Firmicutes* cell for modification of the target sequence. In an example, R1, R1', R2 and R2' are Type II (eg, Type II-C) CRISPR/Cas system repeats of the same bacterial species, eg, a *Bacteroides,* such as *thetaiotamicron* or *fragalis* or *Streptococcus,* such as *thermophilus* or *pyogenes.*

[0176] A thirtieth aspect relates to the method, array, use or vector of the twenty-fourth aspect, wherein R1 and R1' are at least 95% (eg, 96, 97, 98, 99 or 100%) identical to repeat sequences of a CRISPR array (eg, a Type II-C array) of a *Bacteroidetes* (eg, *Bacteroides* or *Prevotella*) or *Firmicutes* (eg, *Streptococcus)* cell. A thirty-first aspect relates to the method, array, use or vector of the twenty-fourth aspect, wherein each of R1 and R1' is at least 95% (eg, 96, 97, 98, 99 or 100%) identical to a sequence selected from SEQ ID NOs: 1 to 5 of Table 2 and optionally the first bacterial cells are *Bacteroides* cells, eg, of a species or strain (eg, the species or strain listed against the selected sequence) in Table 2.

[0177] A thirty-second aspect relates to the method, array, use or vector of the twenty-fourth aspect, wherein each of R1 and R1' is at least 95% (eg, 96, 97, 98, 99 or 100%) identical to a sequence selected from SEQ ID NOs: 6 to 11 Table 2 of and optionally the first bacterial cells are Prevotella cells, eg, of a species or strain (eg, the species or strain listed against the selected sequence) in Table 2.

[0178] A thirty-third aspect relates to the method, array or vector of any preceding aspect, wherein the or each array is in combination with one or more Cas nuclease(s) that function with the crRNA in a said host cell to modify the target sequence. For example, the target sequence comprises a protospacer sequence immediately adjacent to a Protospacer Adjacent Motif (PAM), optionally wherein the PAM is cognate to a Cas nuclease comprised by the *Bacteroidetes* host cells. In an example, the Cas is a Type II-C Cas nuclease.

[0179] A thirty-fourth aspect relates to the method, array or vector of any preceding aspect, wherein the or each array is in combination with nucleic acid sequence(s) encoding one or more Cas nuclease(s) that function with the crRNA in a said host cell to modify the target sequence.

[0180] A thirty-fifth aspect relates to the method, array, use or vector of the twenty-fourth aspect, wherein R1 and R1' are functional with a Type II Cas9 nuclease (eg, a *S pyogenes, S thermophilus* or 5 *aureus* Cas9) to modify the target in a said host cell, optionally wherein the method, array or vector is further according to the thirty-fourth or thirty-fifth aspect wherein the Cas is said Cas9.

[0181] A thirty-sixth aspect relates to an *ex-vivo* mixed population of bacteria obtainable by the method of any one of the first to tenth or fifteenth to thirty-fifth aspects or a use herein. For example, the mixed population is in a container for medical or nutiritional use. For example, the container is a sterilised container.

[0182] A thirty-seventh aspect relates to a composition for administration to a human or non-human animal for therapeutic, prophylactic, cosmetic, human or non-human animal body mass reduction (eg, cosmetic reduction) or nutritional use, the composition comprising the mixed population of the thirty-sixth aspect. In an example, the composition is for oral, systemic, inhaled, intrarectal, ocular, buccal or intravaginal administration. In an example, the composition is for administration to the gut or oral cavity of a human or non-human animal.

[0183] A thirty-eight aspect relates to a foodstuff or beverage for human or non-human animal consumption comprising the the mixed population of the thirty-sixth aspect or the composition of the thirty-seventh aspect.

[0184] A thirty-ninth aspect relates to the foodstuff or beverage of the thirty-eighth aspect, which is a nutritional supplement or a probiotic beverage or foodstuff.

[0185] A fortieth aspect relates to an antibiotic composition for treating or preventing a *Bacteroidetes* infection in a human or non-human animal or in drinking water, wherein the composition comprises an array or vector of any one of the eleventh to thirty-fifth aspects, optionally wherein the modifying is according to the twenty-second aspect (iii) or (iv).

[0186] A forty-first aspect relates to a probiotic composition for increasing the proportion *of gut Bacteroidetes* (eg, to treat or prevent obesity, diabetes (eg, Type I) or a GI inflammatory condition) in a human or non-human animal, wherein the composition comprises an array or vector of any one of the eleventh to thirty-fifth aspects, optionally wherein the

modifying is according to the twenty-second aspect (iii) or (iv).

[0187] A forty-second aspect relates to the composition of the thirty-seventh, fortieth or forty-first aspect for increasing the relative proportions of gut *Bacteroides* to *Firmicutes* in the human or animal, eg for treating or preventing obesity, diabetes (eg, Type I diabetes) or a GI condition (eg, Crohn's disease, IBD, IBS or ulcerative colitis).

[0188] In an alternative, "array" in any configuration as mentioned in claims 1-15 can instead be an engineered nucleotide sequence encoding a HM-crRNA or gRNA for expression in a host cell. The features of any of the aspects herein relating to an array can, therefore, in the alternative apply *mutatis mutandis* to such an engineered sequence.

MOBILE GENETIC ELEMENTS & CRISPR SYSTEMS

[0189] A forty-third aspect relates to a nucleic acid vector (eg, a plasmid, virus, phage or phagemid) comprising an engineered CRISPR array for modifying a target sequence of the genome of a host bacterial cell (eg, *Firmicutes* or pathogenic bacterial cell, such as described above) or the genome of a virus (eg, phage) in a host cell,

(a) wherein the CRISPR array comprises one or more sequences for expression of a crRNA (eg, comprised by a gRNA) and a promoter for transcription of the sequence(s) in the host cell;
(b) wherein the crRNA is capable of hybridising to the target sequence to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence;
(c) wherein the array is comprised by a transposon that is capable of horizontal transfer between first and second bacterial cells of different species.

[0190] Optionally, the Cas nuclease is a wild-type endogenous Cas nuclease of the host cell.

[0191] A forty-fourth aspect relates to the vector of the forty-third aspect, wherein the array is for adminstration to a human or non-huan animal; and the first cell species is non-pathogenic to the human or animal and the second cell species is pathogenic to the human or animal, wherein the array is comprised by the first cell.

[0192] A forty-fifth aspect relates to the vector of the forty-fourth aspect, wherein the first cell species is a species that is commensal or symbiotic with the human or animal, eg, a gut microbiota species.

[0193] A fourth-sixth aspect relates to the vector of the forty-fourth or forty-fifth aspect, wherein the first cell species is selected from the group consisting of a *Lactobacillus* species (eg, *acidophilus* (eg, La-5, La-14 or NCFM), *brevis, bulgaricus, plantarum, rhammosus, fermentum, caucasicus, helveticus, lactis, reuteri* or *casei* eg, *casei* Shirota), a *Bifidobacterium* species (eg, *bifidum, breve, longum* or *infantis), Streptococcus thermophilus* and *Enterococcus faecium.*

[0194] A forty-seventh aspect relates to the vector of any one of the forty-fourth to forty-sixth aspects, wherein the vector is comprised by a beverage (eg, a probiotic drink) or foodstuff for human or animal consumption.

[0195] A forty-eight aspect relates to the vector of any one of the forty-fourth or forty-seventh aspects, wherein the vector comprises at least one repeat-spacer-repeat unit for targeting the target sequence, wherein the repeats are at least 95% (eg, 96, 97, 98, 99 or 100%) identical to repeats of a CRISPR/Cas system of the host cell, whereby the repeats of the vector are operable in the host cell to guide Cas of the host system to modify the target nucleotide sequence.

[0196] A forty-ninth aspect relates to the vector of the forty-eighth aspect, wherein the vector lacks a Cas (eg, Cas nuclease)-encoding sequence.

[0197] Targeting of a nucleotide sequence of the host CRISPR/Cas system according to the invention is useful for removing host cell resistance to a vector (eg, invading virus) or reducing the development or increase of resistance. For example, the invention thereby provides the advantage of targeting and knocking down the activity of an endogenous CRISPR/Cas system so that new vector (eg, phage) spacer acquisition is inhibited.

[0198] A feature of mobilisation is the presence of a cis-acting region (oriT) that is required for transfer. This region is the initiation site of DNA processing at which a site- and strand-specific nick is made in the plasmid to start the transfer event. The disclosure provides further principles employing mobile genetic elements (MGEs) as follows:-

[0199] In a first principle, there is provided an engineered CRISPR nucleic acid vector comprising or consisting of a mobile genetic element (MGE), wherein the MGE comprises an origin of transfer *(ori*T) and a CRISPR array for modifying a target sequence of the genome of a host cell (eg, pathogenic bacterial cell) or the genome of a virus (eg, prophage) in a host cell,

(a) wherein the CRISPR array comprises one or more sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in the host cell;
(b) wherein the crRNA is capable of hybridising to the target sequence to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence;
(c) wherein the vector is capable of transfer between (i) first and second nucleic acid positions of a first host cell, wherein each position is a position on a chromosome or a plasmid and the target sequence is comprised by the host cell, or (ii) first and second host cells, wherein the target sequence is comprised by the first and/or second host

cell.

[0200] Examples of MGEs are ICEs, transposons, plasmids and bacteriophage. An origin of transfer (*oriT*) is a short sequence (eg, up to 500 bp) that is necessary for transfer of the DNA that contains it from a bacterial host to recipient during conjugation. The *oriT* is cis-acting — it is found on the same DNA that is being transferred, and it is transferred along with the DNA. A typical origin of transfer comprises three functionally defined domains: a nicking domain, a transfer domain, and a termination domain.

[0201] Optionally, the promoter is operable for transcription of said sequence(s) in the first and second (and optionally the third) cells.

[0202] Optionally the target sequence is comprised by the second cell. Optionally the target sequence is not comprised by the second cell.

[0203] In an example, the first and second cells are of different bacterial species (eg, species found in a human microbiome population, eg, of the gut, armpit, vagina or mouth). In an example, the first and second cells are *ex vivo.* In another example, the first and second cells are comprised by a human gut, vaginal, armpit or oral microbiome *in vivo* or *ex vivo.*

[0204] A second principle relates to the vector of the first principle, wherein the MGE is or comprises an integrative and conjugative element (ICE). Alternatively, the MGE is a mobilisable MGE (ie, able to use factors encoded by genes not carried by the MGE, in order to be mobilised). The terms "mobilisable" and "conjugative" in relation to MGEs are readily apparent to the skilled addressee.

[0205] Reference is made to the ICEberg database (http://db-mml.sjtu.edu.cn/ICEberg/), which provides examples of suitable ICEs for the invention and sources for suitable *ori*T. In an example, the ICE is a member of an ICE family comprising an ICE selected from the group 1 to 28, or the *ori*T is an *ori*T of a member of such a family: 1=SXT/R391; 2=Tn916; 3=Tn4371; 4=CTnDOT/ERL; 5=ICEclc; 6=ICEBs1; 7=ICEHin1056; 8=PAPI-1; 9=ICEMISym(R7A); 10=ICESt1; 11=SPI-7; 12=ICE6013; 13=ICEKp1; 14=TnGBS1; 15=Tn5253; 16=ICESa2603; 17=ICEYe1; 18=10270-RD.2; 19=Tn1207.3; 20=Tn1806; 21=ICEA5632; 22=ICEF-I/II; 23=ICEAPG2; 24=ICEM; 25=10270-RD.1; 26=Tn5801; 27=PPI-1; 28=ICEF-III. Family descriptions are found in the ICEberg database. For example, the Tn916 family was defined by Roberts et al (2009) (Trends Microbiol. 2009 Jun;17(6):251-8. doi: 10.1016/j.tim.2009.03.002. Epub 2009 May 20; "A modular master on the move: the Tn916 family of mobile genetic elements", Roberts A, Mullany P). Elements belonging to the Tn916 family are defined by the following criteria: they must have the general organization shown in Roberts *et al,* and they must have a core region (conjugation and regulation module) that is similar in sequence and structure to the original Tn916 at the DNA level. Exceptions are some conjugative transposons, such as Tn1549 which have been previously classified in this family and those with a high degree of protein similarity as described in corresponding references. A third principle relates to the vector of the second principle, wherein the ICE is a transposon, eg, a conjugative transposon. In an example, the MGE is a mobilisable transposon that is mobilisable in the presence of a functional helper element, optionally wherein the transposon is in combination with a said helper element.

[0206] A foruth principle relates to the vector of any preceding principle, wherein the vector is a plasmid, optionally wherein the MGE is a transposon comprised by the plasmid. For example, the transposon is a conjugative transposon. In an example the transposon is a mobilisable transposon (eg, mobilisable using one or more factors encoded by the plasmid, eg, by genes outside the transposon sequence of the plasmid). Optionally, the transposon is a Type I transposon. Optionally, the transposon is a Type II transposon.

[0207] A fifth principle relates to the vector of any preceding principle, wherein *ori*T is functional in the first and second host cells. This is useful to promote spread and propogation across bacteria in a bacterial population, eg, when the first and second cells are of different species.

[0208] A sixth principle relates to the vector of the fifth principle when comprised by the first cell, wherein the first cell comprises nucleotide sequences encoding proteins operable to transfer the MGE to the second cell, wherein the sequences are not comprised by the MGE. This is useful to avoid using space in the MGE for such sequences. For example, this enables construction of a more compact MGE for transfer between cells or enables inclusion of larger or more CRISPR arrays, eg, to include a plurality of spacers to target respective sequences in a host cell or to target different sequences in the first and second host cells.

[0209] A seveneth principle relates to the vector of the sixth principle, wherein the sequences are not comprised by the vector. This is useful to avoid using space in the vector or MGE for such sequences. For example, this enables construction of a more compact vector or MGE for transfer between cells or enables inclusion of larger or more CRISPR arrays, eg, to include a plurality of spacers to target respective sequences in a host cell or to target different sequences in the first and second host cells, and/or to include one or more sequences for encoding Cas protein(s), eg a Cas9.

[0210] An eighth principle relates to the vector of the sixth or seventh principle, wherein the sequences are comprised by a conjugative transposon of the first cell. This is useful since it enables harnessing of factors outside the MGE to effect conjugative transposition, for horizontal transfer of the MGE of the invention between first and second host cells (eg, of different bacterial species in a human microbiome). A ninth principle relates to the vector of the eighth principle,

wherein the transposon is operable in *trans* to transfer the MGE to the second cell. This is useful since it enables harnessing of factors outside the MGE to effect conjugative transposition, for horizontal transfer of the MGE of the disclosure between first and second host cells (eg, of different bacterial species in a human micribiome). For example, the *ori*T of the MGE of the disclosure is the same as an *ori*T comprised by a conjugative transposon of the host cell. This is useful to enable the MGE of the disclosure to operate with factors encoded by the host cell for effecting horizontal transfer of the MGE between the first and second host cells (eg, bacterial cells of different species, eg, human microbiome species). This enables the MGE to be more compact or frees up space for CRISPR arrays and/or Cas gene(s) as discussed above.

[0211] The term "operable in trans" means that the MGE (ICE) is operable for horizontal transfer using proteins expressed from host nucleotide sequences outside the vector nucleotide sequences (eg, proteins expressed by a conjugative transposon of the host cell) to transfer the MGE (or the entire vector, such as a plasmid containing the MGE) into the second cell.

[0212] A tenth principle relates to the vector of any preceding principle when comprised by the first cell, wherein the *ori*T of the MGE is the same as an ori*T* comprised by an ICE of the first cell, wherein the ICE is operable in *trans* to transfer the MGE to the second cell.

[0213] An eleventh principle relates to the vector of any preceding principle, wherein the vector *ori*T is an *ori*T of a *Bacteroidetes (eg,* Bacteroidales or *Bacteroides*) or *Prevotella* transposon. This useful when the first and/or second host cell is a *Bacteroidetes (eg,* Bacteroidales or *Bacteroides*) or *Prevotella* cell respectively. For example, the first cell is a cell of such a species and the second cell is a *Firmicutes* cell, the target sequence being comprised by the second cell but not the first cell, whereby the CRISPR array directs Cas in the second cell to cut the target sequence. In an example, the target sequence is comprised by an essential gene or antibiotic resistance gene of the second cell (and for the latter, optionally the vector is in combination with said antibiotic or administered to a human or non-human animal in combination with said antibiotic). Optionally, the transposon is a CTnDot or CTnERL transposon and the vector is in combination with tetracycline or administered to a human or non-human animal in combination with tetracycline.

[0214] A twelfth principle relates to the vector of any preceding principle, wherein the vector *ori*T is a CTnDot, CTnERL SXT/R391, Tn916 or Tn4371 family transposon *ori*T.

[0215] A thirteenth principle relates to the vector of any preceding principle, wherein the MGE comprises first and second terminal repeat sequences and the CRISPR array between the repeat sequences.

[0216] A fourteenth principle relates to the vector of any preceding principle, wherein the MGE leaves behind a transposon copy (1) at the first nucleic acid position when it has transferred to the second position; or (2) in the first cell when the it has transferred to the second cell. This is useful for promoting propogation and maintenance of the MGE in a bacterial population comprising the host cell(s). In an alternative, the MGE does not leave behind a transposon copy (i) at the first nucleic acid position when it has transferred to the second position; or (ii) in the first cell when the it has transferred to the second cell.

[0217] A fifteenth principle relates to the vector of any preceding principle when comprised by the first and/or second cell (eg, first and second copies of the vector comprised by the first and second cells).

[0218] A sixteenth principle relates to the vector of the fifteenth principle, wherein the first and second cells are cells of different species. For example, the first cell is a *Lactobacillus* cell (eg, as described herein) and/or the second cell is a *Bcteroidetes* (eg, *Bacteroides* cell, eg, such a cell described herein) or a *Firmicutes* cell (eg, such a cell described herein). In an example, the first cell is a *Bcteroidetes* (eg, *Bacteroides* cell, eg, such a cell described herein) and the second cell is a *Firmicutes* cell (eg, such a cell described herein), eg, for administration to a gut micribiome of a human for treating or preventing a GI condition or diabetes; or for treating or preventing obesity.

[0219] A seventeenth principle relates to the vector of the fifteenth or sixteenth principle, wherein the first and second cells are bacterial or archaeal cells.

[0220] An eighteenth principle relates to the vector of the sixteenth or seventeenth principle, wherein the first cell is non-pathogenic in a human (eg, a commensal or symbiotic bacterial cell) and optionally the second cell is a pathogenic cell in a human. In an alternative, the second cell is a non-pathogenic cell in a human. The term "non-pathogenic in a human" includes cells, such as certain bacterial species (eg, *Bacteroides* species, such as *fragalis*) that can reside in microbiomes of the human (eg, the gut, vaginal, armpit or oral microbiome) without pathogenicity or substantial pathogenicity, but in other environments of the human are pathogenic. The skilled person will readily understand that the first cell type can be retained in or on a human and the second cell type should be reduced in or on the human. For example, the CRISPR array modifies the genome of the second cell to kill or reduce cell viability or growth in or on the human. For example, the target site is comprised by the second cell and the site is cut by said Cas nuclease, thereby inactivating or down-regulating a gene comprising the target site. For example, the gene is an essential gene or antibiotic resistance gene of the second cell. In an example, the gene is a virulence gene.

[0221] A nineteenth principle relates to the vector of any preceding principle, or any use herein, wherein the second cell (each host cell) is a cell selected from (i) a Staphylococcus *aureus* cell, eg, resistant to an antibiotic selected from methicillin, vancomycin-resistant and teicoplanin; (ii) a *Pseudomonas aeuroginosa* cell, eg, resistant to an antibiotic

selected from cephalosporins (eg, ceftazidime), carbapenems (eg, imipenem or meropenem), fluoroquinolones, aminoglycosides (eg, gentamicin or tobramycin) and colistin; (iii) a *Klebsiella* (eg, *pneumoniae)* cell, eg, resistant to carbapenem; (iv) a *Streptoccocus (eg, pneumoniae* or pyogenes) cell, eg, resistant to an antibiotic selected from erythromycin, clindamycin, beta-lactam, macrolide, amoxicillin, azithromycin and penicillin; (v) a *Salmonella* (eg, serotype Typhi) cell, eg, resistant to an antibiotic selected from ceftriaxone, azithromycin and ciprofloxacin; (vi) a *Shigella* cell, eg, resistant to an antibiotic selected from ciprofloxacin and azithromycin; (vii) a *mycobacterium tuberculosis* cell, eg, resistant to an antibiotic selected from Resistance to isoniazid (INH), rifampicin (RMP), fluoroquinolone, amikacin, kanamycin and capreomycin; (viii) an *Enterococcus* cell, eg, resistant to vancomycin; (ix) an *Enterobacteriaceae* cell, eg, resistant to an antibiotic selected from a cephalosporin and carbapenem; (x) an *E. coli* cell, eg, resistant to an antibiotic selected from trimethoprim, itrofurantoin, cefalexin and amoxicillin; (xi) a *Clostridium* (eg, *dificile)* cell, eg, resistant to an antibiotic selected from fluoroquinolone antibiotic and carbapenem; (xii) a Neisseria *gonnorrhoea* cell, eg, resistant to an antibiotic selected from cefixime (eg, an oral cephalosporin), ceftriaxone (an injectable cephalosporin), azithromycin and tetracycline; (xiii) an *Acinetoebacter baumannii* cell, eg, resistant to an antibiotic selected from beta-lactam, meropenem and a carbapenem; or (xiv) a Campylobacter cell, eg, resistant to an antibiotic selected from ciprofloxacin and azithromycin. Such species can be pathogenic to humans.

[0222] A twentieth principle relates to the vector or use of the nineteenth principle, wherein the target site is comprised by an antibiotic resistance gene of the second cell, wherein the antibiotic is a respective antibiotic recited in the nineteenth principle.

[0223] A twenty-first principle relates to the vector of any one of the fifteenth to twentieth principles, wherein the first cell is a *Bacteroidetes (eg,* Bacteroidales or *Bacteroides)* cell; *Lactobacillus* (eg, *acidophilus* (eg, La-5, La-14 or NCFM), *brevis, bulgaricus, plantarum, rhammosus, fermentum, caucasicus, helveticus, lactis, reuteri* or *casei* eg, *casei* Shirota); *Bifidobacterium* (eg, *bifidum, breve, longum* or *infantis); Streptococcus thermophiles; Enterococcusfaecium; Alistipes; Alkaliflexus; Parabacteroides; Tannerella;* or *Xylanibacter* cell.

[0224] A twenty-second principle relates to the vector of any preceding principle, wherein the first and/or second nucleic acid positions of (i) are comprised by a *Bacteroidetes (eg,* Bacteroidales or *Bacteroides)* cell; or the first and/or second host cells of (ii) are *Bacteroidetes (eg,* Bacteroidales or *Bacteroides)* or *Prevotella* cells.

[0225] A twenty-third principle relates to the vector of the twenty-second principle, wherein the first cell is a *Bacteroidetes* (eg, *Bacteroidales* or *Bacteroides)* cell and the second cell is a *Firmicutes* (eg, Clostridium or Staphylococcus) cell, eg, wherein the vector is for administration to a gut micribiome of a human for treating or preventing a GI condition or diabetes; or for treating or preventing obesity.

[0226] A twenty-fourth principle relates to the vector of the sixteenth or seventeenth principle (or any use herein), wherein the first cell (each first cell) is environmentally-acceptable in an environment (eg, in a water or soil environment) and optionally the second cell (each host cell) is not acceptable in the environment. The water environment will be readily apparent to the skilled person and can, for example, be a marine or waterway (eg, lake, canal, river or reservoir) environment. In an example, the water environment is drinking water intended for human consumption or sewage water. In an example, the soil environment is soil of farming land or soil at a mining site (eg, a mineral or metal mining site).

[0227] By "acceptable" and "not acceptable" the skilled person will readily understand that the first cell type can be retained in the environment and the second cell type should be reduced in the environment. For example, the CRISPR array modifies the genome of the second cell to kill or reduce cell viability or growth in the environment. For example, the target site is comprised by the second cell and the site is cut by said Cas nuclease, thereby inactivating or down-regulating a gene comprising the target site. For example, the gene is an essential gene or antibiotic resistance gene of the second cell. In an example, the gene is a virulence gene.

[0228] In an example, the environment is a microbiome of a human, eg, the oral cavity microbiome or gut microbiome or the bloodstream. In an example, the environment is not an environment in or on a human. In an example, the environment is not an environment in or on a non-human animal. In an embodiment, the environment is an air environment. In an embodiment, the environment is an agricultural environment. In an embodiment, the environment is an oil or petroleum recovery environment, eg, an oil or petroleum field or well. In an example, the environment is an environment in or on a foodstuff or beverage for human or non-human animal consumption.

[0229] In an example, the vector, system, vector, array, crRNA, gRNA, method or any use herein is for use in an industry or the environment is an industrial environment, wherein the industry is an industry of a field selected from the group consisting of the medical and healthcare; pharmaceutical; human food; animal food; plant fertilizers; beverage; dairy; meat processing; agriculture; livestock farming; poultry farming; fish and shellfish farming; veterinary; oil; gas; petrochemical; water treatment; sewage treatment; packaging; electronics and computer; personal healthcare and toiletries; cosmetics; dental; non-medical dental; ophthalmic; non-medical ophthalmic; mineral mining and processing; metals mining and processing; quarrying; aviation; automotive; rail; shipping; space; environmental; soil treatment; pulp and paper; clothing manufacture; dyes; printing; adhesives; air treatment; solvents; biodefence; vitamin supplements; cold storage; fibre retting and production; biotechnology; chemical; industrial cleaning products; domestic cleaning products; soaps and detergents; consumer products; forestry; fishing; leisure; recycling; plastics; hide, leather and suede;

waste management; funeral and undertaking; fuel; building; energy; steel; and tobacco industry fields.

[0230] A twenty-fifth principle relates to the vector of any preceding principle in combination with a nucleic acid (eg, a DNA) for incorporation at the modified target site.

[0231] In an example, the modification is cutting of the target site and the nucleic acid (eg DNA) is incorporated by homologous recombination in the host cell. This is useful for effecting precise targeted modification of the host cell genome using the vector of the invention.

[0232] A twenty-sixth principle relates to the vector of the twenty-fifth principle, wherein the nucleic acid for incorporation is or comprises a regulatory element or exon sequence, eg a human sequence.

[0233] A twenty-seventh principle relates to the vector of any preceding principle in combination with a transposase for mobilisation of the MGE.

[0234] A twenty-eighth principle relates to the vector or any preceding principle, wherein the vector or MGE comprises a toxin-antioxin module that is operable in the first host cell; optionally wherein the toxin-antitoxin module comprises an anti-toxin gene that is not operable or has reduced operation in cells other than the first cell.

[0235] A twenty-ninth principle relates to the vector or any preceding principle, wherein the vector or MGE comprises a toxin-antioxin module that is operable in the second host cell; optionally wherein the toxin-antitoxin module comprises an anti-toxin gene that is not operable or has reduced operation in cells other than the second cell.

[0236] A thirtieth principle relates to the vector or any preceding principle, wherein the vector or MGE comprises a toxin-antioxin module that is operable in the first and second host cells ; optionally wherein the toxin-antitoxin module comprises an anti-toxin gene that is not operable or has reduced operation in cells other than the first and second cells. The use of a toxin-antitoxin module is useful to confer selective advantages and thus MGE retention and spread. For example, the module is a Type I module, eg, a Hok-Sok module. For example, the module is a Type II module, eg, a HiCa-HicB module. For example, the module is a tad-ata-type toxin-antitoxin module. For example, the module is a plasmid addiction module. In an example, the first and/or second cell is a *Bacteroides* cell and the module is a module of a *Bacteroides* species, eg, the Txe/YoeB family addiction module (see, eg, http://www.uniprot.org/uniprot/F0R9D1); RelE/StbE family addiction module (see, eg, http://www.uniprot.org/uniprot/F0R9A0); HigA family addiction module (see, eg, http://www.uniprot.org/uniprot/D7J8V2 or http://www.uniprot.org/uniprot/D2ESD0); RelE/StbE family addiction module (see, eg, http://www.uniprot.org/uniprot/F0R5F4). Use of a toxin-antitoxin in the vector or MGE can be useful to allow for destruction of a vector-bearing cell other than a cell that is desired (eg, the first and second and/or third bacterial cell). In this example, the MGE or vector comprises a toxin gene of a bacterial toxin-antitoxin module and a cognate anti-toxin gene, wherein the expression of the toxin and anti-toxin genes are separately regulated, eg, from different promoters. For example, the toxin gene can comprise a promoter that is constitutively active in the first, second (and third) cells so that the toxin is always produced. The anti-toxin gene can comprise a promoter that is inducible by one or more factors (eg, a protein expressed) in the first and/or second cells, but not in non-target cells of different strain or species. As is known, the anti-toxin is inherently less stable than the toxin in a bacterial toxin/anti-toxin system, and thus transfer of the vector or MGE to a cell that is not a target cell (eg, not the first and/or second cell) will lead to toxin expression in the absence of anti-toxin expression or lower anti-toxin activity, thus leading to cell death of the non-target cell. This, therefore creates a selection pressure for the target cells (first, second and third cells) to take up and retain the vector of the invention so that it can have the desired CRISPR array activity therein and also be propagated across target cells in a population (such as the gut microbiota). This also limits the spread of the vector or MGE to non-target cells so that the effect of the array is controlled in the population - in this respect there will be a pressure for non-target cells not to take up the vector and if they do, the recipient cells will not survive in the population, thereby limiting replication of non-target cells with the MGE and array.

[0237] A thirty-first principle relates to the vector of any preceding principle wherein the first and second cells are of the same phylum (eg, both bacterial cells) and the vector is replicable or operable (d) in the first cell and/or second cell but not in another cell of the same phylum; (e) in the first cell and/or second cell but not in another cell of the same order; (f) in the first cell and/or second cell but not in another cell of the same class; (g) in the first cell and/or second cell but not in another cell of the same order; (h) in the first cell and/or second cell but not in another cell of the same family; (i) in the first cell and/or second cell but not in another cell of the same genus; (j) in the first cell and/or second cell but not in another cell of the same species; (k) in the first cell and/or second cell but not in another cell of the same strain.

[0238] This affords selectivity of the vector of the invention (eg, for selective killing of the second host cell type in a mixed bacterial population) in a microbiome. This can be achieved, for example, by engineering the MGE or array (eg, the promoter thereof) so that it requires expression of a particular protein for replication or operation (eg, expression to produce crRNA). For example, the promoter can be selected from a promoter that operates in the first and/or second cell but not in other cells, or wherein the MGE is engineered so that one or more of the replication initiation sites thereof are dependent upon a protein or other factor produced in the first and/or second cell but in not other cells. A thirty-second principle relates to first and second copies of the vector of any preceding principle in a mixed population of cells, wherein the first vector is comprised by the first cell, the second vector is comprised by the second cell, the cells are cells of different species (eg, different bacterial species) and the one or both of the vector MGEs is capable of transferring to a

third cell (eg, a bacterial cell), wherein the third cell species is the same as the species of the first or second cell or is a species that is different from the first and second cell species. This is useful, since the first cell can act as a carrier (eg, when it is non-pathogenic it can be adminstered to a huma or animal so that it populates the human or animal, such as a microbiome thereof). By horizontal transfer, the carrier can transfer and propogate CRISPR arrays of the invention to third cells (directly or via second cells, the latter acting as a reservoir for arrays). The arrays can then mediate Cas modification (eg, cutting) of the target sequence in the third cells, eg, to inactivate or down-regulate an essential or antibiotic resistance gene of the third cells.

[0239] Generally herein, when the target sequence is comprised by an antibiotic resistance gene of a cell, the vector, engineered sequence or array of the invention can be administered to a human or animal together with (simultaneously or sequentially) the antibiotic. This is useful to kill or reduce proliferation of cells comprising the target sequence. In this respect, the vector, engineered sequence or array is comprised by a composition comprising an antibiotic, wherein the target sequence is a sequence of a gene encoding for resistance to said antibiotic.

[0240] Optionally, the mixed population comprises the third cell.

[0241] In an example, there is a provided a plurality of the first cells, each comprising a vector of the invention. In an example, there is a provided a plurality of the second cells, each comprising a vector of the invention. In an example, there is a provided a plurality of the first cells in combination with a plurality of the second cells, each cell comprising a vector of the invention. In an example, there is a provided a plurality of the first cells in combination with a plurality of the second cells and a plurality of the third cells, cells of at least 2 (or all of) said pluralities comprising a vector of the invention.

[0242] A thirty-third principle relates to the vectors of the thirty-second principle, wherein the vector or MGE comprises a toxin-antioxin module that is operable in the first, second and third host cells; optionally wherein the toxin-antitoxin module comprises an anti-toxin gene that is not operable or has reduced (ie, lesser) operation in cells other than the first, second and third cells.

[0243] A thirty-fourth principle relates to the vector of any preceding principle, wherein the MGE is a conjugative transposon, *ori*T is functional in the first and second host cells, the MGE comprises first and second terminal repeat sequences and the CRISPR array between the repeat sequences, and wherein the first and second cells are bacterial cells, the second cell being of a human microbiota cell species (eg, a pathogenic species), wherein the target site is comprised by the second cell but not the first cell, and wherein said modifying inactivates or down-regulates a gene or regulatory sequence comprising said target in the second cell.

[0244] Usefully, the first cells can thereby act as carriers and reservoirs for the arrays of the invention, which can be transferred by horizontal transfer of the MGEs.

[0245] In an example, the MGE is a conjugative *Bacteroidetes* transposon, *ori*T is a *Bacteroidetes ori*T functional in the first and second host cells, the MGE comprises first and second terminal repeat sequences and the CRISPR array between the repeat sequences, and wherein the first and second cells are bacterial cells, the first cell being a *Bacteroidetes* cell and the second cell being a *Firmicutes* cell (eg, Clostridium or Staphylococcus cell), wherein the target site is comprised by the second cell but not the first cell, and wherein said modifying inactivates or down-regulates a gene or regulatory sequence comprising said target in the second cell.

[0246] A thirty-fifth principle relates to the vector of the thirty-fourth principle when comprised by the first or second cell.

[0247] A thirty-sixth principle relates to the vector of any preceding principle, wherein the first and second cells are comprised by a mixed bacterial cell population, eg, a population of cells of human or non-human animal (eg, dog, cat or horse) gut, vaginal, armpit or oral microbiota species. As explained above, the population is useful for administration to a human or animal to populate a microbiome thereof.

[0248] A thirty-seventh principle relates to an *ex vivo* composition comprising a plurality of cells as defined in the twenty-second principle, wherein each cell comprises a vector according to any one of the first to thirty-sixth principles. Alternatively, the composition is *in vivo,* eg, in a non-human animal.

[0249] A thirty-eighth principle relates to a beverage or foodstuff for human or non-human animal consumption comprising a vector of any one of the first to thirty-sixth principles or the compositon of the thirty-seventh principle. The beverage can be, for example, a probiotic drink, eg, for consumption daily, once every two days or weekly by a human or animal, eg, to treat or prevent obesity or a GI condition in the human or animal.

[0250] A thirty-ninth principle relates to a composition comprising a plurality of *Bacteroides* cells, wherein each cell comprises a vector according to any one of the first to thirty-sixth principles.

[0251] Usefully, the cells can act as carriers and a reservoir of arrays of the invention, for administration to a microbiome (eg, gut microbiome) of a human or animal, eg, to treat or prevent obesity or a GI condition in the human or animal,

[0252] A fortieth principle relates to a mixed population of bacterial cells comprising a sub-population of first cells and a sub-population of second cells, wherein the first cells comprise vectors according to any one of the first to thirty-sixth principles, wherein the vectors are capable of horizontal transfer between the first and second cell sub-populations. Such a population is useful as it can be adminstered (eg, intranasally) to a human or animal so that the bacteria populate one or more microbiomes (eg, gut microbiome) of the human or animal. The first (and optionally also the second) cells

can act as carriers of the CRISPR arrays of the invention, especially when those cells are non-pathogenic to the human or animal (eg, non-pathogenic in the gut microbiome). The microbiome can be any other micribiome or microbiota population disclosed herein.

[0253] A forty-first principle relates to the population of the fortieth principle, wherein one or both of the first and second bacterial species is capable of populating the gut microbiota of a human or non-human animal, and optionally the first bacteria are commensal or symbiotic with humans or animals. Usefully, the first bacteria can be safely administered to the human or animal and can act as a carrier of the arrays of the invention for transfer thereafter to other cells of the microbiota.

[0254] A forty-second principle relates to the population of the fortieth principle, wherein the mixed population is harboured by a beverage or water (eg, a waterway or drinking water for human consumption) or soil. Provision of the population in water or soil is useful for treating such in the environment or (for water) in heating, cooling or industrial systems, or in drinking water storage containers.

[0255] In an example of any principle, the second cell is a cholera cell comprising the target sequence, wherein when the target sequence is modified the cell is killed or cell proliferation is reduced. In an example, the second cell is comprised by water for human consumption (eg, such water before or after processing for human consumption). In an example, the vector is comprised by a pharmaceutical compostion for administration to a human to treat or prevent cholera in the human.

[0256] A forty-third principle relates to a composition comprising a plurality of vectors according to any one of the first to thirty-sixth principles *in vitro.* For example, the composition is mixed with a multi-species bacterial population in an industrial apparatus or container (eg, for food, consumer goods, cosmetics, personal healthcare product, petroleum or oil production).

[0257] A forty-fourth principle relates to the vector, composition, foodstuff, beverage or population of any preceding principle for administration to a human or non-human animal for therapeutically or prophylactically populating and re-balancing a microbiome thereof or for cosmetically changing the human or animal (eg, for cosmetic weight-loss).

[0258] A forty-fifth principle relates to a method of modifying a target nucleotide sequence in a host cell, the method comprising

(1) combining the host cell with a carrier cell,

(a) wherein the carrier cell comprises a CRISPR nucleic acid vector comprising a CRISPR array for modifying the target,
(b) wherein the CRISPR array comprises one or more sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in the host cell;
(c) wherein the crRNA is capable of hybridising to the target sequence to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence; and

(2) culturing the cells together, wherein the vector is transfered from the carrier cell to the host cell, whereby the crRNA hybridises to the target sequence to guide Cas in the host cell and the target is modified.

[0259] In an example, the method is carried out *ex vivo.* In an example, the method is a cosmetic method and is not a therapetic or prophylactic medical method.

[0260] A forty-sixth principle relates to the method of the forty-fifth principle, wherein the vector is according to any one of principles 1 to 36.

[0261] A forty-seventh principle relates to the method of the forty-fifth or forty-sixth principle, wherein the host cell is a cell of a human or non-human animal microbiome bacterial species, optionally wherein the host cell is a cell of a pathogenic bacterial species. In an example, any microbiome herein is selected from a gut, vaginal, armpit, scalp, skin or oral microbiome.

[0262] A forty-eighth principle relates to the method of any one of the forty-fifth to forty-seventh principles, wherein the carrier cell is of a species that is a commensal or symbiotic human or non-human animal microbiome bacterial species. In an example, the carrier cell is non-pathogenic to humans, eg, when administered intranasally, topically or orally.

[0263] In any configuration, concept, aspect, embodiment or example etc herein the vector, composition, array or population of the disclosure is administered intranasally, topically or orally to a human or non-human animal, or is for such administration. The skilled person aiming to treat a microbiome of the human or animal will be able to determine the best route of administration, depending upon the microbiome of interest. For example, when the microbiome is a gut microbiome, administration can be intranasally or orally. When the microbiome is a scalp or armpit microbiome, administration can be topically. When the microbiome is in the mouth or throat, the administration can be orally.

[0264] A forty-ninth principle relates to the method of any one of the forty-fifth to forty-eighth principles, wherein the

host cell is of a gut microbiome bacterial species of a human or non-human animal.

[0265] A fiftieth principle relates to a method of altering the relative ratio of sub-populations of first and second bacteria host cell species in a mixed population of bacteria comprising said sub-populations, the method comprising

A: providing said first bacterial host cells;
B: providing the second bacterial host cells, wherein the second cells are cells of a different species or strain to the first cells;
C: introducing engineered CRISPR arrays into the first bacterial host cells, wherein wherein each CRISPR array comprises one or more sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in a said second host cell, wherein the crRNA is capable of hybridising to a target sequence comprised by said second cell to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence;
D: combining the first and second bacterial cells together to produce a mixed bacterial population; and
E: allowing bacterial growth in the mixed population such that horizontal transfer of CRISPR arrays from first bacterial cells to second bacterial cells occurs, wherein target sequences in second cells are Cas modified, whereby the relative ratios of said first and second bacteria is altered.

[0266] A fifty-first principle relates to the method of the fiftieth principle, wherein each CRISPR array is according to any one of the first to twenty-sixth principles 1 to 26.

[0267] A fifty-second principle relates to the method of the fiftieth or fifty-first principle, further comprising obtaining a first sample of the mixed population of step E and optionally comparing the proportion of second cells in the first sample to the proportion of second cells in a second sample of cells, wherein the second sample is a sample of a mixed population of bacterial cells used to provide the second cells in step B and the comparison shows that the proportion of second cells has increased or decreased after step E.

[0268] A fifty-third principle relates to the method of the fifty-second principle, wherein the second sample is a sample of a human or animal microbiome (eg, gut, vaginal, scalp, armpit, skin or oral cavity cells).

[0269] A fifty-fourth principle relates to the method of any one of the fiftieth to fifty-third principles, wherein a sample of a human or animal microbiome (eg, gut, vaginal, scalp, armpit, skin or oral cavity cells) is used to provide the second cells of step B.

[0270] A fifty-fifth principle relates to the method of any one of the fiftieth to fifty-fourth principles, wherein a recombinant, cultured population of the first cells is used for step A.

[0271] A fifty-sixth principle relates to the method of any one of the fiftieth to fifty-fifth principles, wherein plasmid, ICE or transposon horizontal transfer is used in step E, wherein each plasmid, ICE or transposon comprises a said CRISPR array.

[0272] A fifty-seventh principle relates to the method of any one of the fiftieth to fifty-sixth principles for therapeutically or prophylactically rebalancing the microbiota of a human or non-human animal, eg, for treating or preventing obesity, diabetes IBD, a GI tract condition or an oral cavity condition. The diabetes can be Type I or II. In an example, the prophylaxis is medical. In an example, the prophylaxis herein is non-medical, eg, cosmetic or for hygiene purposes. For example, the microbiota is an armpit microbiota and the method is for preventing or reducing body odour of a human. For example, in this case the method down-regulates growth or viability of host bacterial cells that mediate the generation and/or persistence of human body odour.

[0273] A fifty-eighth principle relates to the method of any one of the fiftieth to fifty-seventh principles, comprising providing third bacterial host cells of a species or strain that is different to the carrier and host cells, wherein the third cells are comprised by the mixed population in step E or combined with said population after step E, wherein horizontal transfer of CRISPR arrays to third host cells occurs.

[0274] A fifty-ninth principle relates to the method of the fifty-eighth principle, wherein the third cells do not comprise a said target sequence.

[0275] In this way, the third cells can act as carriers of the arrays and are capable of horizontally transferring arrays to host cells comprising the target sequence.

[0276] A sixtieth principle relates to the method of the fifty-eighth principle, wherein the third cells do comprise a target sequence for Cas modification.

[0277] A sixty-first principle relates to the method of any one of the fiftieth to sixtieth principles, wherein the carrier (and optionally also the third) cells are of a species recited in the twenty-first principle, eg, *Bacteroidetes* cells.

[0278] A sixty-second principle relates to the method of any one of the fiftieth to sixtieth principles, wherein the host cells are of a species recited in the nineteenth principle or *Firmicutes* cells.

[0279] A sixty-third principle relates to the vector, composition, foodstuff, beverage, population or method of any preceding principle wherein each vector is or is comprised by a plasmid, phage (eg, a packaged phage) or phagemid.

[0280] A sixty-fourth principle relates to the vector, composition, foodstuff, beverage, population or method of any preceding principle, wherein the modifying is (i) cutting of the target sequence, (ii) down-regulating transcription of a

gene comprising the target sequence, (iii) up-regulating transcription of a gene comprising the target sequence, or (iv) adding, deleting or substituting a nucleic acid sequence at the target.

[0281] A sixty-fifth principle relates to the vector, composition, foodstuff, beverage, population or method of any preceding principle, wherein each target sequence is a sequence comprised by a regulatory element or gene of the host cell, wherein the gene is an essential gene, a CRISPR gene or an antibiotic resistance gene, optionally wherein the regulatory element is an element of such a gene. In an alternative, the gene is a virulence gene.

[0282] A sixty-sixth principle relates to the vector, composition, foodstuff, beverage, population or method of any preceding principle, wherein each target sequence is a sequence comprised by a phage genome, wherein the phage is comprised by the host cell. In an example, the target sequence is comprised by a phage gene required for host cell infectivity, the phage lysogenic or lytic cycle, or phage viability, eg, an essential gene or coat protein gene.

[0283] In an example, the *Bacteroidetes* phage is a *Bacteroides* phage selected from a crAssphage, a GB-124 phage, a GA-17 phage, a HB-13 phage, a H16-10 phage, a B40-8 phage and B fragalis phage ATCC51477-B1. This is useful, for example, for providing a survival advantage to *Bacteroidetes* in the gut microbiome of a human or animal. In this way, the ratio of *Bacteroidetes* to *Firmicutes* can be altered to increase the proportion of the former versus the latter (eg, for treating or preventing obesity). In an example, the target sequence is comprised by a BACON (*Bacteroidetes*-associated carbohydrate-binding) domain-encoding sequence (eg, wherein the host is a *Bacteroides* host) or an endolysin-encoding sequence.

[0284] A sixty-seventh principle relates to the vector, composition, foodstuff, beverage, population or method of any preceding principle, wherein each CRISPR array comprises a sequence R1-S1-R1' for expression and production of the respective crRNA in the host cell,

(i) wherein R1 is a first CRISPR repeat, R1' is a second CRISPR repeat, and R1 or R1' is optional; and
(ii) S1 is a first CRISPR spacer that comprises or consists of a nucleotide sequence that is 95% or more identical to said target sequence.

[0285] A sixty-eighth principle relates to the vector, composition, foodstuff, beverage, population or method of the sixty-seventh principle, wherein R1 and R1' are at least 95% identical respectively to the first and second repeat sequences of a CRISPR array of the second host cell species.

[0286] A sixty-ninth principle relates to the vector, composition, foodstuff, beverage, population or method of the sixty-seventh to sixty-eighth principle, wherein R1 and R1' are functional with a CRISPR/Cas system of said host cell for modification of the target sequence.

1. A seventieth principle relates to the vector, composition, foodstuff, beverage, population or method of any preceding principle, wherein the or each array is in combination with one or more Cas nuclease(s) that function with the respective crRNA in a host cell to modify the target sequence. The target sequence comprises a protospacer sequence immediately adjacent to a Protospacer Adjacent Motif (PAM).

[0287] A seventy-first principle relates to the vector, composition, foodstuff, beverage, population or method of any preceding principle, wherein the or each array is in combination with nucleic acid sequence(s) encoding one or more Cas nuclease(s) that function with the respective crRNA in a host cell to modify the target sequence.

[0288] A seventy-second principle relates to the vector, composition, foodstuff, beverage, population or method of any one of the sixty-seventh to seventy-first principles, wherein R1 and R1' are functional with a Type II Cas9 nuclease (eg, a *S pyogenes* or *S aureus* Cas9) to modify the target in a said host cell, optionally wherein the vector, composition, foodstuff, beverage, population or method is further according to the seventieth or seventy-first principle wherein the Cas is said Cas9.

[0289] A seventy-third principle relates to an *ex-vivo* mixed population of bacteria obtainable by the method of any one of the fiftieth to seventy-second principles.

[0290] A seventy-fourth principle relates to a composition for administration to a human or non-human animal for therapeutic, prophylactic, cosmetic, human or non-human animal body mass reduction (eg, cosmetic reduction) or nutritional use, the composition comprising the mixed population of the seventy-third principle.

[0291] A seventy-fifth principle relates to a foodstuff or beverage for human or non-human animal consumption comprising the mixed population of the seventy-third principle or the composition of the seventy-fourth principle.

[0292] A seventy-sixth principle relates to the foodstuff or beverage of the seventy-fifth principle, which is a nutritional supplement or a probiotic beverage or foodstuff.

[0293] A seventy-seventh principle relates to an antibiotic composition for treating or preventing a bacterial infection in a human or non-human animal or in drinking water or in soil, wherein the composition comprises a vector of any one of the first to thirty-sixth and sixty-third to seventy-second principles 2.

[0294] A seventy-eighth principle relates to a probiotic composition for increasing the proportion of gut *Bacteroidetes*

(eg, to treat or prevent obesity, diabetes or a GI inflammatory condition) in a human or non-human animal, wherein the composition comprises a vector of any one of the first to thirty-sixth and sixty-third to seventy-second principles.

[0295] A seventy-ninth principle relates to the composition of the seventy-fourth, seventy-seventh and seventy-eighth principle for increasing the relative proportions of gut *Bacteroides* to *Fermicutes* in a human or animal, eg for treating or preventing obesity, diabetes or a GI condition.

[0296] An eightieth principle relates to the vector, composition, foodstuff, beverage, population or method of any preceding principle, wherein the vector does not comprise a Cas nuclease-encoding sequence operable with the array. This is useful to save space in the vector (eg, to allow for inclusion of larger arrays or more arrays for host cell targeting - this is useful to target multiple genome locations to reduce likelihood of evolution of resistance to the arrays of the invention).

[0297] An eighty-first principle relates to the vector, composition, foodstuff, beverage, population or method of any preceding principle, wherein the MGE does not comprise a Cas nuclease-encoding sequence operable with the array. This is useful to save space in the MGE (eg, to allow for inclusion of larger arrays or more arrays for host cell targeting - this is useful to target multiple genome locations to reduce likelihood of evolution of resistance to the arrays of the invention). For example, it is possible to avoid including the large sequence encoding Cas9 endounclease.

[0298] An eighty-second principle relates to the vector, composition, foodstuff, beverage, population or method of the eightieth to eighty-first principle, wherein the array is operable with a Cas endonuclease found in cells of the same species or strain as the first and/or second cell. In an example, the array is operable with a Cas endonuclease found in cells of the same species or strain as a host cell or third cell. This is useful to save space in the vector or MGE (eg, to allow for inclusion of larger arrays or more arrays for host cell targeting - this is useful to target multiple genome locations to reduce likelihood of evolution of resistance to the arrays of the invention).

[0299] An eighty-third principle relates to the vector, composition, foodstuff, beverage or population of any preceding principle, wherein the first and second cells are bacterial cells of different species, wherein the second cell is of a human microbiota species and the first cell is of a species that is non-pathogenic in said human microbiota, wherein the target sequence is not comprised by the genome of the first cell, the MGE comprising an *ori*T that is operable in the first and second cells, wherein the MGE is capable of horizontal transfer from the first cell to the second cell.

[0300] In an alternative, there is provided:-

The method of any preceding principle, wherein the carrier and host cells are bacterial cells of different species, wherein the host cell is of a human microbiota species and the carrier cell is of a species that is non-pathogenic in said human microbiota, wherein the target sequence is not comprised by the genome of the carrier cell, the MGE comprising an *ori*T that is operable in the carrier and host cells, wherein the MGE is capable of horizontal transfer from the carrier cell to the host cell.

[0301] An eighty-fourth principle relates to the vector, composition, foodstuff, beverage, population or method of the eighty-third principle, wherein the vector is comprised by a bacteriophage, the bacteriophage being capable of infecting the first cell (carrier) to introduce the MGE into the first (carrier) cell.

[0302] An eighty-fifth principle relates to the vector, composition, foodstuff, beverage, population or method of the eighty-third or eighty-fourth principle, wherein the target sequence is comprised by the genome of the second (host) cell (eg comprised by an essential or antibiotic resistance gene of the genome).

[0303] An eighty-sixth principle relates to the vector, composition, foodstuff, beverage, population or method of the eighty-fifth principle, wherein the second (host) cell species is pathogenic in said human microbiota, wherein the target sequence is modified by cutting of the target sequence or down-regulating a gene comprising said target sequence. In an example, the second (host) cell is a cell according to any one of features (i) to (xiv) of the nineteenth principle. In an example the second (host) cell is a *Firmicutes* cell, eg, wherein the vector is for treating or preventing obesity in a human.

[0304] An eighty-seventh principle relates to the vector, composition, foodstuff, beverage, population or method of the eighty-third, eighty-fourth or eighty-fifth principle, wherein the second (host) cell species is non-pathogenic in said human microbiota.

[0305] An eighty-eighth principle relates to the vector, composition, foodstuff, beverage, population or method of any one of the eighty-third to eighty-seventh principle, wherein the second (hosst) cell is a *Bacteroidetes* or *Prevotella* cell; optionally wherein the MGE is capable of horizontal transfer from the second (hosst) cell species to *Firmicutes* species of said human microbiota. The latter is useful, for example, for treating or preventing obesity in a human when the target sequence is comprised by the *Firmicutes,* but not the first (carrier) or second (host) cell.

[0306] An eighty-ninth principle relates to the vector, composition, foodstuff, beverage, population or method of any one of the eighty-third or eighty-eighth principle, wherein the MGE is capable of horizontal transfer from the second (host) cell species to a third bacterial cell species of said human microbiota, wherein the third cell species is pathogenic in said human microbiota and comprises said target sequence. In an example, the first (carrier) and second (host) cells do not comprise the target sequence. A ninetieth principle relates to the vector, composition, foodstuff, beverage, population or method of the eighty-ninth principle, wherein the third cell is a cell according to any one of features (i) to (xiv) of the nineteenth principle.

[0307] A ninety-first principle relates to the vector, composition, foodstuff, beverage, population or method of any preceding principle, wherein the MGE is devoid of a sequence encoding a Cas endonuclease that is operable with repeat sequences of the array, and wherein the vector comprises such a sequence (eg, encoding a Cas9) outside the MGE.

[0308] Any of the general features also may apply to the present configuration. Any of the features of any other configuration, aspect, paragraph, example, emodiment or concept herein also may be combined with the present configurations employing MGEs.

[0309] Thus, the invention provides the following features, numbered as paragraphs; these paragraphs apply to any of the aspects as recited, or to any of the first to ninety-first principles, or to any other configuration herein:-

In a first paragraph, there is provided a vector of any one of the forty-third to forty-ninth aspects, wherein the target sequence is a nucleotide sequence of a host CRISPR/Cas system, whereby the crRNA guides Cas to the target to modify the host CRISPR/Cas system in the host cell.

[0310] A second paragraph relates to the vector of the first paragraph 1, wherein the host CRISPR/Cas system is a Type I, II or III system and the target sequence is a nucleotide sequence conserved in said Type of system in at least one, two or three additional host strains or species, wherein said additional strains or species are different from said host.

[0311] A third paragraph relates to the vector of any preceding paragraph, wherein the target sequence is identical to a *Streptococcus* species (eg, *S thermophilus or S pyogenes*) CRISPR/Cas system sequence.

[0312] A fourth paragraph relates to the vector of any preceding paragraph, wherein the target sequence of the host CRISPR/Cas system comprises

i. a CRISPR array leader or leader promoter sequence contiguous with the 5'-most nucleotide of the first repeat (and optionally comprising said 5'-most nucleotide of the repeat, eg, comprising the first 3 nucleotides at the 5' end of the first repeat.);
ii. a sequence of up to 20 (eg, 3, 5, 7, 9, 10, 12, 15, 20, 30 or 32) nucleotides contiguous nucleotides immediately 5' of the first repeat;
iii. a sequence of up to 20 (eg, 3, 5, 7, 9, 10, 12, 15, 20, 30 or 32) contiguous nucleotides of the 5'-most nucleotides of the first repeat; or
iv. a sequence of up to 20 (eg, 3, 5, 7, 9, 10, 12, 15, 20, 30 or 32) contiguous nucleotides immediately 3' of the first spacer repeat (and optionally wherein the sequence comprises the 3'-most nucleotide of the first spacer, eg, comprising the last 3 nucleotides at the 3' end of the first repeat).

[0313] A fifth paragraph relates to the vector of the first, second or third paragraph, wherein the array is comprised by a nucleic acid vector (eg, a virus, virion, phage, phagemid or prophage) and

i. the crRNA comprises or consists of the structure R-S-R, wherein R=a CRISPR repeat and S=a CRISPR spacer, wherein S comprises, (in 5' to 3' direction) V-H_R or H_R-V or, wherein V=a sequence identical to a DNA sequence of the vector and H_R=a DNA sequence of a repeat of a CRISPR array of said host cell CRISPR array;
ii. wherein the sequence of H_R is immediately contiguous with the sequence of V in the host CRISPR array; and
iii. wherein the crRNA is capable of hybridising to a spacer of the host CRISPR array to guide Cas to the host target for modification of the host CRISPR array in the cell.

[0314] For example, V is a sequence of a phage vector coat protein-encoding sequence. In this respect Heler *et al* found in a study of bacteial resistance that three CRISPR-independent, bacteriophage-resistant mutants displayed a marked defect in phage adsorption (about 50%), indicating that most likely they carry envelope resistance mutations.

[0315] A sixth paragraph relates to the vector of the fifth paragraph, wherein the first crRNA does not or does not substantially hybridise to the nucleic acid present in the vector. For example, the first crRNA does not hybridise to V in the vector or hybridises less strongly than it hybridises to the spacer of the host array. Hybridisation testing is routine for the skilled person. For example, it can be determined *in vitro* by isolating or synthesizing the vector DNA and incubating it with the crRNA. Standard techniques, eg, using PCR can be used to detect whether or not hybridisation has occurred (eg, tested under pH and temperature conditions that would be found in host cell).

[0316] A seventh paragraph relates to the vector of the fifth or sixth paragraph, wherein V=one or up to 40 (eg, up to 15) contiguous nucleotides of vector DNA. The seed sequence immediately 5' of the PAM in the protospacer found in a target sequence is important for crRNA pairing and functioning of the CRISPR/Cas system to cut. This seed sequence includes around 15 or 12 continguous nucleotides immediately 5' of the PAM.

[0317] An eighth paragraph relates to the method, array or vector of any preceding aspect or paragraph, wherein the array is comprised by a vector and comprises (in 5' to 3' direction) a first repeat sequence, a first spacer sequence and a second repeat sequence, wherein the spacer sequence comprises a sequence that is capable of hybridising (eg, is identical to or has greater than 90% identity) to the target sequence in the host cell, the array further comprising a promoter for transcription of the repeats and spacer in the host cell, and optionally the vector comprises a Cas nuclease-

encoding sequence and/or a tracrRNA-encoding sequence for encoding a functional Cas and/or tracrRNA sequence in the host cell, wherein the tracrRNA sequence comprises a sequence that is complementary to the first or second repeat.

[0318] A ninth paragraph relates to the method, array or vector of any preceding aspect or paragraph, wherein the CRISPR array is comprised by a vector and comprises (in 5' to 3' direction) a first repeat sequence, a first spacer sequence and a second repeat sequence, wherein the spacer sequence comprises a sequence that is capable of hybridising (eg, is identical to or has greater than 90% identity) to the target sequence in the host cell, the array further comprising a promoter for transcription of the repeats and spacer in the host cell, and wherein the vector does not comprise a Cas nuclease-encoding sequence and/or a tracrRNA-encoding sequence for encoding a tracrRNA sequence in the host cell wherein the tracrRNA sequence comprises a sequence that is complementary to the first or second repeat, wherein the HM-CRISPR array is functional in the host cell to guide Cas (eg, endogenous host Cas nuclease) to the host target site, optionally using a host tracrRNA.

[0319] A tenth paragraph relates to the method, array or vector of the eighth or ninth paragraph, wherein the repeats are identical to repeats in a host array, wherein the CRISPR array of the invention does not comprise a PAM recognised by a Cas (eg, a Cas nuclease, eg, Cas9) of a host CRISPR/Cas system. The ability to omit Cas sequences frees up space in the array of the invetion.

[0320] An "essential gene" is a gene in the host whose presence or expression is required for host cell growth or for promoting or sustaining cell viability. A resistance gene is a gene in the host whose presence or expression is required for providing complete or partial resistance to an anti-host drug, eg, an antibiotic, eg, a beta-lactam antibiotic. A virulence gene is a gene in the host whose presence or expression is required for infectivity of an organism that the host cell is capable of infecting, eg, wherein the host is a pathogen (eg, of a plant, animal, human, livestock, companion pet, plant, bird, fish or insect).

[0321] An eleventh paragraph relates to the method, array or vector of any preceding aspect or paragraph, wherein the CRISPR array is in combination with a non-host cell Cas (eg, a Type I system Cas wherein the host system is a Type II or III; a Type II system Cas wherein the host system is a Type I or III; or a Type III system Cas wherein the host system is a Type I or II), optionally wherein the host cell does not comprise or express a Cas of a Type that is the same as the Type of the non-host Cas. This is useful since the CRISPR array does not target a sequence in itself (such as in the vector) or a vector-encoded Cas in the host.

[0322] A twelfth paragraph relates to the method, array or vector of any preceding aspect or paragraph, wherein the CRISPR array is in combination with a tracrRNA sequence or a sequence encoding a tracrRNA sequence (eg, on same nucleic acid as the array), optionally wherein the tracrRNA sesquence and HM-crRNA are comprised by a single guide RNA (gRNA)).

[0323] A thirteenth paragraph relates to the method, array or vector of any preceding aspect or paragraph, wherein the CRISPR array is in combination with a Cas or a sequence encoding a Cas, optionally wherein the array is integrated in a host cell genome and the Cas is endogenous to the host cell or encoded by an exogenous sequence. In an example, the Cas-encoding sequence is an exogenous sequence that has been introduced into the host, eg, from a plasmid or virus, such as a phage.

[0324] A fourteenth paragraph relates to the method, array or vector of any preceding aspect or paragraph, wherein the CRISPR array is comprised by a nucleotide sequence of a plasmid, virus, virion, phage, phagemid or prophage. The phagemid is a packaged phage. The prophage is a phage integrated into the host chromosome or episomal in the cell.

[0325] A fifteenth paragraph relates to the method, array or vector of any preceding aspect or paragraph, wherein the CRISPR array is integrated in a host cell genome, eg, in a chromosome or episomal nucleic acid.

[0326] In one example the array is in combination with a dead Cas (eg, dCas9) conjugated to a transcription or translation activator that acts on the target sequence or a gene comprising the target sequence. This is useful, for example, for switching on gene expression in the host cell (eg, of a desired gene, eg, an exogenous gene sequence that has previously been engineered into the host cell, eg, to encode an antibiotic where the host is a microbe, or to encode a desired exogenous protein for production in host culture, eg, for food, drink, medicine or any other application of the invention as disclosed herein).

[0327] A sixteenth paragraph relates to a virus (eg, a virion, phage, phagemid or prophage) comprising a CRISPR array of any preceding aspect or paragraph, eg, for infecting a cell, eg, a microbe or for use in medicine or dentistry.

[0328] A seventeenth paragraph relates to a population of virions according to the sixteenth paragraph, a first and a second virion thereof comprising different array leaders or promoters and/or for targeting different target sequences in the host cell or in different host strains.

[0329] An eighteenth paragraph relates to a collection of CRISPR arrays, each array being according to any preceding aspect or paragraph, wherein a first array comprises a first promoter for crRNA transcription; a second array comprises a second promoter for crRNA transcription that is different from the first promoter; and wherein each promoter is identical to a host promoter or is a homologue thereof; optionally wherein the first or both promoters is identical to a host Cas (eg, Cas1, 2, 9 or Csn2) promoter or a host CRISPR array promoter. For example, the first promoter is an endogenous Cas nuclease promoter or endogenous Cas1 or Cas2 promoter; or the promoter of an endogeous gene that is highly or

constitutively expressed or is an essential, virulence or resistance gene of the host cell. By using endogenous promoters, there will be pressure during evolution of the host to preserve the host promoters, and thus this decreases the likelihood of the host CRISPR/Cas defence system targeting one or more promoters of the arrays.

[0330] A nineteenth paragraph relates to a collection of CRISPR arrays of the invention, wherein a first array comprises one or more spacers (eg, 2, 3, 4, 5, 6, 7, 8, 9,10, 20, 30, 40, 50 or more spacers); and the second array comprises more than one spacer (eg, 2, 3, 4, 5, 6, 7, 8, 9,10, 20, 30, 40, 50 or more spacers), wherein said spacers of the second array are identical to the one or more spacers of the first array. This is useful for evading host resistance by homologous recombination of HM-array spacers, as proving many of such spacers in the HM-array (or furthermore distributing the spacers across a plurality of arrays) increases the chances that some HM-array spacers will remain in the host cell even if the host cell does delete some of the spacers. The defence against deletion is also enhanced by using different repeats flanking identical copies of the spacers in different arrays. Thus the disclosure provides the following:-

[0331] A twentieth paragraph relates to the collection of the eighteenth or nineteenth paragraph, wherein spacers (or said spacers) of the first array are flanked by first repeats that are identical; spacers (or said spacers) of the second array are flanked by second repeats that are identical; and wherein the first repeats are different from the second repeats.

[0332] A twenty-first paragraph relates to the collection of the twentieth paragraph, wherein the first repeats are identical to repeats in a host cell CRISPR/Cas system.

[0333] A twenty-second paragraph relates to the collection of the twentieth paragraph, wherein the first repeats are different from repeats in a host CRISPR/Cas system.

[0334] A twenty-third paragraph relates to the collection of any one of the eighteenth to twenty-second paragraphs, wherein the first and second arrays are contained in the same host cell or in the same vector (eg, plasmid, virus, virion, phage, phagemid or prophage).

[0335] A twenty-fourth paragraph relates to the collection of any one of the eighteenth to twenty-second paragraphs, wherein the first array is contained in a first vector and the second array is contained in a second vector which does not contain the first array (eg, wherein the vectors are plasmids or virions (eg, of the same virus type) or packaged phage (eg, of the same phage type).

[0336] In an embodiment, the vectors used in the method of the invention are vectors comprised by an array of any one of the eighteenth to the twenty-fourth paragraphs.

[0337] A twenty-fifth paragraph relates to a host cell comprising an array, virus, virion, phage, phagemid, prophage, population or collection according to any preceding paragraph.

[0338] Any of the general features (see below) also may apply to the present configuration.

[0339] An example of the invention provides the following for reducing the risk of host adaptation and resistance to the array:-

The CRISPR array or vector of the invention for modifying a target nucleotide sequence of a host cell,

a. wherein the host cell comprises a first endogenous promoter (first host promoter) for transcription of the target sequence;

b. wherein the CRISPR array comprises a sequence encoding a crRNA and a first promoter for transcription of the crRNA, the crRNA being optionally comprised by a single guide RNA (gRNA) and capable of hybridising to the host target sequence to guide Cas to the target in the host cell to modify the target sequence;

c. wherein the sequence of the first promoter is the sequence of a second endogenous host promoter that is different to the sequence of the first host promoter.

[0340] In an example, a promoter is used for each vector (eg, phage) CRISPR unit that is a promoter of an essential gene in the host - that way the host will express the crRNA well (and constitutively if the promoter is from a host gene that must always or often be switched on). The host will not easily adapt away from that promoter so will not easily gain resistance. Optionally it is possible to use different essential promoters for different vector CRISPR units to decrease the chance of host adaptation (resistance). One can use the promoter of the virulence or essential or resistance gene being targeted in the host by the array (or a different array). To gain resistance to the phage the host would need to mutate the endogenous gene promoter and the gene targeting site (which may, for example, be in an coding sequence that is essential for cell growth, viability or anti-host drug (eg, antibiotic) resistance) and thus risk inactivating the gene that way too.

[0341] The provision as per the invention of multiple copies of nucleic acid sequences encoding crRNAs, wherein the copies comprise the same spacer sequence for targeting a host cell sequence as per the invention is advantageous for reducing the chances of host removal (eg, by host cell homologous recombination) of useful targeting spacers from the vector. Multiple targeting spacers can be provided flexibly, on the same or multiple HM-arrays of the invention to provide alternative ways of evading resistance.

[0342] Thus, the disclosure provides the following concepts:-

In a first concept, there is provided a host modifying (HM) CRISPR/Cas system (eg, Type I, II or III) for modifying a target

nucleotide sequence of a host cell, the system comprising components according to (i) to (iv):-

(i) at least one nucleic acid sequence encoding a Cas nuclease (eg, a Cas9);
(ii) an engineered host modifying (HM) CRISPR array (eg, an array as described above) comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to a host target sequence to guide Cas to the target in the host cell to modify the target sequence;
(iii) an optional tracrRNA sequence or a DNA sequence for expressing a tracrRNA sequence;
(iv) wherein said components of the system comprises two, three or more of copies (eg, 2, 3, 4, 5, 6, 7, 8, 9,10, 20, 30, 40, 50 or more); of nucleic acid sequences encoding crRNAs, wherein the copies comprise the same spacer sequence for targeting a host cell sequence (eg, a host virulence, resistance or essential gene sequence or a sequence of a host CRISPR/Cas system component that mediates vector adaptation).

[0343] For example, the system comprises 4 or more; or 5 or more; of said copies of nucleic acid sequences encoding crRNAs comprising the same spacer. This is advantageous to increase the expression of desired cRNAs in the host. Additionally, this provides greater chance of avoiding host resistance as more than one sequence will need to be targeted (especially if there are may copies such as 5, 10, 15, 20, 30, 40, 50 or 100 or more). Distribution of the copies over different arrays, eg, the vector comprises these spaced on the same DNA strand, is useful to reduce the chances of recombination between spacers or between flanking repeats which could then lead to excision of the desired cRNA-encoding sequences. The chances of the host excising all copies is reduced by providing copies distributed across many vector arrays, it is also reduced by including many copies of the desired spacers (eg, many copies in a first vector array and many copies in a second vector array - it is possible to include at least 2, 3, 4, 5, 6, 10 or more such arrays, each comprising 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50 or 100 or more copies of the desired spacer).

[0344] A second concept relates to the system of the first concept, wherein said components of the system comprises 4, 5, 10, 15 or 20 more of said copies of nucleic acid sequences encoding crRNAs comprising the same spacer.

[0345] A third concept relates to the system of the first or second concept, wherein the copies are split between two or more nucleic acid vector CRISPR arrays.

[0346] A fourth concept relates to the system of the third concept, wherein the system comprises first and second HM-arrays, wherein first and second vector CRISPR arrays are contained in the same host cell or in the same vector (eg, a plasmid, virus, virion, phage, phagemid or prophage).

[0347] A fifth concept relates to the system of the third or fourth concept, wherein the first array is contained in a first vector and the second array is contained in a second vector which does not contain the first array (eg, wherein the vectors are plasmids or virions (eg, of the same virus type) or phagemids (eg, of the same phage type).

[0348] A sixth concept relates to the system of any preceding concept, wherein the repeats are identical to repeats in a host CRISPR array.

[0349] A seventh concept relates to the system of any one of the first to fifth concepts, wherein the repeats are not identical to repeats in a host CRISPR array.

[0350] An eighth concept relates to a host cell comprising a system, vector, virus, virion, phage, phagemid or prophage according to any preceding concept.

[0351] A ninth concept relates to an antimicrobial composition (eg, an antibiotic, eg, a medicine, disinfectant or mouth-wash), comprising a system, vector, virus, virion, phage, phagemid or prophage according to any one of the first or eighth concepts.

[0352] Any of the general features (see below) also may apply to the present concepts.

SPLIT CRISPR/CAS9 SYSTEM

[0353] This configuration is advantageous to free up space in target vectors, for example viruses or phage that have restricted capacity for carrying exogenous sequence. By freeing up space, one is able to include more targeting spacers or arrays, which is useful for evading host resistance. It is advantageous, for example to harness the endogenous Cas endonuclease rather than encode it in the vector - especially for bulky Cas sequences such as sp or saCas9. Additionally, there is not chance of inferior compatibility as may be seen with some exogenous Cas from non-host sources. The ability to reduce virus, eg, phage genome size, may also be beneficial for promoting host cell uptake (infection and/or maintenance of the virus in host cells). In some examples, an advantage is that invasion of the host by the vector (eg, phage) may upregulate host CRISPR/Cas activity, including increased expression of host Cas nucleases - in an attempt of the host to combat invading nucleic acid. This, however, is also useful to provide endogenous Cas for use with the arrays, vectors, systems and other aspects of this configuration disclosure when these comprise one or more repeats that are recognised by the host Cas. In the case where the invention involves one or more spacers targeting a host CRISPR array, this then promotes inactivation of the host CRISPR array itself, akin to a "suicidal" host cell which then uses its own Cas nuclease to inactivate its own CRISPR systems.

[0354] Thus, the disclosure provides the following features, numbered as positions:-

In a first position, there is provided a host modifying (HM) CRISPR/Cas9 system (eg, Type I, II or III) for modifying a target nucleotide sequence of a host cell, the system comprising components according to (i) to (iv):-

(i) at least one nucleic acid sequence encoding a Cas nuclease (eg, a Cas9);
(ii) an engineered host modifying (HM) CRISPR array (eg, an array of the invention described above) comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to a host target sequence to guide said Cas to the target in the host cell to modify the target sequence;
(iii) an optional tracrRNA sequence or a DNA sequence for expressing a tracrRNA sequence;
(iv) wherein said components of the system are split between the host cell and at least one nucleic acid vector that can transform the host cell, whereby the HM-crRNA guides Cas to the target to modify the target sequence in the host cell.

[0355] By "split" here it is meant that the vector comprises one or more (but not all) of the components of the system and the host cell comprises one or more (but not all) of the components, and the vector comprises one or more components that are not comprised by the host cell. In an embodiment, the vector and host cell do not share in common any of the components, eg, the host cell comprises component (i) and the vector comprises component (ii), and either the vector comprises component (iii) and/or the host cell comprises component (iii). When the vector is inside the host cell (eg, as an integrated or episomal vector, eg, a prophage), it is intended that the vector is the nucleic acid that has been provided by a vector that has transformed the host cell (and components of the system provided by such nucleic acid are not in that case be construed as host cell components). This can readily be determined by sequencing of nucleic acid (eg, chromosome and episomal nucleic acid) of the transformed host and comparing this against the sequences from a non-transformed host of the same type (eg, from the same host parental colony or clone, eg, when the host is a microbe, eg, a bacterium or archaeon).

[0356] Optionally, the system is a CRISPR/Cas9 system. Optionally, the nuclease of (a) is a Type I Cas nuclease. Optionally, the nuclease of (a) is a Type II Cas nuclease (eg, a Cas9). Optionally, the nuclease of (a) is a Type III Cas nuclease.

[0357] A second position relates to the system of the first position, wherein at least one of the components is endogenous to the host cell.

[0358] A third position relates to the system of the first or second position, wherein component (i) is endogenous to the host cell.

[0359] A fourth position relates to the system of any one of the first to third positions, wherein component (iii) is endogenous to the host cell.

[0360] A fifth position relates to a host modifying (HM) CRISPR/Cas system (eg, Type I, II or III) for modifying a target nucleotide sequence of a host cell, the system comprising components according to (a) to (e):-

a. at least one nucleic acid sequence encoding a Cas nuclease (eg, a Cas9);
b. an engineered host modifying (HM) CRISPR array comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to a host target sequence to guide said Cas to the target in the host cell;
c. an optional tracrRNA sequence or a DNA sequence for expressing a tracrRNA sequence;
d. wherein said components of the system are split between at least a first and a second nucleic acid vector, wherein at the first vector comprises component (a) but the second vector lacks component (a); and
e. wherein the vectors can co-transform simultaneously or sequentially the host cell, whereby the HM-crRNA guides Cas to the target to modify the target sequence in the host cell.

[0361] The definition of "split" provided above applies *mutatis mutandis* to the present example comprising first and second vectors.

[0362] In an embodiment a tracrRNA sequence is not provided by the vectors, but is a tracrRNA sequence of an endogenous host cell CRISPR/Cas system, wherein the tracrRNA is capable of hybridising with the HM-crRNA in the cell for subsequent processing into mature crRNA for guiding Cas to the target in the host cell.

[0363] A sixth position relates to the system of the fifth position, wherein the first vector comprises component (a) and the second vector comprises components (b) and (c).

[0364] A seventh position relates to the system of the fifth or sixth position, wherein the first and/or second vector each comprises one, two, three or more further engineered HM-CRISPR-arrays.

[0365] An eighth position relates to the system of any one of the fifth to seventh position, wherein one of the first and

second vectors is a phagemid and the other vector is a helper phage.

[0366] A ninth position relates to the system of any preceding position (eg, third or sixth positione), wherein the crRNA sequence and tracrRNA sequence are comprised by a single guide RNA (gRNA), eg provided by the vector.

[0367] A tenth position relates to the system of any preceding position, wherein each vector has a restricted capacity for insertion of exogenous nucleic acid.

[0368] An eleventh position relates to the system of any preceding position, wherein the vector or vectors are viruses (eg, virions, packaged phage, phagemid or prophage).

[0369] A twelfth position relates to the system of any preceding position, wherein the host cell comprises a deoxyribonucleic acid strand with a free end (HM-DNA) encoding a HM-sequence of interest and/or wherein the system comprises a sequence encoding the HM-DNA (eg, integrated in the vector or in the host cell genome or an episome thereof), wherein the HM-DNA comprises a sequence or sequences that are homologous respectively to a sequence or sequences in or flanking the target sequence.

[0370] The strand comprises a free end, ie, an end not integrated into the host or vector DNA such that the strand has one or two free ends, ie, the DNA is unbonded to a neighbouring nucleotide immediately 5' and or 3' repectively.

[0371] A thirteenth position relates to the system of the twelfth position, wherein the target site is cut in the host cell by Cas (eg, by Cas9 when said Cas nuclease is a Cas9), and the HM-DNA comprise first and second sequences that are homologous 5' and 3' respectively flanking the cut for inserting the HM-DNA into the host genome (eg, into a chromosomal or episomal site).

[0372] A fourteenth position relates to the system of the thirteenth position, wherein the insertion is by homology directed recombination (HDR).

[0373] A fifteenth position relates to the system of the thirteenth position, wherein the insertion is by non-homologous end joining (NHEJ).

[0374] A sixteenth position relates to the system of any one the twelfth to fifteenth positions, wherein the HM-sequence is or encodes a regulatory element (eg, a promoter, eg, an inducible promoter that replaces an endogenous promoter), a transcription inhibiting sequence, a transcription enhancing sequence, a label, or a sequence that encodes an exogenous protein or domain.

[0375] A seventeenth position relates to the system of any one of the twelfth to sixteenth positions, wherein the system comprises first and second HM-DNAs wherein a sequence of the first HM-DNA is complementary to a sequence of the second DNA whereby the DNAs are able to combine in the host cell by homologous recombination to form a combined HM-DNA for insertion into the host cell genome (eg, into a chromosomal or episomal site).

[0376] An eighteenth position relates to the system of any preceding position, wherein the vector or vectors are capable of infecting the host cell to introduce vector nucleic acid comprising a system component into the cell.

[0377] A nineteenth position relates to the system of any preceding position, wherein said Cas nuclease is a nickase.

[0378] A twentieth position relates to the system of any preceding position, wherein the cell is a bacteria or archaea and said Cas nuclease is provided by an endogenous Type II CRISPR/Cas system of the bacteria or archaea.

[0379] A twenty-first position relates to the system of any preceding position, wherein the vector or vectors are inside a said host cell, optionally integrated into a host DNA.

[0380] A twenty-second position relates to the system of any preceding position, wherein the vector or vectors lack a Cas nuclease (eg, aCas9)-encoding sequence.

[0381] A twenty-third position relates to an engineered nucleic acid viral vector (eg, a vector, virion or packaged phage as described above) for infecting a microbe host cell comprising an endogenous CRISPR/Cas system, the vector

(a) comprising nucleic acid sequences for expressing a plurality of different crRNAs for use in a CRISPR/Cas system according to any preceding position; and

(b) lacking a nucleic acid sequence encoding a Cas nuclease (eg, a Cas9), wherein a first of said crRNAs is capable of hybridising to a first nucleic acid sequence in said host cell; and a second of said crRNAs is capable of hybridising to a second nucleic acid sequence in said host cell, wherein said second sequence is different from said first sequence; and

(c) the first sequence is comprised by an anti-microbe (eg, antibiotic) resistance gene (or RNA thereof) and the second sequence is comprised by an anti-microbe resistance gene (or RNA thereof); optionally wherein the genes are different;

(d) the first sequence is comprised by an anti-microbe resistance gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof);

(e) the first sequence is comprised by an essential gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof); or

(f) the first sequence is comprised by a virulence gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof).

[0382] A twenty-fourth position relates to an engineered (directly engineered or isolated from a vector in a host cell, where that vector was derived from an engineered vector that transformed the host) nucleic acid vector for transforming a host cell comprising an endogenous CRISPR/Cas system, the vector optionally being a vector as described above and

(a') comprising nucleic acid sequences for expressing a plurality of different crRNAs for use in a CRISPR/Cas system according to any preceding position; and
(b') lacking a nucleic acid sequence encoding a Cas nuclease (eg, a Cas9),

wherein a first of said crRNAs is capable of hybridising to a first nucleic acid sequence in said host cell; and a second of said crRNAs is capable of hybridising to a second nucleic acid sequence in said host cell,
wherein said second sequence is different from said first sequence; and the first and/or second sequence is a target sequence of the host CRISPR/Cas system which sequence is or comprises

(c') a repeat DNA or RNA sequence (eg, wherein the repeat is the 5'-most repeat (the first repeat) in said host CRISPR array;
(d') a tracrRNA sequence or a tracrRNA-encoding DNA sequence;
(e') a CRISPR array leader sequence;
(f') a Cas gene promoter (eg, a Cas1, Cas2 or Csn2 promoter);
(g') a CRISPR array leader promoter sequence; or
(h') a Cas-encoding DNA or RNA sequence (eg, wherein the Cas is Cas9, Cas1, Cas2 or Csn2), eg, wherein a first of said crRNAs is capable of targeting a host Cas1 gene sequence (or a sequence of an RNA thereof) and a second of said crRNAs is capable of targeting a host Cas2 gene sequence (or a sequence of an RNA thereof).

[0383] A twenty-fifth position relates to the vector of the twenty-fourth position, wherein the first and/or second target sequence is or comprises

i. a CRISPR array leader or leader promoter sequence contiguous with the 5'-most nucleotide of the first repeat (and optionally comprising said 5'-most nucleotide of the repeat), eg, comprising the first 3 nucleotides at the 5' end of the first repeat;
ii. a sequence of up to 20 (eg, 3, 5, 7, 9, 10, 12, 15, 20, 30 or 32) contiguous nucleotides immediately 5' of the first repeat;
iii. a sequence of up to 20 (eg, 3, 5, 7, 9, 10, 12, 15, 20, 30 or 32) contiguous nucleotides of the 5'-most nucleotides of the first repeat; or
iv. a sequence of up to 20 (eg, 3, 5, 7, 9, 10, 12, 15, 20, 30 or 32) contiguous nucleotides immediately 3' of the first spacer (and optionally wherein the sequence comprises the 3'-most nucleotide of the first spacer), eg, comprising the last 3 nucleotides at the 3' end of the first repeat.

[0384] A twenty-sixth position relates to the vector of the twenty-fourth or twenty-fifth position, wherein the or each target sequence is comprised by a sequence selected from the group consisting of SEQ ID NO: 1 to 44, or a complement thereof.

[0385] A twenty-seventh position relates to the vector of any one of the twenty-fourth to twenty-sixth positions, wherein the first crRNA comprises or consists of the structure R-S-R, wherein R=a CRISPR repeat and S=a CRISPR spacer, wherein S comprises, (in 5' to 3' direction) V-H_R or H_R-V or, wherein V=a sequence identical to a DNA sequence of the vector and H_R=a DNA sequence of a repeat of a CRISPR array of said host cell CRISPR/Cas system, wherein the first crRNA is capable of hybridising to a spacer of the host CRISPR array to guide Cas to the target of the crRNA for modification of the host CRISPR array in the cell.

[0386] A twenty-eighth position relates to the vector of the twenty-seventh position, wherein the first crRNA does not substantially hybridise to the nucleic acid present in the vector, eg, wherein the first crRNA does not hybridise to V in the vector or hybridises less strongly than it hybridises to the spacer of the host array. The discussion above on determining this applies to this position too.

[0387] A twenty-ninth position relates to the vector of the twenty-seventh or twenty-eighth position, wherein V=one or up to 40 (eg, up to 15) contiguous nucleotides of vector DNA. For example, V=1, 2, 3, 4, 5, 6, 7 8 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 or 40 contiguous nucleotides of vector DNA.

[0388] A thirtieth position relates to the vector of the twenty-ninth position, wherein

i. the host CRISPR/Cas system is able to recognise a cognate PAM;
j. wherein the vector DNA comprises such a PAM immediately 3' of a protospacer sequence;
k. wherein V=one or up to 40 (eg, up to 15) nucleotides of the protospacer; and

l. wherein H_R=a sequence identical to a contiguous sequence of the repeat of the host CRISPR array.

[0389] A thirty-first position relates to the vector of the thirtieth position, wherein said contiguous sequence of the repeat of the host array is a sequence of at least 50% of a host repeat (eg, including the 5'-most or 3'-most nucleotide of the host repeat).

[0390] A thirty-second position relates to the vector of the thirtieth or thirty-first position, wherein V=from 1 to 40 (eg, up to 15) of the 3'-most protospacer contiguous nucleotides; and optionally said contiguous sequence of the repeat includes the 5'-most nucleotide of the host repeat.

[0391] A thirty-third position relates to the vector of the thirtieth or thirty-first position, wherein V=from 1 to 40 (eg, up to 15) of the 5'-most protospacer contiguous nucleotides; and optionally said contiguous sequence of the repeat includes the 3'-most nucleotide of the host repeat.

[0392] A thirty-fourth position relates to the vector of any one of the twenty-seventh to thirty-third positions, wherein R=a repeat that is recognised by the host CRISPR/Cas system. Alternatively, R=a repeat that is not recognised by the host CRISPR/Cas system. In this case, preferably the vector comprises a nucleotide sequence of a Cas nuclease (and optionally a tracrRNA) that is cognate to R, ie, is capable of functioning with R in the host cell.

[0393] A thirty-fifth position relates to a vector according to any one of the twenty-fourth to thirty-fourth positions, wherein the first sequence is according to any one of (c') to (h') and the second sequence is selected from a host essential gene, virulence gene or resistance gene.

[0394] A thirty-sixth position relates to an engineered nucleic acid viral vector (eg, a virion or packaged phage) for use in the system of any one of the first to twenty-second positions for infecting a microbe host cell comprising an endogenous CRISPR/Cas system,

a. the vector comprising a first nucleic acid sequence for expressing a first crRNA in the host; and
b. wherein the first sequence comprises (in 5'to 3' direction) R1a-S1-R1b, wherein R1a=a first CRISPR repeat, wherein R1a is optional; R1b=a second CRISPR repeat and S1=a CRISPR spacer complementary to a host sequence (eg, a host sequence recited in position 23 or 24), wherein R1a and R1b are recognised by a host Cas nuclease (eg, a Type II nuclease, eg, a Cas9);
c. wherein the vector lacks (i) a nucleic acid sequence encoding a Cas nuclease (eg, a Cas9) that recognises the repeat(s) of (b) and/or (ii) a nucleic acid sequence encoding a tracrRNA sequence that is complementary to a crRNA sequence encoded by the first sequence.

[0395] For example, the vector is a nucleic acid vector comprised by a phage.

[0396] A thirty-seventh position relates to the vector of the thirty-sixth position, wherein

d. the vector comprises a second nucleic acid sequence for expressing second crRNA in the host, wherein the second crRNA is different from the first crRNA;
e. wherein the second sequence comprises (in 5'to 3' direction) R2a-S2-R2b, wherein R2a=a first CRISPR repeat, wherein R2a is optional; R2b=a second CRISPR repeat and S2=a CRISPR spacer complementary to a host sequence (eg, a host sequence recited in position 23 or 24), wherein R2a and R2b are recognised by a host Cas nuclease (eg, a Type I or II nuclease, eg, a Cas6).

[0397] Thus, for example, the first and second nucleic acid sequences are comprised by the same packaged phagemid, eg, in the same or different CRISPR arrays.

[0398] A thirty-eighth position relates to the vector of the thirty-seventh position, wherein the vector lacks (iii) a nucleic acid sequence encoding a Cas (eg, a Cas6) that recognises the repeat(s) of (e) and/or (iv) a nucleic acid sequence encoding a tracrRNA sequence that is complementary to a crRNA sequence encoded by the second sequence.

[0399] A thirty-ninth position relates to a collection of engineered nucleic acid viral vectors (eg, vectors, virions or packaged phages as described above) for use in the system of any one of the first to twenty-second positions for co-infecting a microbe host cell comprising an endogenous CRISPR/Cas system, the collection comprising a first vector and a second vector,

f. wherein the first vector is according to the thirty-sixth position;
g. wherein the second vector comprises a second nucleic acid sequence for expressing second crRNA in the host, wherein the second crRNA is different from the first crRNA;
h. wherein the second sequence comprises (in 5'to 3' direction) R2a-S2-R2b, wherein R2a=a first CRISPR repeat, wherein R2a is optional; R2b=a second CRISPR repeat and S2=a CRISPR spacer complementary to a host sequence, wherein R2a and R2b are recognised by a host Cas nuclease (eg, a Type I or II nuclease, eg, a Cas6).

[0400] For example, the first vector is comprised by a first packaged phagemid and the second vector is comprised by a second packaged phagemid.

[0401] A fortieth position relates to the collection of the thirty-ninth position, wherein the second vector comprises (v) a nucleic acid sequence encoding a Cas (eg, a Cas9) that recognises the repeat(s) of (b) and/or (vi) a nucleic acid sequence encoding a tracrRNA sequence that is complementary to a crRNA sequence encoded by the first sequence.

[0402] For example, in this case the Cas functions are provided by the endogenous host system. This saves vector space (eg, for inclusion of more host-targeting HM-array spacers) and simplifies vector and array construction.

[0403] A forty-first position relates to the collection of the thirty-ninth or fortieth position, wherein the second vector lacks (vii) a nucleic acid sequence encoding a Cas (eg, a Cas6) that recognises the repeat(s) of (h) and/or (viii) a nucleic acid sequence encoding a tracrRNA sequence that is complementary to a crRNA sequence encoded by the second sequence.

[0404] For example, in this case the Cas functions are provided by the endogenous host system.

[0405] A forty-second position relates to the collection of the thirty-ninth position, wherein the first and second vectors each lacks (ix) a nucleic acid sequence encoding a Cas (eg, a Cas9) that recognises the repeat(s) of (b) and (x) a nucleic acid sequence encoding a Cas (eg, a Cas6) that recognises the repeat(s) of (h); optionally wherein the collection is comprised by a host cell comprising one or more Cas that recognise the repeat(s) of (b) and (h).

[0406] A forty-third position relates to the collection of the forty-second position, further comprising a third vector (eg, a virion or a phage) comprising a nucleic acid sequence according to (ix) and/or (x).

[0407] A forty-fourth position relates to the collection of any one of the thirty-ninth to forty-third positions, wherein each vector is comprised by a respective packaged virion or phagemid, or a respective virion or phage nucleic acid.

[0408] A forty-fifth position relates to the vector or collection of any one of the thirty-sixth to forty-fourth positions, wherein R1a and R1b comprise the same repeat sequence.

[0409] A forty-sixth position relates to the vector or collection of any one of the thirty-seventh to forty-fifth positions, wherein R2a and R2b comprise the same repeat sequence.

[0410] A forty-seventh position relates to the vector or collection of any one of the thirty-seventh to forty-sixth positions, wherein the repeat(s) of (b) are recognised by a Cas nuclease that is different from the Cas nuclease that recognises the repeat(s) of (e).

[0411] A forty-eighth position relates to the vector or collection of any one of the thirty-seventh to forty-seventh positions, wherein the host comprises CRISPR/Cas systems of different types (eg, a Type I and a

[0412] Type II system; a Type I and a Type III system; a Type II and a Type III system; or Type I, II and III systems).

[0413] A forty-ninth position relates to the vector or collection of any one of the thirty-sixth to forty-eighth positions, wherein the repeat(s) of (b) are recognised by a Type II Cas nuclease, eg, a Cas9.

[0414] A fiftieth position relates to the vector or collection of any one of the thirty-seventh to forty-ninth positions, wherein the repeat(s) of (e) are recognised by a Type I or III Cas nuclease, eg, a Cas6.

[0415] A fifty-first position relates to the vector or collection of any one of the twenty-third to fiftieth positions, wherein the vector is a virus, a virion, phage, phagemid or prophage.

[0416] A fifty-second position relates to the vector or collection of any one of the twenty-third to fifty-first positions inside a host cell comprising one or more Cas that are operable with cRNA encoded by the vector(s).

[0417] A fifty-third position relates to the vector or collection of any one of the twenty-third to fifty-second positions inside a host cell comprising a Cas9.

[0418] A fifty-fourth position relates to the vector or collection of any one of the twenty-third to fifty-third examples, in combination with a HM-DNA (eg, integrated in the vector, on a plasmid or in the host cell genome or an episome thereof), wherein the HM-DNA is as recited in any of the twelfth to seventeenth positions.

[0419] A fifty-fifth position relates to the system, vector or collection of any preceding position, comprising nucleic acid sequences for expressing a plurality of different crRNAs, wherein said crRNAs are capable of targeting at least 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50 or 100 DNA sequences in the host cell.

[0420] A fifty-sixth position relates to the system, vector or collection of any preceding position, comprising a first crRNA or a nucleic acid sequence encoding a first cRNA that is capable of targeting a DNA sequence of a Cas nuclease (or sequence of an RNA thereof) which is not said Cas nuclease (eg, Cas9) but which mediates host vector adaptation; optionally comprising a second crRNA or a nucleic acid sequence encoding a second cRNA that is capable of targeting a sequence of a resistance, virulence or essential host gene (or RNA thereof) in the host.

[0421] A fifty-seventh position relates to the system, vector or collection of any preceding position, comprising two, three or more of copies of nucleic acid sequences encoding crRNAs, wherein the copies comprise the same spacer sequence for targeting a host cell sequence (eg, a virulence, resistance or essential gene sequence or a sequence of a host CRISPR/Cas system component that mediates vector adaptation, but which is not said Cas nuclease).

[0422] A fifty-eighth position relates to the system, vector or collection of the fifty-seveneth position, wherein the copies are split between two or more vector CRISPR arrays.

[0423] A fifty-ninth position relates to the system, vector or collection of any preceding position, wherein the vector

repeats are identical to repeats in a or the host CRISPR array (eg, each vector repeat has at least 95% sequence identity to a host repeat).

[0424] A sixtieth position relates to the system, vector or collection of any one of the first to fifty-eighth positions, wherein the vector repeats are not identical to repeats in a or the host CRISPR array.

[0425] A sixty-first position relates to the system, vector or collection of any preceding position, comprising first and second vector CRISPR arrays which are contained in the same host cell or by the same vector (eg, plasmid or virus or virion or phage or prophage or phagemid).

[0426] A sixty-second position relates to the system, vector or collection of the sixty-first position, wherein the first array is contained in a first vector and the second array is contained in a second vector which does not contain the first array (eg, wherein the vectors are plasmids or virions (eg, of the same virus type) or phagemids (eg, of the same phage type).

[0427] A sixty-third position relates to a host cell comprising a system, vector, collection, virus, virion, phage, phagemid or prophage according to any preceding position.

[0428] A sixty-fourth position relates to an antimicrobial composition (eg, an antibiotic, eg, a medicine, disinfectant or mouthwash), comprising a system, vector, virus, virion, phage, phagemid or prophage according to any one of the first to sixty-second positions.

CONDITIONING MICROBES TOGETHER

[0429] The disclosure provides for methods of producing microbes (eg, phage and/or bacterial populations) that involves conditioning hosts and viruses together to facilitate co-evolution and thus conditioning of the hosts to the viruses (eg, phage) and *vice versa.* Using repressible control of crRNA expression or activity the disclosure purposely modulates the co-evolution in a controllable manner where a desired spacer activity can be toggled on or off to enable tuning to occur with or without stress imposed by spacer-guided Cas action in the host, eg, with or without antibiotic resistance gene targeting. In this way, the bacterial populations can be tuned for use in situations (eg, dairy or food production cultures) where phage inactivation of desirable genes may be encountered; or for use in tuning phage to be used to kill or modulate bacteria, eg, to knock-down antibiotic resistance. This configuration further enables, in one embodiment, culturing of antibiotic-resistant bacterial host with virus, eg, phage, harbouring one or more CRISPR arrays of the invention that target the antibiotic resistance gene of the host, since the method purposely represses the antibotic resistance gene inactivation activity of the array during culturing with the host. Thus, a resistant bacterial host population can be used to grow up phage in culture (eg, in an industrial culture vessel or plant) allowing the phage and host to co-evolve and mutually tune without the antibiotic resistance inactivation effect hampering the growth and thus culturing ability of the host cells (which would otherwise minimise phage expansion) and whilst still enabling all other components of the desired phage to tune to the cultured host population. Testing of a sample of the resultant phage population can be carried out, eg, at lab scale, using an antibotic resistant host celll population but with the test phage de-repressed for the array targeting of the antibiotic resistance gene of the host cells. Naturally-occurring and synthetic repression of gene expression in prokaryotic cell and phage settings is well known to the skilled person, eg, *tet* systems or light-inducible systems.

[0430] Thus, the disclosure provides the following features, numbered as paragraphs:-

In a first paragraph, there is provided a microbe production method, the method comprising

(a) providing a host cell that comprises a host CRISPR/Cas system for nucleotide sequence targeting in the host cell;
(b) providing a virus that is capable of infecting the host cell, wherein

(i) the virus comprises one or more engineered host modifying (HM) CRISPR arrays (eg, an array as described above) for modifying target nucleotide sequences of the host cell;
(ii) a first said HM-array encodes a first HM-crRNA comprising a spacer sequence (HM-spacer) that is capable of hybridising to a first host target sequence to guide Cas to the target in the host cell to modify the target sequence, optionally wherein the modification of the first target sequence reduces host cell growth or viability; and
(iii) the first HM-array is reversibly repressible for the transcription of the first HM-crRNA and/or first HM-crRNA activity is repressible;

(c) infecting the host cell with the virus to introduce the one or more HM-CRISPR arrays into the cell;
(d) repressing the transcription of the first HM-crRNA and/or first HM-crRNA activity in the cell;
(e) culturing the infected host cell to produce a population (PH1) of host cells comprising a population (PV1) of virus; and
(f) obtaining the virus population PV1 and/or the cultured host cell population.

[0431] In an example, the first HM-crRNA comprises a HM-spacer that is capable of hybridising to the first host target

sequence to guide Cas to the target in the host cell to modify the target sequence, wherein the target sequence is a nucleotide sequence of the host CRISPR/Cas system, whereby the first HM-crRNA guides Cas to the target to modify the host CRISPR/Cas system in the host cell, wherein the modification of the target sequence reduces or eliminates functioning of the host CRISPR/Cas system.

[0432] In an alternative, the modification enhances or inhibits epression of a gene in the host. In an embodiment, the gene is an essential gene, virulence gene or resistance gene (eg, an antibiotic resistance gene). In an embodiment, the modification enhances the expression of a gene product that is endogenous or exogenous to the host. In an example, the host is an engineered host comprising an exogenous nucleotide sequence (eg, for producing a desired protein) and the modification enhances or inhibits expression of the desired protein in the host cell. In an example, the desired protein is an antibiotic and host cell is a microbe, eg, bacterial or archaeal cell. Thus, the method enables culturing of culturing of host cells to produce the viral population, wherein the antibotic is not expressed which would otherwise hamper the expansion of the host cell population. Thereafter, one or more viruses of the the isolated virus population can be used in an antimicrobial composition for reducing host cell growth or viability, since the first HM-crRNA repression can be removed after isolation, thereby providing an actively antibiotic virus composition. The disclosure therefore also provides such a method and such an antibiotic composition comprising virus that are capable of expressing an antibiotic in a host cell. Modification to activate the expression can be effected, for example, by providing a Cas (eg, Cas9) conjugated to a transcription activator, wherein the Cas is a cognate Cas for the first HM-crRNA and the activator activates the transcription of the desired exogenous or endogenous gene. Modification to inhibit the expression can be effected, for example, by providing a dead Cas (eg, dCas9), wherein the CAs is a cognate Cas for the first HM-crRNA and inhibits transcription of the desired exogenous or endogenous gene.

[0433] Repression of the crRNA transcription or activity can be partial or complete (ie, no activity or no transcription of the crRNA from the array in the host). Activity refers to the ability of the crRNA to hybridise to the cognate host sequence for guiding of Cas to the first host target site for modification.

[0434] In an example, the virus is not so repressed when introduced into the cell, the method comprising carrying out step (d) after the virus has infected the cell, eg, by using a chemical, physical, mechanical, magnetic, light or other agent to cause repression. In an embodiment, the first HM-array comprises a repressible promoter (HM-promoter) for transcription of the first HMcrRNA and the promoter is repressed (eg, by binding a repressor agent, eg, a chemical or protein, to the promoter) after the first HM-array is introduced into the cell.

[0435] In another example, the virus is so repressed before step (c) is carried out, eg, by using a chemical, physical, mechanical, magnetic, light or other agent to cause repression. In an embodient, the first HM-array comprises a repressible promoter (HM-promoter) for transcription of the first HMcrRNA and the promoter is repressed (eg, by binding a repressor agent, eg, a chemical or protein to the promoter) before the first HM-array is introduced into the cell, wherein subsequently the repressed first HM-array is introduced into the cell.

[0436] In one embodiment, step (f) comprises isolating PV1. In an embodiment, the step comprised separating PV1 or a virus thereof from host cells of PH1.

[0437] A second paragraph relates to the method of the first paragraph 1, further comprising de-repressing the transcription of first HM-crRNA and/or first HM-crRNA activity in the virus population after step (e) or (f), and optionally thereafter further culturing the host cells.

[0438] A third paragraph relates to the method of any preceding paragraph, comprising

A. obtaining a population (PH2) of host cells that are optionally identical to the host cell of (a), (f) or the further cultured cells of paragraph 2;
B. infecting the host cells of A with virus from the population PV1;
C. repressing the transcription of the first HM-crRNA and/or first HM-crRNA activity in the cells;
D. culturing the infected host cells to produce a population (PH3) of host cells comprising a population of virus (PV2); and
E. obtaining the virus population PV2 (or a virus thereof) and/or the cultured host cell population. A fourth paragraph relates to the method of the third paragraph, further comprising de-repressing the transcription of first HM-crRNA and/or first HM-crRNA activity in the virus population after step (D) or (E), and optionally thereafter further culturing the host cells.

[0439] A fifth paragraph relates to the method of any preceding paragraph, comprising testing an isolated sample of the virus population PV1 or PV2 on a further host cell or population (PH4) of host cells, optionally wherein the further cell or population PH4 is identical to the cell of (a), the testing comprising infecting the further cell or population PH4 with virus of said sample, waiting a period of time to allow any host cell growth to occur, and determining if a predetermined activity of the further cell or population PH4 (eg, cell growth or viability) has been modified (eg, reduced, such as reduced host cell growth or viability*) or occurred, wherein virus inside the cell or cells have de-repressed transcription of first HM-crRNA and/or first HM-crRNA activity during said period of time.

[0440] * This can be tested using a standard assay for plaque formation when the virus of the sample are added to the cell or PH4 plated on agar).

[0441] A sixth paragraph relates to the method of any preceding the fifth paragraph, wherein all of the host cells are microbial cells (eg, bacterial or archaeal cells) and the modification of the first target sequence reduces host cell growth or viability, and said determining determines that antimicrobial activity** has occurred.

[0442] **This can be determined using a standard plaque assay.

[0443] A seventh paragraph relates to the method of the fifth or sixth paragraph, wherein the period of time is at least one,5, 10, 30, 60 or 120 minutes.

[0444] An eighth paragraph relates to the method of any one of the fifth to seventh paragraphs, wherein the cell of (a) and optionally PH1, PH2 and/or PH3 cells do not comprise the first target sequence, wherein the further cell or population PH4 cells comprise the first target sequence.

[0445] A ninth paragraph relates to the method of any one of the first to eighth paragraphs, wherein the cell of

(a) and optionally PH1, PH2 and/or PH3 cells do not comprise a gene that confers resistance to a first antibiotic, wherein the first target sequence is a target sequence of such a gene; optionally wherein the further cell or population PH4 cells comprise said gene.

[0446] A tenth paragraph relates to the method of any one of the first to seventh paragraphs, wherein the cell of (a) and optionally PH1, PH2 and/or PH3 cells comprise a gene that confers resistance to a first antibiotic, wherein the first target sequence is a target sequence of such a gene.

[0447] An eleventh paragraph relates to the method of any preceding paragraph, wherein all of the host cells are microbial cells (eg, bacterial or archaeal cells) and the modification of the first target sequence reduces host cell growth or viability, or reduces host cell resistance to an antibiotic.

[0448] A twelfth paragraph relates to the method of any preceding paragraph, wherein all of the host cells are infectious disease pathogens of humans, an animal (eg, non-human animal) or a plant.

[0449] A thirteenth paragraph relates to the method of any preceding paragraph, wherein all of the host cells are of the same species, eg, selected from a species of Escherichia (eg, *E coli* O157:H7 or O104: H4), Shigella (eg, *dysenteriae*), Salmonella (eg, *typhi* or *enterica,* eg, serotype typhimurium, eg, DT 104), Erwinia, Yersinia (eg, *pestis*), Bacillus, Vibrio, Legionella (eg, *pneumophilia*), Pseudomonas (eg, *aeruginosa*), Neisseria (eg, *gonnorrhoea* or *meningitidis*), Bordetella (eg, *pertussus*), Helicobacter (eg, *pylori*), Listeria (eg, *monocytogenes*), Agrobacterium, Staphylococcus (eg, *aureus,* eg, MRSA), Streptococcus (eg, *pyogenes* or *thermophilus*), Enterococcus, Clostridium (eg, *dificile* or *botulinum*), Corynebacterium (eg, *amycolatum*), Mycobacterium (eg, *tuberculosis*), Treponema, Borrelia (eg, *burgdorferi*), Francisella, Brucella, Campylobacter (*eg, jejuni*), Klebsiella (eg, *pneumoniae*), Frankia, Bartonella, Rickettsia, Shewanella, Serratia, Enterobacter, Proteus, Providencia, Brochothrix, Bifidobacterium, Brevibacterium, Propionibacterium, Lactococcus, Lactobacillus, Pediococcus, Leuconostoc, *Vibrio* (eg, *cholera,* eg, O139, or *vulnificus*), *Haemophilus* (eg, *influenzae*), *Brucella* (eg, *abortus*), *Franciscella, Xanthomonas, Erlichia* (eg, *chaffeensis*), *Chlamydia* (eg, *pneumoniae*), *Parachlamydia,* Enterococcus (eg, faecalis or faceim, eg, linezolid-resistant), Oenococcus and *Acinetoebacter* (eg, *baumannii,* eg, multiple drug resistant).

[0450] A fourteenth paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are Staphylococcus *aureus* cells, eg, resistant to an antibiotic selected from methicillin, vancomycin-resistant and teicoplanin.

[0451] A fifteenth paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are *Pseudomonas aeuroginosa* cells, eg, resistant to an antibiotic selected from cephalosporins (eg, ceftazidime), carbapenems (eg, imipenem or meropenem), fluoroquinolones, aminoglycosides (eg, gentamicin or tobramycin) and colistin.

[0452] A sixteenth paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are *Klebsiella* (eg, *pneumoniae*) cells, eg, resistant to carbapenem.

[0453] A seventeenth paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are *Streptoccocus (eg, pneumoniae* or pyogenes) cells, eg, resistant to an antibiotic selected from erythromycin, clindamycin, beta-lactam, macrolide, amoxicillin, azithromycin and penicillin.

1. An eighteenth paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are *Salmonella* (eg, serotype Typhi) cells, eg, resistant to an antibiotic selected from ceftriaxone, azithromycin and ciprofloxacin.

2. A nineteenth paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are *Shigella* cells, eg, resistant to an antibiotic selected from ciprofloxacin and azithromycin.

3. A twentieth paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are *mycobacterium tuberculosis* cells, eg, resistant to an antibiotic selected from Resistance to isoniazid (INH), rifampicin (RMP), fluoroquinolone, amikacin, kanamycin and capreomycin. A twenty-first paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are *Enterococcus* cells, eg, resistant to vancomycin.

[0454] A twenty-second paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are *Enterobacteriaceae* cells, eg, resistant to an antibiotic selected from a cephalosporin and carbapenem.

[0455] A twenty-third paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are *E. coli* cells, eg, resistant to an antibiotic selected from trimethoprim, itrofurantoin, cefalexin and amoxicillin.

[0456] A twent-fourth paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are *Clostridium* (eg, *dificile*) cells, eg, resistant to an antibiotic selected from fluoroquinolone antibiotic and carbapenem.

[0457] A twenty-fifth paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are Neisseria *gonnorrhoea* cells, eg, resistant to an antibiotic selected from cefixime (eg, an oral cephalosporin), ceftriaxone (an injectable cephalosporin), azithromycin and tetracycline.

[0458] A twenty-sixth paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are *Acinetoebacter baumannii* cells, eg, resistant to an antibiotic selected from beta-lactam, meropenem and a carbapenem.

[0459] A twenty-seventh paragraph relates to the method of the thirteenth paragraph, wherein all of the host cells are Campylobacter cells, eg, resistant to an antibiotic selected from ciprofloxacin and azithromycin. A twenty-eighth paragraph relates to the method of any preceding paragraph, wherein the host cells produce Beta (β)-lactamase.

[0460] A twenty-ninth paragraph relates to the method of any preceding paragraph, wherein the host cells are resistant to an antibiotic recited in any one of the fourteenth to twenty-seventh paragraphs.

[0461] A thirtieth paragraph relates to the method of the twenty-ninth paragraph, wherein the first target sequence is a sequence of a gene encoding a product conferring host cell resistance to said antibiotic.

[0462] A thirty-first paragraph relates to the method of any preceding paragraph, wherein the first target sequence is a sequence of an antibiotic resistance gene (ie, for conferring host cell resistance to an antibiotic eg, methicillin resistance) and/or one, more or all of the the population PH1, the population PH2, the population PH3 and the population PH4 is resistant to an antibiotic or said antibiotic (eg, an antibiotic recited in any one of the thirteenth to twenty-seventh paragraphs).

[0463] A thirty-second paragraph relates to the method of any preceding paragraph, wherein de-repressed virus of the virus population PV1 or PV2 have antimicrobial activity (eg, antibacterial activity, such as when the virus are phage); optionally wherein the host cell or cells comprise the first target sequence as recited in paragraph 30, wherein modification of the first target provides said antimicrobial activity.

[0464] A thirty-third paragraph relates to the method of any preceding paragraph when dependant from paragraph 5, wherein the cells of PH4 are resistant to an antibiotic (eg, an antibiotic recited in any one of paragraphs 13 to 27) and the cells of (a) and PH2 are not resistant to said antibiotic. This aids manufacturing of the virus for drug use, since culturing and expansion can be performed relatively safety without the risk of having to deal with antibiotic-resistant host cells (and risk of inadequante containment of these and escape from drug manufacturing plant, for example). Nevertheless, testing against PH4 can be performed in a containment lab or other facility that is set up for use of antibiotic-resistant host strains. When testing against PH4, the first HM-crRNA is de-repressed so that modificaion of the resistance gene in the host cells is possible by the HM-array of the invention.

[0465] A thirty-foruth paragraph relates to the method of any preceding paragraph, wherein the host CRISPR/Cas system is a Type I, II or III system and the target sequence is a nucleotide sequence conserved in said Type of system in at least one, two or three additional host strains or species of the same genus as the host cell of (a).

[0466] A thirty-fifth paragraph relates to the method of any preceding paragraph, wherein the virus is a phage or phagemid.

[0467] A thirty-sixth paragraph relates to the method of the thirty-fifth paragraph, wherein the virus of (b) is a Corticoviridae, Cystoviridae, Inoviridae, Leviviridae, Microviridae, Myoviridae, Podoviridae, Siphoviridae, or Tectiviridae virus.

[0468] A thirty-seventh paragraph relates to the method of the thirty-fifth to thirty-sixth paragraph, wherein the virus of (b) is a naturally occurring phage, eg, a phage induced from a cell that is of the same strain as the cell of (a).

[0469] A thrty-eighth paragraph relates to the method of the thirty-fifth, thirty-sixth or thirty-seventh paragraph, wherein the phage of (b) is a mutated phage obtained through selective pressure using a phage-resistant bacterium.

[0470] A thirty-ninth paragraph relates to the method of any preceding paragraph, wherein in (b)

(iv) said one or more HM-arrays comprise a HM-array that encodes a second HM-crRNA comprising a HM-spacer that is capable of hybridising to a second host target sequence to guide Cas to the second target in the host cell to modify the target sequence, wherein the second target sequence is a nucleotide sequence of the host CRISPR/Cas system, whereby the second HM-crRNA guides Cas to the second target to modify the host CRISPR/Cas system in the host cell, wherein the modification of the second target sequence reduces or eliminates functioning of the host CRISPR/Cas system; and

(v) wherein the HM-array of (iv) is active in the cell of (a) for the transcription of second HM-crRNA capable of hybridising to the second host target sequence.

[0471] In an embodiment, the HM-array of (ii) and (iv) are the same HM-array. In another embodiment, they are different

HM-arrays (eg, arrays of different CRISPR/Cas types, eg, Type I and II, or Type II and III, or Type I and III, or different Type II arrays).

[0472] A fortieth paragraph relates to the method of the thirty-ninth paragraph, wherein the cells of any one or all of PH1-4 comprise said second target sequence.

[0473] A forty-first paragraph relates to the method of the thirty-ninth or fortieth paragraph, wherein the second target sequence is identical to a CRISPR/Cas system sequence of a genus or species of cell as recited in any one of the eleventh to twenty-fourth paragraphs (eg, *S thermophilus or S pyogenes* or *S aureus*).

[0474] A forty-second paragraph relates to the method of any one of the thirty-ninth to forty-first paragraphs, wherein the second target sequence is comprised by a sequence selected from the group consisting of SEQ ID NO: 1 to 44, or a complement thereof.

[0475] A forty-third paragraph relates to the method of any one of the thirty-ninth to forty-second paragraphs, wherein the second target sequence comprises

A. a repeat DNA or RNA sequence (eg, wherein the repeat is the 5'-most repeat (the first repeat) in said host CRISPR array;
B. a tracrRNA sequence or a tracrRNA-encoding DNA sequence;a CRISPR array leader sequence;
C. a Cas gene promoter (eg, a Cas1, Cas2 or Csn2 promoter);
D. a CRISPR array leader promoter sequence; or
E. a Cas-encoding DNA or RNA sequence (eg, wherein the Cas is Cas9, Cas1, Cas2 or Csn2).

[0476] A forty-fourth paragraph relates to the method of any one of the thirty-ninht to forty-third paragraphs, wherein the second target sequence comprises

F. a CRISPR array leader or leader promoter sequence contiguous with the 5'-most nucleotide of the first repeat (and optionally comprising said 5'-most nucleotide of the repeat);
G. a sequence of up to 20 contiguous nucleotides immediately 5' of the first repeat;
H. a sequence of up to 20 contiguous nucleotides of the 5'-most nucleotides of the first repeat; or
I. a sequence of up to 20 contiguous nucleotides immediately 3' of the first spacer repeat (and optionally wherein the sequence comprises the 3'-most nucleotide of the first spacer).

[0477] A forty-fifth paragraph relates to the method of any one of the thirty-ninth to forty-fourth paragraphs, wherein

J. the second HM-crRNA comprises or consists of the structure R-S-R, wherein R=a CRISPR repeat and S=a CRISPR spacer, wherein S comprises, (in 5' to 3' direction) V-H_R or H_R-V or , wherein V=a sequence at least 95, 96, 97, 98 or 99% identical to a DNA sequence of the virus of (b) and H_R=a DNA sequence of a CRISPR repeat of said host cell CRISPR/Cas system;
K. wherein the sequence of H_R is immediately contiguous with the sequence of V in the host CRISPR/Cas system; and
L. wherein the second HM-crRNA is capable of hybridising to a spacer of the host CRISPR/Cas system to guide Cas to the spacer for modification (eg, cleavage or inactivation) of the host CRISPR/Cas system in the cell.

[0478] A forty-sixth paragraph relates to the method of the forty-fifth paragraph, wherein V=one or up to 40 (eg, up to 15) contiguous nucleotides of virus DNA.

[0479] A forty-seventh paragraph relates to the method of any one of the thirty-ninth to forty-sixth paragraphs, wherein the second HM-crRNA does not substantially hybridise to nucleic acid of the virus of (b).

[0480] A forty-eighth paragraph relates to the method of any one of the forty-fifth to forty-sevneth paragraphs, wherein

a. the host CRISPR/Cas system is able to recognise a cognate PAM;
b. wherein the nucleic acid of the virus of (b) comprises such a PAM immediately 3' of a protospacer sequence;
c. wherein V=one or up to 40 (eg, up to 15) nucleotides of the protospacer; and
d. wherein H_R=a sequence identical to a contiguous sequence of the repeat of the host CRISPR/Cas system.

[0481] A forty-ninth paragraph relates to the method of the forty-eighth paragraph, wherein said contiguous sequence of the repeat of the host system is a sequence of at least 50% of a host repeat (eg, including the 5'-most or 3'-most nucleotide of the host repeat).

[0482] A fiftieth paragraph relates to the method of the forty-fifth or forty-sixth paragraph, wherein V=from 1 to 40 (eg, up to 15) of the 3'-most protospacer contiguous nucleotides; and optionally said contiguous sequence of the repeat includes the 5'-most nucleotide of the host repeat.

[0483] A fifty-first paragraph relates to the method of the forty-eighth or forty-ninth paragraph, wherein V=from 1 to 40

(eg, up to 15) of the 5'-most protospacer contiguous nucleotides; and optionally said contiguous sequence of the repeat includes the 3'-most nucleotide of the host repeat.

[0484] A fifty-second paragraph relates to the method of any one of the forty-fifth to fifty-first paragraphs, wherein R=a repeat that is recognised by the host CRISPR/Cas system.

[0485] A fifty-third paragraph relates to the method of any preceding paragraph, wherein the or each HM-CRISPR comprises (in 5' to 3' direction) a first repeat sequence, a first spacer sequence and a second repeat sequence, wherein the spacer sequence comprises a sequence that is capable of hybridising to the respective target sequence in the host cell, the array further comprising a promoter for transcription of the repeats and spacer in the host cell, and optionally the nucleic acid of the virus of (b) comprises a Cas nuclease-encoding sequence and/or a tracrRNA-encoding sequence for encoding a functional Cas and/or tracrRNA sequence in the host cell, wherein the tracrRNA sequence comprises a sequence that is complementary to the first or second repeat.

[0486] A fifty-fourth paragraph relates to the method of any preceding paragraph, wherein the or each HM-CRISPR array comprises (in 5' to 3' direction) a first repeat sequence, a first spacer sequence and a second repeat sequence, wherein the spacer sequence comprises a sequence that is capable of hybridising to the respective target sequence in the host cell, the array further comprising a promoter for transcription of the repeats and spacer in the host cell, and wherein the vector does not comprise a Cas nuclease-encoding sequence and/or a tracrRNA-encoding sequence for encoding a tracrRNA sequence in the host cell wherein the tracrRNA sequence comprises a sequence that is complementary to the first or second repeat, wherein the HM-CRISPR array is functional in the host cell to guide Cas (eg, endogenous host Cas nuclease) to the respective host target site, optionally using a host tracrRNA.

[0487] A fifty-fifth paragraph relates to the method of the fity-third or fifty-fourth paragraph , wherein the repeats are identical to repeats in the host CRISPR/Cas system, wherein the or each HM-CRISPR array does not comprise a PAM recognised by a Cas (eg, a Cas nuclease, eg, Cas9) of the host CRISPR/Cas system.

[0488] A fifty-sixth paragraph relates to the method of any preceding paragraph, wherein the or each HM-CRISPR array comprises more than one copy of a HM-spacer (eg, at least 2, 3 or 4 copies).

[0489] A fifty-seventh paragraph relates to the method of any preceding paragraph, encoding a second or third HM-crRNA (further HM-crRNA), wherein the further HM-crRNA comprises a nucleotide sequence that is capable of hybridising to a host target sequence to guide Cas to the target in the host cell; optionally wherein the target sequence is a nucleotide sequence of an essential, virulence or resistance gene of the host cell, or of an essential component of the CRISPR/Cas system of the host cell.

[0490] A fifth-eighth paragraph relates to the method of any preceding paragraph, wherein the or each HM-CRISPR array comprises CRISPR repeat sequences that are identical to endogenous CRISPR repeat sequences of the host cell for producing the respective HM-crRNA in the host cell.

[0491] A fifty-ninth the method of any preceding paragraph, wherein the virus of (b) comprises a nucleotide sequence encoding a Cas (non-host Cas) that is functional in the host cell of (a) (eg, wherein the non-host Cas is a Type I system Cas wherein the host system is a Type II or III; a Type II system Cas wherein the host system is a Type I or III; or a Type III system Cas wherein the host system is a Type I or II), optionally wherein the host cell does not comprise or express a Cas of a Type that is the same as the Type of the non-host Cas.

[0492] A sixtieth paragraph relates to the method of any preceding paragraph, wherein the virus of (b) comprises a nucleotide sequence encoding a tracrRNA sequence, optionally wherein the tracrRNA sequence and first HM-crRNA are comprised by a single guide RNA (gRNA)).

[0493] A sixty-first paragraph relates to the method of any preceding paragraph, wherein the or each HM-crRNA is comprised by a respective single guide RNA (gRNA).

[0494] A sixty-second paragraph relates to the method of of any preceding paragraph, wherein the first HM-array is operable to cause Cas cleavage in the first target sequence, activation of the first target sequence (or gene comprising the first target sequence), knock-down of the first target sequence (or gene comprising the first target sequence) or mutation of the first target sequence.

[0495] A sixty-third paragraph relates to a virus, host cell or virus population obtainable by the method of any preceding paragraph, optionally wherein the population is identical to PV1 or PV2 or the virus is obtainable from such a population.

[0496] A sixty-fourth paragraph relates to t host cell (eg, bacterial cell) population obtainable by the method of any preceding paragraph, optionally wherein the population is identical to PH1, PH2, PH3 or PH4 or a cultured cell population recited in any preceding paragraph.

[0497] A sixty-fifth paragraph relates to the host cell population of the sixty-fourth paragraph wherein the population does not comprise nucleic acid of a virus of (b), or does not comprise said first HM-array or said second HM-array (eg, as determined by PCR).

[0498] A sixty-sixth paragraph relates to the virus, host cell or population of any one of the sixty-third to sixty-fifth paragraphs, for medical or dental or opthalmic use (eg, for treating or preventing an infection in an organism or limiting spread of the infection in an organism.

[0499] A sixty-seventh paragraph relates to a composition comprising a virus, host cell or population according to any

one of the sixty-third to sixty-sixth paragraphs for food, beverage, dairy or cosmetic use (eg, use in a cosmetic product, eg, make-up), or for hygiene use (eg, use in a hygiene product, eg, soap).

[0500] A sixty-eighth paragraph relates to use of a composition a virus, host cell or population according to any one of the sixty-third to sizty-seventh paragraphs, in medicine or for dental therapeutic or prophylactic use.

[0501] A sixty-ninth paragraph relates to use of a composition a virus, host cell or population according to any one of the sixty-third to sixty-eighth paragraphs, in cosmetic use (eg, use in a cosmetic product, eg, make-up), or for hygiene use (eg, use in a hygiene product, eg, a soap).

[0502] A seventieth paragraph relates to the use, virus, host cell or population of any one of the sicty-third to sixty-ninth paragraphs for modifying a microbial host cell (eg, for killing or reducing growth of the cell or a culture of microbe cells).

[0503] A seventy-first paragraph relates to the method, virus or virus population of any one of the first to sixty-third and sixty-sixth to seventieth paragraphs, wherein the virus or virus in said population express a holin and/or an endolysin for host cell lysis, optionally wherein the endolysin is a phage phi11, phage Twort, phage P68, phage phiWMY or phage K endolysin (eg, MV-L endolysin or P-27/HP endolysin).

[0504] A seventy-seventh paragraph relates to the method, virus or virus population of any one of the first to sixty-third and sixty-sixth to seventieth paragraphs, wherein the virus or virus in said population does no express a holin and/or an endolysin for host cell lysis.

[0505] A seventy-fourth paragraph relates to the method, virus or virus population of any one of the first to sixty-third and sixty-sixth to seventieth paragraphs, wherein the virus (eg, virus of (b)) or virus in each said population is in combination with an antimicrobial functional in the host cell of (a), eg, antibiotic agent, eg, a beta-lactam antibiotic (eg, an antibiotic recited in any one of the thirteenth ro twenty-seventh paragraphs).

CONTROL OF CORROSION, BIOFILMS & BIOFOULING

[0506] The invention relates *inter alia* to methods of controlling microbiologically influenced corrosion (MIC) or biofouling of a substrate or fluid in an industrial or domestic system. The disclosure also relates to treated fluids and vectors for use in the methods.

[0507] Corrosion is the result of a series of chemical, physical and (micro) biological processes leading to the deterioration of materials such as metal (eg, steel or iron), plastic and stone. It is a world-wide problem with great societal and economic consequences. Current corrosion control strategies based on chemically produced products are under increasing pressure of stringent environmental regulations. Furthermore, they are rather inefficient and may be hampered by microbial (eg, bacterial) resistance to the agents used. Therefore, there is an urgent need for environmentally friendly and sustainable corrosion control strategies. Corrosion is influenced by the complex processes of different microorganisms performing different electrochemical reactions and secreting proteins and metabolites that can have secondary effects.

[0508] The severity of microbial corrosion processes is evident from the fact that many of the industrially and domestically used metals and alloys such as stainless steels, nickel and aluminium-based alloys and materials such as concrete, asphalt and polymers are readily degraded by microorganisms. Protective coatings, inhibitors, oils and emulsions are also subject to microbial degradation.

[0509] Microbially influenced corrosion (MIC) is a costly problem that impacts hydrocarbon production and processing equipment, water distribution systems, ships, railcars, and other types of metallic and non-metallic industrial and domestic systems. In particular, MIC is known to cause considerable damage to hydrocarbon fuel infrastructure including production, transportation, and storage systems, oftentimes with catastrophic environmental contamination results. Around 40% of pipe corrosion in the oil industry is attributed to microbiological corrosion and leads to huge financial losses in production, transportation and storage of oil every year. Pipe biofilms can cause the reduction in fluid velocity in equipment due to the process of incrustation on walls. Furthermore, pipe leaks are generated as a result of the corrosion, with consequent impacts on the environment and productivity.

[0510] MIC takes place in environments such as soil, fresh water and sea water and is estimated to be responsible for more than 30 percent of all corrosion damage. MIC occurs due to the fixation of microbes such as bacteria, release of metabolites and usually formation of biofilms that induce or accelerate the corrosion process. Among the groups of bacteria involved in the corrosion process are included: sulphur- or sulphate-reducing bacteria (SRB), extracellular polymeric substance-producing bacteria (EPSB), acid-producing bacteria (APB), sulphur- or sulphide-oxidising bacteria (SOB); iron- or manganese-oxidising bacteria (IOB), ammonia prouducing bacteria (AmPB) and acetate producing bacteria (AcPB). Small subunit ribosomal RNA gene pyrosequencing surveys indicate that acetic-acid-producing bacteria (*Acetobacter* spp. and *Gluconacetobacter* spp.) are prevalent in environments exposed to fuel-grade ethanol and water.

[0511] Microbial growth under environmental conditions influences electrochemical reactions directly or indirectly. Microbe-substrate interactions lead to initial adhesion and biofilm formation. The attachment of microbes such as bacteria to substrate, release of metabolites and formation of biofilms influences the electrochemical conditions at substrate surfaces, inducing or accelerating the corrosion process, thereby mediating the process of MIC. The formation of a

bacterial biofilm on a metallic substrate comprises the following stages: I - formation of a film, through the adsorption of organic and inorganic molecules on the metal, which modifies the load distribution on the metallic surface and, also serves as a nutritional source for the bacteria, facilitating the adherence of free-floating microorganisms present in the liquid; II - adhesion and multiplication of aerobic bacteria forming microcolonies; III - production of extracellular polymeric substances (EPS) by some sessile bacteria; IV - colonisation by aerobic free-floating microbial cells, that will consume the oxygen by respiration, creating a local anaerobic environment in the biofilm as required by strict anaerobic bacteria and ; V - increase of biofilm thickness, which may favour the shedding of the outer layers. The EPS produced by the bacteria adhered to the biofilm capture essential ions for their growth; they are used as a means of attachment and protect bacteria against biocides interfering with the mechanisms of corrosion by favouring the creation of differential aeration areas, besides serving as a nutritional source in case of low nutrient availability. The process of corrosion by differential aeration occurs due to uneven distribution of the biofilm on the metal substrate with aerated regions (surrounding the biofilm) and non-aerated regions (below the biofilm). The biofilm formation on the metal surface decreases the oxygen content, reaching levels of almost total anaerobiosis. *Pseudomonas* is the main EPS producer genus.

[0512] An example of a MIC biocorrosion process mediated by corrosive bacteria is as follows: (A) Aerobic corrosive bacteria from fresh water, sea water, industrial/domestic systems or storage tanks reach out equipment and pipelines of industrial or domestic systems, that have a conditioning film on the surface. (B) EPS-producing bacteria attach to equipment/pipeline walls and produce EPS, which creates a favourable environment for adhesion by other microorganisms. (C) Adhesion of other groups of corrosive bacteria to pipeline walls takes place, which release their metabolites, developing into a microcolony through cell division, consuming oxygen available. Action of iron-oxidising bacteria results in a large accumulation of ferric precipitation leading to blockage in the equipment/pipeline; sulphuric acid released by sulphur-oxidising bacteria promotes the acidification of the environment. (D) The low oxygen concentration and organic acids released by acid-producing bacteria favour attachment and development of sulphate-reducing bacteria producing hydrogen sulphide (H_2S), thereby accelerating the corrosion process and reducing the local pH. (E) A corroded equipment/pipeline results, which is partially blocked by iron precipitates with micro-leaks and containing a bacterial biofilm. The H_2S poses a serious health risk to personnel operating the system affected. Furthermore, the production of thick biofilms and sludges lead to biofouling and hampering of the functioning of the system.

[0513] Similarly, bacterial populations may propogate in fluids, such as water stores or reservoirs (eg, in drinking water or in water of cooling systems), thereby mediating biofouling of the fluid. This may also be referred to as souring of the fluid. An example is waterway or drinking water reservoir souring.

[0514] The disclosure addresses such problems of MIC and biofouling by providing the following Aspects 1 *et seq*:-

[0515] In a first aspect, there is provided a method of controlling microbiologically influenced corrosion (MIC) or biofouling of a substrate in an industrial or domestic system, wherein a surface of the substrate is in contact with a population of first host cells of a first microbial species that mediates MIC or biofouling of the substrate, the method comprising

(i) contacting the population with a plurality of vectors that are capable of transforming or transducing the cells, each vector comprising a CRISPR array whereby CRISPR arrays are introduced into the host cells, wherein

(a) each CRISPR array comprises one or more nucleotide sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in a host cell; and

(b) each crRNA is capable of hybridising to a target sequence of a host cell to guide Cas (eg, a Cas nuclease, eg, a Cas9 or Cpf1) in the host cell to modify the target sequence (eg, to cut the target sequence); the target sequence being a gene sequence for mediating host cell viability; and

(ii) allowing expression of said cRNAs in the presence of Cas in host cells, thereby modifying target sequences in host cells, resulting in reduction of host cell viability and control of MIC or biofouling of said substrate.

[0516] In an example, the system comprises equipment (eg, for use in an industrial process) and the surface is a surface of said equipment.In an example, each array is an engineered array, eg, any engineered array disclosed herein. In an embodiment, the vector is an engineered CRISPR nucleic acid vector as described herein. In an example, the biofouling comprises microbial biofilm and/or sludge formation, proliferation or maintenance. In an example, the first host cells are sessile. In an example of Aspect 1 or 4 (below), "controlling" comprises preventing, reducing or eliminating said MIC or biofouling, or reducing spread of said MIC or biofouling in the system. Non-limiting examples of how bacteria mediate MIC or biofouling are described above. Cell growth or proliferation or maintenance is, for example, a characteristic of cell viability. Thus, in an example, the method reduces host cell proliferation and/or maintenance. In an example, the method kills host cells.

[0517] A second aspect relates to the method of the first Aspect, wherein said host cells are comprised by a microbial biofilm that is in contact with said substrate.

[0518] A third aspect relates to the method of any preceding Aspect, wherein said surface and host cells are in contact

with a fluid, such as an aqueous liquid (eg, sea water, fresh water, stored water or potable water).

[0519] Fresh water is naturally occurring water on the Earth's surface in ice sheets, ice caps, glaciers, icebergs, bogs, ponds, lakes, rivers and streams, and underground as groundwater in aquifers and underground streams. Fresh water is generally characterized by having low concentrations of dissolved salts and other total dissolved solids. The term specifically excludes sea water and brackish water, although it does include mineral-rich waters such as chalybeate springs. In an example said fresh water is any of these fresh water types. Potable water is water for human or animal (eg, livestock) consumption. In an example, the fluid is selected from industrial cooling water wherein the system is a cooling system; sewage water wherein the system is a sewage treatment or storage system; drinking water wherein the system is a drinking water processing, storage, transportation or delivery system; paper making water wherein the system is a paper manufacture or processing system; swimming pool water wherein the system is a swimming pool or swimming pool water teatment or storage system; fire extinguisher water wherein the system is a fire extinguishing system; or industrial process water in any pipe, tank, pit, pond or channel.

[0520] A fourth aspect relates to a method of controlling microbial biofouling of a fluid in an industrial or domestic system (eg, for controlling bacterial souring of a liquid in a reservoir or container), wherein the fluid comprises a population of first host cells of a first microbial species that mediates said biofouling, the method comprising

(i) contacting the population with a plurality of vectors that are capable of transforming or transducing the cells, each vector comprising a CRISPR array whereby CRISPR arrays are introduced into the host cells, wherein

(a) each CRISPR array comprises one or more sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in a host cell; and
(b) each crRNA is capable of hybridising to a target sequence of a host cell to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence (eg, to cut the target sequence); the target sequence being a gene sequence for mediating host cell viability; and

wherein the method comprises allowing expression of said cRNAs in the presence of Cas in host cells, thereby modifying target sequences in host cells, resulting in reduction of host cell viability and control of said biofouling.

[0521] In an example, the fluid is a liquid. In an example, the fluid is a gaseous fluid.

[0522] Systems: An example system for any Aspect is selected from the group consisting of a:-Petrochemical recovery, processing, storage or transportation system; hydrocarbon recovery, processing, storage or transportation system; crude oil recovery, processing, storage or transportation system; natural gas recovery, processing, storage or transportation system, (eg, an oil well, oil rig, oil drilling equipment, oil pumping system, oil pipeline, gas rig, gas extraction equipment, gas pumping equipment, gas pipeline, oil tanker, gas tanker, oil storage equipment or gas storage equipment); Water processing or storage equipment; water reservoir (eg, potable water reservoir); Air or water conditioning (eg, cooling or heating) equipment, eg, a coolant tube, condenser or heat exchanger; Medical or surgical equipment; Environmental (eg, soil, waterway or air) treatment equipment; Paper manufacturing or recycling equipment; Power plant, eg, a thermal or nuclear power plant; Fuel (eg, hydrocarbon fuel, eg, petroleum, diesel or LPG) storage equipment; Mining or metallurgical, mineral or fuel recovery system, eg, a mine or mining equipment; Engineering system; Shipping equipment; Cargo or goods storage equipment (eg, a freight container); Food or beverage manufacturing, processing or packaging equipment; Cleaning equipment (eg, laundry equipment, eg, a washing machine or dishwasher); Catering (eg, domestic or commercial catering) equipment; Farming equipment;Construction (eg, building, utilities infrastructure or road construction) equipment; Aviation equipment; Aerospace equipment; Transportation equipment (eg, a motor vehicle (eg, a car, lorry or van); a railcar; an aircraft (eg, an aeroplane) or a marine or waterway vehicle (eg, a boat or ship, submarine or hovercraft)); Packaging equipment, eg, consumer goods packaging equipment; or food or beverage packaging equipment; Electronics (eg, a computer or mobile phone or an electronics component thereof); or electronics manufacture or packaging equipment; Dentistry equipment; Industrial or domestic piping (eg, a sub-sea pipe) or storage vessel (eg, a water tank or a fuel tank (eg, gasoline tank, eg, a gasoline tank of a vehicle));Underground equipment; Building (eg, a dwelling or office or commercial premises or factory or power station); Roadway; Bridge; Agricultural equipment; Factory system; Crude oil or natural gas exploration equipment; Office system; and a Household system.

[0523] In an example, the system is used in an industry or business selected from the group consisting of agriculture, oil or petroleum industry, food or drink industry, clothing industry, packaging industry, electronics industry, computer industry, environmental industry, chemical industry, aerospace industry, automotive industry, biotechnology industry, medical industry, healthcare industry, dentistry industry, energy industry, consumer products industry, pharmaceutical industry, mining industry, cleaning industry, forestry industry, fishing industry, leisure industry, recycling industry, cosmetics industry, plastics industry, pulp or paper industry, textile industry, clothing industry, leather or suede or animal hide industry, tobacco industry and steel industry. In an example, the surface or fluid to be treated is a surface or fluid of equipment used in said selected industry. In an example, the system is used in the crude oil industry. In an example, the system is used in the natural gas industry. In an example, the system is used in the petroleum industry. In an example,

the system is a sea container, platform or rig (eg, oil or gas platform or rig for use at sea or at sea), ship or boat. In an embodiment, such a system is anchored at sea; eg, non-temporarily anchored at sea, eg, has been anchored at sea for 1, 2, 3, 4, 5, 6,7,8, 9,10, 11, 12, 13,14,15, 16, 17, 18, 19, 20, 21, 22 , 23, 24 or more months (eg, contiguous months). In an embodiment, such a system is in the waters of a country or state; eg, non-temporarily at sea in such waters, eg, has been in waters of said country for 1, 2, 3, 4, 5, 6,7,8, 9,10, 11, 12, 13,14,15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or more months (eg, contiguous months).

[0524] In an example, the substrate surface to be treated comprises stainless steel, carbon steel, copper, nickel, brass, aluminium, concrete, a plastic or wood. In an example, the substrate is a metal weld or join. In an example, the surface is a metallic (eg, steel or iron) or non-metallic (eg, plastic, concrete, asphalt, wood, rubber or stone) surface. In an example, the metal is an alloy (eg, stainless steel, brass or a nickel-, zinc-, copper-, nickel- or aluminium- alloy). In an example, the surface is a man-made polymer surface. In an example, the surface is a substrate coating. In an example, the substrate is in contact with soil, fresh water or sea water.

[0525] In an example, the fluid is potable water; a waterway; brackish water; or a liquid fuel, eg, gasoline or diesel (eg, for a car or motorised vehicle), LPG, kerosine, an alcohol (eg, ethanol, methanol or butanol), liquid hydrogen or liquid ammonia), in an example, the fuel is stored liquid fuel. In an example the fluid is an oil or non-aqueous liquid. In an example, the fluid is a liquid comprised by a waterway or body of water, eg, sea water, fresh water, potable water, a river, a stream, a pond, a lake, a reservoir, stored water (eg, in a water storage tank or cooling equipment), groundwater, well water, water in a rock formation, soil water or rainwater. In an example, the liquid is sea water. In an example, the substrate is in contact with a liquid mentioned in this paragraph. In an example, the fluid or liquid is selected from the group consisting of an oil, an aqueous solution, a hydraulic fracturing fluid, a fuel, carbon dioxide, a natural gas, an oil/water mixture, a fuel/water mixture, water containing salts, ocean or sea water, brackish water, sources of fresh water, lakes, rivers, stream, bogs, ponds, marshes, runoff from the thawing of snow or ice, springs, groundwater, aquifers, precipitation, any substance that is a liquid at ambient temperature (eg, at rtp) and is hydrophobic but soluble in organic solvents, hexanes, benzene, toluene, chloroform, diethyl ether, vegetable oils, petrochemical oils, crude oil, refined petrochemical products, volatile essential oils, fossil fuels, gasoline, mixtures of hydrocarbons, jet fuel, rocket fuel, biofuels. In an example the fluid is an oil/water mixture.

[0526] The terms "microbiologically influenced corrosion" or "MIC" as used herein, unless otherwise specified, refer to processes in which any element (substrate) of a system is structurally compromised due to the action of at least one member of a microbial population, eg, bacterial or archaeal population. The term "biofouling" as used herein, unless otherwise specified, refers to processes in which microorganisms (such as bacteria and/or archaea) accumulate on a substrate surface in contact with a fluid (eg, water or an aqueous liquid, or a hydrocarbon, or a petrochemical). Also included is the undesirable accumulation and proliferation of microorganisms (such as bacteria and/or archaea) in a fluid (eg, water or an aqueous liquid, or a hydrocarbon, or a petrochemical), ie, "souring" of the fluid. In an example, the bacteria are comprised by ship or boat ballast water and the bacteria are environmentally undesirable. The term "substrate" as used herein refers to any type of surface on which cells can attach and a biofilm can form and grow or on which biofouling (eg slime or sludge formation) can occur. The substrate may be an "industrial" substrate such as the surface of equipment in an petrochemical, fuel, crude oil or gas piping system, or a "non-industrial" (eg, domestic, eg, household or office) substrate such as a kitchen counter or a shower substrate or a garden substrate.

[0527] In an alternative of any of the Aspects, instead of a population of host bacterial cells, the population is a population of archaeal cells of a first species.

[0528] A fifth aspect relates to the method of the fourth Aspect, wherein said fluid is an aqueous liquid (eg, sea water, fresh water, stored water or potable water).

[0529] A sixth aspect relates to the method of any one of the third to fifth Aspects, wherein the method comprises mixing the fluid with the vectors, thereby contacting the host cells with vectors. For example, the vectors can be pre-mixed with a liquid (optionally with an antibiotic or biocide too) and the mixture then added to the fluid that is in contact with the surface (Aspect 1) or the fluid of the fourth Aspect.

[0530] A seventh aspect relates to the method of any one of the first to sixth Aspects, wherein each target sequence is a host cell virulence, resistance or essential gene sequence, eg, an exon or reguatory sequence thereof. Resistance can be antibiotic resistance. In an example, the host cells are contacted with said antibiotic and said vectors to reduce host cell viability.

[0531] An eighth aspect relates to the method of any one of the first to seventh Aspects, wherein the modification of target sequences results in host cell killing and/or a reduction in host cell growth or proliferation. Proliferation is, for example, cell expansion or cell distribution in contact with the surface. A ninth aspect relates to the method of any one of the first to eighth Aspects, wherein the vectors comprise identical CRISPR arrays.

[0532] A tenth aspect relates to the method of any one of the first to ninth Aspects, wherein the host cells are bacterial or archaeal cells. In an alternative, instead the first cells are algal cells.

[0533] An eleventh aspect relates to the method of any one of the firsth to tenth Aspects, wherein the first host cells are sulphate reducing bacteria (SRB) cells (eg, *Desulfovibrio* or *Desulfotomaculum* cells). In an example, the cells are

selected from the group consisting of *Desulfotomaculum nigrificans, Desulfacinum infernum, Thermodesulfobacterium mobile, Thermodesulforhabdus norvegicus, Archaeoglobus fulgidus, Desulfomicrobium apsheronum, Desulfovibrio gabonensis, Desulfovibrio longus, Desulfovibrio vietnamensis, Desulfobacterium cetonicum, Desulphomaculum halophilum, Desulfobacter vibrioformis and Desulfotomaculum thermocisternum* cells. In an example, the population comprises a mixture of two or more of these cell species.

[0534] A twelfth aspect relates to the method of the eleventh Aspect, wherein the surface or fluid is comprised by a crude oil, gas or petrochemicals recovery, processing, storage or transportation equipment. Crude oil is one of the most important energetic resources in the world. It is used as raw material in numerous industries, including the refinery-petrochemical industry, where crude oil is refined through various technological processes into consumer products such as gasoline, oils, paraffin oils, lubricants, asphalt, domestic fuel oil, vaseline, and polymers. Oil-derived products are also commonly used in many other chemical processes. In an alternative, the fluid is a said consumer product or the surface is in contact with such a consumer product.

[0535] A thirteenth aspect relates to the method of the eleventh or twelfth Aspect, wherein the surface is in contact with sea water, a fracking liquid or liquid in a well; or wherein the fluid is sea water, a fracking liquid or liquid in a well.

[0536] A fourteenth aspect relates to the method of any one of the first to thirteenth Aspects , wherein step (i) of the method comprises providing a population of microbial cells of a second species (second host cells), the second cells comprising said vectors, wherein the vectors are capable of transfer from the second host cells to the first host cells; and combining the second host cells with the first host cells, whereby vectors are introduced into the first host cells. In an example, the second cell(s) are environmentally-, industrially-, or domestically-acceptable in an environment (eg, in a water or soil environment) and the first host cell(s) are not acceptable in the environment.

[0537] A fifteenth aspect relates to the method of the fourteenth aspect, wherein the first host cells are comprised by a mixture of microbial cells (eg, comprised by a microbial biofilm) before contact with said vectors, wherein the mixture comprises cells of said second species.

[0538] A sixteenth aspect relates to the method of the fourteenth or fifteenth Aspect, wherein said second species is a species of *Bacillus* or nitrate-reducing bacteria or nitrate reducing sulfide oxidizing bacteria (NRB).

[0539] A seventeenth aspect relates to the method of the sixteenth Aspect, wherein the NRB is selected from the group consisting of *Campylobacter* sp., *Nitrobacter* sp., *Nitrosomonas* sp., *Thiomicrospira* sp., *Sulfurospirillum* sp., *Thauera* sp., *Paracoccus* sp., *Pseudomonas* sp., *Rhodobacter* sp. and *Desulfovibrio* sp; or comprises at least 2 of said species.

[0540] An eighteenth aspect relates to the method of the seventeenth Aspect wherein NRB is selected from the group consisting of *Nitrobacter vulgaris, Nitrosomonas europea, Pseudomonas stutzeri, Pseudomonas aeruginosa, Paracoccus denitrificans, Sulfurospirillum deleyianum, and Rhodobacter sphaeroides.*

[0541] A nineteenth aspect relates to *the* method of any one of the first to eighteenth Aspects, wherein the method comprises contacting the host cells of said first species with a biocide simultaneously or sequentially with said vectors. In an example, the vectors and biocide are provided pre-mixed in a composition that is contacted with the host cells.

[0542] A twentieth aspect relates to the method of the nineteenth Aspect, wherein the biocide is selected from the group consisting of tetrakis hydroxymethyl phosphonium sulfate (THPS), glutaraldehyde, chlorine monoxide, chlorine dioxide, calcium hypochlorite, potassium hypochlorite, sodium hypochlorite, dibromonitriloproprionamide (DBNPA), methylene bis(thiocyanate) (MBT), 2-(thiocyanomethylthio) benzothiazole (TCMTB), bronopol, 2-bromo-2-nitro-1,3-propanediol (BNPD), tributyl tetradecyl phosphonium chloride (TTPC), taurinamide and derivatives thereof, phenols, quaternary ammonium salts, chlorine-containing agents, quinaldinium salts, lactones, organic dyes, thiosemicarbazones, quinones, carbamates, urea, salicylamide, carbanilide, guanide, amidines, imidazolines, acetic acid, benzoic acid, sorbic acid, propionic acid, boric acid, dehydroacetic acid, sulfurous acid, vanillic acid, p-hydroxybenzoate esters, isopropanol, propylene glycol, benzyl alcohol, chlorobutanol, phenylethyl alcohol, formaldehyde, iodine and solutions thereof, povidone-iodine, hexamethylenetetramine, noxythiolin, 1-(3-chloroallyl)-3,5,7-triazo-l-azoniaadamantane chloride, taurolidine, taurultam, N-(5-nitro-2-furfurylidene)-l-amino-hydantoin, 5-nitro-2-furaldehyde semicarbazone, 3,4,4'-trichlorocarbanilide, 3,4',5-tribromosalicylanilide, 3-trifluoromethyl-4,4'- dichlorocarbanilide, 8-hydroxyquinoline, l-cyclopropyl-6-fluoro-l,4-dihydro-4-oxo-7-(l- piperazinyl)-3-quinolinecarboxylic acid, 1 ,4-dihydro- 1 -ethyl-6-fluoro-4-oxo-7-(l - piperazinyl)-3-quinolinecarboxylic acid, hydrogen peroxide, peracetic acid, sodium oxychlorosene, parachlorometaxylenol, 2,4,4'-trichloro-2'-hydroxydiphenol, thymol, chlorhexidine, benzalkonium chloride, cetylpyridinium chloride, silver sulfadiazine, silver nitrate, bromine, ozone, isothiazolones, polyoxyethylene (dimethylimino) ethylene (dimethylimino) ethylene dichloride, 2-(tert-butylamino)-4-chloro-6-ethylamino-5'- triazine (terbutylazine), and combinations thereof. In an example the biocide is tetrakis hydroxymethyl phosphonium sulfate (THPS). In an example, the biocide is a quaternary ammonium compound.

[0543] A twenthy-first aspect relates to the method of any one of the first to twentieth Aspects, wherein the system is used in an industry operation selected from the group consisting of mining; shipping; crude oil, gas or petrochemicals recovery or processing; hydraulic fracturing; air or water heating or cooling; potable water production, storage or delivery; transportation of hydrocarbons; and wastewater treatment.

[0544] A twenty-second aspect relates to the method of the twenthy-first Aspect, wherein the surface is a surface of equipment used in said selected industry; or wherein the fluid is a fluid comprised by equipment used in said selected industry.

[0545] A twenty-second aspect relates to the method of any one of the first to twenty-second Aspects, wherein the surface is a surface of kitchen, bathing or gardening equipment; or wherein the fluid is comprised by kitchen, bathing or gardening equipment. For example, the equipment is used in a domestic setting.

[0546] A twenty-fourth aspect relates to the method of any one of the first to twenty-third Aspects when dependent from the third Aspect, wherein the fluid is a potable liquid contained in a container (eg, water tank or bottle) and the surface is a surface of the container in contact with the liquid.

[0547] A twenty-fifth aspect relates to the method of any one of the first to twenty-fourth Aspects, wherein each vector comprises a mobile genetic element (MGE), wherein the MGE comprises an origin of transfer (*ori*T) and a said CRISPR array; wherein the MGE is capable of transfer between a host cell of said first species and a further microbial host cell in said industrial or domestic system. For example, the further cell(s) are environmentally-, industrially-, or domestically-acceptable in an environment (eg, in a water or soil environment) and the first host cell(s) are not acceptable in the environment.

[0548] A twenty-sixth aspect relates to the method of the twenty-fifth Aspect, wherein *ori*T is functional in the first and further host cells.

[0549] A twenty-seventh aspect relates to the method of the twenty-fifth or twenty-sixth Aspect, wherein said first and further host cells are comprised by a biofilm of fluid in contact with said surface; or wherein said cells are comprised by said fluid.

[0550] A twenty-eighth aspect relates to the method of the twenty-fifth, twenthy-sixth or twenty-seventh Aspect, wherein said further cell is a cell of a species as recited in any one of the sixteenth ro eighteenth Aspects. In an example, the MGE is capable of transfer from the further cell to the first host cell and/or *vice versa.*

[0551] A twenty-ninth aspect relates to the method of any one of the twenty-fifth to twenty-seventh Aspects, wherein the further cell is a cell of said first species.

[0552] For example, in this embodiment the MGE is capable of transfer amongst first cells in a population in said system. When the M*GE leaves* a copy of itself in the transfer process to the other cell, this then provides means for propagating and spreading the MGE and thus CRISPR arrays through cell populations in the system, thereby spreading the target sequence modifying effect of the arrays. This can be effective, for example, to create spread of arrays in a biofilm in contact with the surface or in the fluid, and is useful as penetration of biofilms with conventional biocides can be sub-optimal.

[0553] A thirtieth aspect relates to the method of any one of the twenty-fifth to twenty-ninth Aspects, wherein each MGE is or comprises an integrative and conjugative element (ICE); or wherein each vector is a phage that is capable of infecting host cells of said first species and each MGE is a phage nucleic acid that is capable of said transfer between the cells.

[0554] A thirty-first aspect relates to the method of the thirtieth Aspect, wherein each ICE is a transposon, eg, a conjugative transposon.

[0555] A thirty-second aspect relates to the method of any one of the first to thirty-first Aspects, wherein each vector is a plasmid, optionally comprising an MGE according to any one of the twenty-fifth to thirty-first Aspects.

[0556] A thirty-third aspect relates to the method of any one of the twenty-fifth to thirty-second Aspects, wherein the first and/or further cell comprises nucleotide sequences encoding proteins operable to transfer the MGE to the other cell, wherein the sequences are not comprised by the MGE.

[0557] A thirty-fourth aspect relates to the method of the thirty-third Aspect, wherein the sequences are not comprised by the vector.

[0558] A thirty-fifth aspect relates to the method of the thirty-third Aspect, wherein the sequences are comprised by a conjugative transposon of the first cell and/or further cell.

[0559] A thirty-sixth aspect relates to the method of the thirty-fifth Aspect, wherein the transposon is operable in *trans* to transfer the MGE between the first and further cells.

[0560] A thirty-seventh aspect relates to the method of any one of the twenty-fifth to thirty-sixth Aspects, wherein the *ori*T of the MGE is the same as an ori*T* comprised by an ICE of the first cell and/or further cells, wherein the ICE is operable in *trans* to transfer the MGE between the first and further cells.

[0561] A thirty-eighth aspect relates to the method of any one of the twenty-fifth to thirty-seventh Aspects, wherein the vector *ori*T is an *ori*T of a SRB or NRB transposon.

[0562] A thirty-ninth aspect relates to the method of any one of the twenty-fifth to thirty-eighth Aspects, wherein each MGE comprises first and second terminal repeat sequences and a said CRISPR array between the repeat sequences.

[0563] A fortieth aspect relates to the method of any one of the twenty-fifth to thirty-ninth Aspects, wherein the MGE leaves behind a CRISPR array copy (1) in the genome of a first host cell when it has transferred to a said further host cell; or (2) in a said further host cell when it has transferred to a first host cell. For example, the copy is comprised by a

transposon or prophage left in the genome of the cell from which transfer takes place.

[0564] A forty-first aspect relates to tmethod of any one of the twenty-fifth to fortieth Aspects, wherein the first and further cells are bacterial cells of different species (eg, SRB and NRB; or SRB and *Bacillus* cells respectively).

[0565] A forty-second aspect relates to the method of any one of the twenty-fifth to forty-first Aspects when dependent from Aspect 30 in combination with a transposase for mobilisation of the MGE.

[0566] A forty-third aspect relates to the method of any one of the first to forty-second Aspects, wherein the vector or MGE comprises a toxin-antioxin module that is operable in a host cell of said first species; optionally wherein the toxin-antitoxin module comprises an anti-toxin gene that is not operable or has reduced operation in cells of another species. These embodiments are useful to create a selective pressure that favours retention of the vector/MGE (and thus CRISPR arrays) in the first host cells comprising the target sequences.

[0567] A forty-fourth aspect relates to the method of any one of the first to forty-third Aspects, wherein the vector or MGE comprises a toxin-antioxin module that is operable in a said second or further cell; optionally wherein the toxin-antitoxin module comprises an anti-toxin gene that is not operable or has reduced operation in cells other than the second or further cell. This is useful to maintain a population of CRISPR arrays in the second or further cells (eg, when such cells are present in a biofilm also comprising the first cells), but wherein the toxin-antitoxin module provides additional killing (over and above the action of the target sequence modification) in first host cells. In an example, the vector or MGE comprises a toxin-antioxin module that is operable in a first host cell and in said second or further cell.

[0568] A forty-fifth aspect relates to the method of any one of the forty-third or forty-fourth Aspects, wherein the toxin-antitoxin module is not operable or has reduced operation in cells other than the first and second or further cells. Thus, there can be a selective pressure in both the first and second (or further) cells to maintain the CRISPR arrays. Usefully, this then provides a reservoir for horizontal transfer of the arrays in MGEs between cells in a mixed population (eg, a biofilm contacting the surface or a population comprised by the fluid).

[0569] A forty-sixth aspect relates to the method of any one of the twenth-fifth to forty-fifth Aspects wherein the first and second cells (or first and further cells) are of the same phylum (eg, both bacterial cells) and the vector is replicable or operable (A) in the first cell and/or second (or further) cell but not in another cell of the same phylum; (B) in the first cell and/or second (or further) cell but not in another cell of the same order; (C) in the first cell and/or second (or further) cell but not in another cell of the same class; (D) in the first cell and/or second (or further) cell but not in another cell of the same order; (E) in the first cell and/or second (or further) cell but not in another cell of the same family; (F) in the first cell and/or second (or further) cell but not in another cell of the same genus; or (G) in the first cell and/or second (or further) cell but not in another cell of the same species.

[0570] A forty-seventh aspect relates to the method of the twenty-fifth Aspect or any one of the twenth-sixth to forty-sixth Aspects when dependent from the twenth-fifth Aspect, wherein each MGE is a conjugative transposon, *ori*T is functional in the first and further (or second) host cells, the MGE comprises first and second terminal repeat sequences and a said CRISPR array between the repeat sequences, and wherein the first and further (or second) cells are bacterial cells, wherein the target site is comprised by the first cells but not the further (or second) cells, and wherein said modifying inactivates or down-regulates a gene or regulatory sequence comprising said target in the first cells, resulting in reduction of first host cell viability and control of said MIC or biofouling.

[0571] A forty-eighth aspect relates to the method of any one of the first to forty-seventh Aspects, wherein each CRISPR array comprises a sequence R1-S1-R1' for expression and production of the respective crRNA in a first host cell,

(i) wherein R1 is a first CRISPR repeat, R1' is a second CRISPR repeat, and R1 or R1' is optional; and
(ii) S1 is a first CRISPR spacer that comprises or consists of a nucleotide sequence that is 95% or more identical to a target sequence of a said first host cell.

[0572] A forthy-ninth aspect relates to the method of the forty-eighth Aspect, wherein R1 and R1' are at least 95, 96, 97, 98 or 99% identical respectively to the first and second repeat sequences of a CRISPR array of the first host cell species. In an embodiment, both R1 and R1' are present.

[0573] A fiftieth aspect relates to the method of the forty-eighth or forty-ninth Aspect, wherein R1 and R1' are functional with a CRISPR/Cas system of said host cells of said first species for modification of target sequences.

[0574] fifty-first aspect relates to the method of any one of the forty-eight to fiftieth Aspects, wherein the first host cells are sulphate reducing bacteria (SRB) cells and R1 and R1' are least 95, 96, 97, 98 or 99% identical respectively to a repeat sequence (eg, the first repeat) of a CRISPR array of the first host cell species.

[0575] A fifty-second aspect relates to tmethod of the fifty-first Aspect, wherein R1 and R1' are least 95, 96, 97, 98 or 99% identical respectively to a repeat sequence selected from the group consisting of SEQ ID NOs: 50-74. See Table 1. In an embodiment, both R1 and R1' are present.

[0576] A fifty-third aspect relates to the method of the fifty-first Aspect, wherein R1 and R1' are least 95, 96, 97, 98 or 99% identical respectively to a repeat sequence selected from the group consisting of SEQ ID NOs: 51, 54 and 69. SEQ ID NOs: 51, 54 and 69 are found in more than one SRB species. This is particularly useful for targeting more than one

SRB type with the CRISPR array of the invention, eg, when the SRB types co-exist in the industrial or domestic system to be treated, for example co-existing in a population or biofilm that is in contact with the substrate or in the fluid to be treated. In an embodiment, both R1 and R1' are present.

[0577] A fifty-fourth aspect relates to the method of any one of the forty-eighth to fifty-third Aspects, wherein the sequences of R1 and R1' are identical.

[0578] A fifty-fifth aspect relates to the method of any one of the first to fifty-foruth Aspects , wherein each array introduced into a first host cell is introduced in combination with one or more Cas nuclease(s) (eg, a Cas9 and/or Cfp1) that function with the respective crRNA in a host cell to modify a target sequence thereof.

[0579] In an example, Cas herein in any configuration is deactivated for nuclease activity and optionally comprises a target sequence activator or depressor. A Cas 9 herein is, for example *S pyogenes* or *S aureus* Cas9.

[0580] A fifty-sixth aspect relates to the method of any one of the first to fifty-fifth Aspects, wherein each array introduced into a first host cell is introduced in combination with nucleic acid sequence(s) encoding one or more Cas nuclease(s) (eg, a Cas9 and/or Cfp1) that function with the respective crRNA in a host cell to modify the target sequence.

[0581] A fifty-seventh aspect relates to the method of any one of the forty-eighth to fifty-sixth Aspects, wherein R1 and R1' are functional with a Type II Cas9 nuclease to modify a target sequence in a said first host cell, optionally wherein the method is further according to the fifty-first to fifty-sixth Aspect wherein the Cas is said Cas9.

[0582] A fifth-eighth aspect relates to the method of any one of the first to fifty-seventh Aspects, wherein all or some of said vectors or MGEs do not comprise a Cas nuclease-encoding sequence operable with the respective array.

[0583] A fifty-ninth aspect relates to the method of the fifty-eighth Aspect, wherein each said respective array is operable with a Cas endonuclease found in cells of the first species.

[0584] A sixtieth aspect relates to the method of the twenty-fifth Aspect, or any one of the twenty-sixth to fifty-ninth Aspects when dependent from the twenty-fifth Aspect, wherein each MGE is devoid of a sequence encoding a Cas endonuclease that is operable with repeat sequences of the array, and wherein the respective vector comprises such a sequence (eg, encoding a Cas9 of Cfp1) outside the MGE.

[0585] A sixty-first aspect relates to a method of controlling microbiologically influenced corrosion (MIC) or biofouling of a substrate comprised by a crude oil, gas or petrochemicals recovery, processing, storage or transportation equipment (eg, a crude oil tanker, oil rig or oil drilling equipment), wherein a surface of the substrate is in contact with a population of first host cells, wherein the first host cells are sulphur- or sulphate-reducing bacteria (SRB), extracellular polymeric substance-producing bacteria (EPSB), acid-producing bacteria (APB), sulphur- or sulphide-oxidizing bacteria (SOB), iron-oxidising bacteria (IOB), manganese-oxidising bacteria (MOB), ammonia producing bacteria (AmPB) or acetate producing bacteria (AcPB) of a first species that mediates MIC or biofouling of the substrate, wherein the surface and cell population are in contact with a liquid selected from sea water, fresh water, a fracking liquid or liquid in a well (eg, oil or natural gas well), the method comprising

(i) contacting the cell population with vectors by mixing the liquid with a plurality of vectors that are capable of transforming or transducing first host cells, each vector comprising a CRISPR array whereby CRISPR arrays are introduced into the host cells, wherein

(a) each CRISPR array comprises one or more sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in a host cell;
(b) each crRNA is capable of hybridising to a target sequence of a host cell to guide Cas (eg, a Cas nuclease, eg, a Cas9 or Cfp1) in the host cell to modify the target sequence (eg, to cut the target sequence); the target sequence being a gene sequence for mediating host cell viability;
(c) wherein each sequence of (a) comprises a sequence R1-S1-R1' for expression and production of the respective crRNA in a first host cell, wherein R1 is a first CRISPR repeat, R1' is a second CRISPR repeat, and R1 or R1' is optional; and S1 is a first CRISPR spacer that comprises or consists of a nucleotide sequence that is 70, 75, 80, 85, 90 or 95% or more identical to a target sequence of a said first host cell and

(ii) allowing expression of said cRNAs in the presence of Cas in host cells, thereby modifying target sequences in host cells, resulting in reduction of host cell viability and control of MIC or biofouling of said substrate. In an embodiment, both R1 and R1' are present.

[0586] A sixty-second aspect relates to the method of the sixty-first Aspect, wherein the method is according to the first Aspect or any preceding Aspect when dependent from the first Aspect.

[0587] A sixty-third aspect relates to the method of the sixty-first to sixty-second Aspect, wherein each vector is a phage capable of infecting a first host cell or is a vector comprising a MGE (eg, a transposon) that comprises a said CRISPR array, wherein the MGE is capable of transfer into a first host cell.

[0588] A sixty-fourth aspect relates to the method of the sixty-first, sixty-second or sixty-third Aspect, wherein the first

cells are sulphate reducing bacteria (SRB) cells, eg, *Desulfovibrio* or *Desulfotomaculum* cells.

[0589] A sixty-fifth aspect relates to the method of the sixty-fourth Aspect, wherein R1 and R1' are at least 95, 96, 97, 98 or 99% identical respectively to a repeat sequence (eg, the first repeat) of a CRISPR array of the first host cell species and the vector arrays are operable with a Cas endonuclease found in cells of the first species. In an example, R1 and R1' are identical sequences.

[0590] A sixty-sixth aspect relates to the method of the sixty-fifth Aspect, wherein R1 and R1' are at least 95, 96, 97, 98 or 99% identical respectively to a repeat sequence selected from the group consisting of SEQ ID NOs: 50-74. In an example, R1 and R1' are identical sequences.

[0591] A sixty-seventh aspect relates to the method of the sixty-sixth Aspect, wherein R1 and R1' are at least 95, 96, 97, 98 or 99% identical respectively to a repeat sequence selected from the group consisting of SEQ ID NOs: 51, 54 and 69. See Table 1. This is particularly useful for targeting more than one SRB type with the CRISPR array of the invention, eg, when the SRB types co-exist in the industrial or domestic system to be treated, for example co-existing in a population or biofilm that is in contact with the substrate or in the fluid to be treated. In an example, R1 and R1' are identical sequences.

[0592] A sixty-eighth aspect relates to the method of any one of the first to sixty-seventh Aspects, wherein said plurality of vectors comprise additional vectors, wherein each additional vector comprises one or more CRISPR arrays for targeting additional host cells comprised by said population, wherein the additional host cell species is different from the first host cell species, wherein in step (i) said additional cells of the population are contacted with a plurality of said additional vectors that are capable of transforming or transducing the additional cells, each vector comprising a CRISPR array whereby CRISPR arrays are introduced into the additional host cells, wherein

(a) each CRISPR array comprises one or more sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in a host cell; and
(b) each crRNA is capable of hybridising to a target sequence of a said additional host cell to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence (eg, to cut the target sequence); the target sequence being a gene sequence for mediating host cell viability; and step (ii) comprises allowing expression of said cRNAs in the presence of Cas in said additional host cells, thereby modifying target sequences in additional host cells.

[0593] A sixty-ninth aspect relates to the method of the sixty-eighth Aspect, wherein the additional host cells mediate MIC or biofouling of said substrate or fluid, wherein step (ii) results in reduction of additional host cell viability and control of MIC or biofouling of said substrate or fluid.

[0594] A seventieth aspect relates to a method of controlling bacterial biofouling in ballast water of a ship or boat, wherein the water comprises a population of first host cells of a first microbial species that mediates said biofouling, the method comprising

(i) contacting the population with a plurality of vectors that are capable of transforming or transducing the cells, each vector comprising a CRISPR array whereby CRISPR arrays are introduced into the host cells, wherein

(a) each CRISPR array comprises one or more sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in a host cell; and
(b) each crRNA is capable of hybridising to a target sequence of a host cell to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence (eg, to cut the target sequence); the target sequence being a gene sequence for mediating host cell viability; and

(ii) allowing expression of said cRNAs in the presence of Cas in host cells, thereby modifying target sequences in host cells, resulting in reduction of host cell viability and control of said biofouling.

[0595] A seventy-first aspect relates to the method of the seventieth Aspect, wherein the first host cells are *Vibrio cholerae, E coli* or *Enterococci* sp cells.

[0596] A seventy-second aspect relates to the method of the seventieth or seventy-first Aspect, wherein step (i) comprises mixing the ballast water with the vectors, eg, in the hull of a ship or boat.

[0597] A seventy-third aspect relates to the method of any one of the seventieth to seventy-second Aspects, wherein the ship or boat is a marine vehicle and the water is sea water.

[0598] A seventy-first aspect relates to the method of any one of the seventieth and seventy-second Aspects, wherein instead of a ship or boat, the ballast water is comprised by a container or a drilling platform at sea, eg, an oil platform or oil rig. In an example, the ship, boat, container, platform or rig is anchored at sea (ie, not temporarily in its location).

[0599] A seventy-fifth aspect relates to a method of discharging ballast water from a ship or boat, wherein the discharged ballast water comprises water treated by the method of any one of the seventieth to seventy-fourth Aspects.

[0600] A seventy-sixth aspect relates to the method of the seventy-fifth Aspect, wherein the water is discharged into a body of water, eg, a sea, ocean or waterway (eg, a river, canal, lake or reservoir) or into a container.

[0601] A seventy-seventh aspect relates to ballast sea water comprising CRISPR arrays, wherein the ballast water is obtained or obtainable by the method of any one of the seventieth to seventy-sixth Aspects.

[0602] A seventy-eighth aspect relates to a ship, boat, container or rig comprising the ballast sea water of the seventy-seventh Aspect.

[0603] A seventy-ninth aspect relates to a vector for use in the method of any one of the sixty-first to sixty-ninth Aspects, wherein the first cells are sulphate reducing bacteria (SRB) cells, eg, *Desulfovibrio* or *Desulfotomoculum* cells, each vector comprising one or more CRISPR arrays for targeting the SRB, wherein each array is as defined in (a)-(c) of the sixty-first Aspect.

[0604] An eightieth aspect relates to the vector of the seventy-ninth Aspect, wherein R1 and R1' are according to any one of the sixty-fifth to sixty-seventh Aspects.

[0605] An eighty-first aspect relates to a vector for use in the method of any one of *the seventieth to seventy-sixth Aspects, wherein the first* cells are *Cholera* (eg, *vibrio,* eg, O1 or O139), *E coli or Enterococci sp cells,* the vector comprising one or more CRISPR arrays for targeting the cells, wherein each array is as defined in (a) and (b) of the seventieth Aspect.

[0606] An eighty-second aspect relates to the vector of any one of the seventy-ninth to eight-first Aspects, wherein the vector is a bacteriophage capable of infecting a said cell.

[0607] An eighty-third aspect relates to the vector of any one of the seventy-ninth to eighty-first Aspects, wherein the vector is a transposon or MGE capable of transfer into a said cell.

[0608] An eighty-fourth aspect relates to a plurality vectors, wherein each vector is according to Aspect 82 or 83, optionally in combination with a biocide or antibiotic that is capable of reducing viability of said cells.

[0609] Bacteria that Mediate MIC or Biofouling: In an example, the first host cells are selected from the group consisting of sulphur- or sulphate-reducing bacteria (SRB), extracellular polymeric substance-producing bacteria (EPSB, eg, Pseudomonas), acid-producing bacteria (APB), sulphur- or sulphide-oxidising bacteria (SOB); iron- or manganese-oxidising bacteria (IOB), ammonia prouducing bacteria (AmPB) and acetate producing bacteria (AcPB). For example, the first host cells are AcPB (eg, Acetobacter spp. and/or Gluconacetobacter spp) and the surface is in contact with a hydrocarbon fuel (eg, fuel-grade ethanol) and/or water.

[0610] The following are examples of relevant bacteria for the present invention (in an example, the first host cells are cells of any of the following species). *Acidithiobacillus* bacteria produce sulphuric acid. *Acidithiobacillus thiooxidans,* a subgenus of *Acidithiobacillus* bacteria, frequently damages sewer pipes. *Ferrobacillus ferrooxidans* directly oxidises iron to iron oxides and iron hydroxides. Other bacteria produce various acids, both organic and mineral, or ammonia. In the presence of oxygen, aerobic bacteria like *Thiobacillus thiooxidans, Thiobacillus thioparus,* and *Thiobacillus concretivorus,* all three widely present in the environment, are the common corrosion-causing factors resulting in biogenic sulphide corrosion. Without presence of oxygen, anaerobic bacteria, especially *Desulphovibrio* and *Desulphotomaculum,* are common. *Desulphovibrio salixigens* requires at least 2.5% concentration of sodium chloride, but *D. vulgaris* and *D. desulphuricans* can grow in both fresh and salt water. D. *africanus* is another common corrosion-causing microorganism. The *Desulphotomaculum* genus comprises sulphate-reducing spore-forming bacteria. *Desulphotomaculum orientis* and *nigrificans* are involved in corrosion processes. Sulphate-reducers require a reducing environment, and an electrode potential of at least -100 mV is required for them to thrive. However, even a small amount of produced hydrogen sulphide can achieve this shift, so the growth, once started, tends to accelerate.

[0611] In an Example the first host cells are *Serratia marcescens, Gallionella sp., Pseudomonas sp., Bacillus sp.(eg, B. subtilis, B. cereus, B. pumilus or B. megaterium), Thiobacillus sp., Sulfolobus sp., Klebsiella oxytoca, Pseudomonas aeruginosa, P. stutzeri, Micrococcus, Enterococcus, Staphylococcus (eg, S. aureus), E. faecalis* or M. *luteus* cells. In an example, the first host cells comprise a mixture of two or more of said species. These species have been isolated from diesel and naphtha-transporting pipelines located in the northwest and southwest regions in India; the association with localized corrosion of the pipeline steel in the presence of these consortia was corroborated. A joint project of different european aircraft manufacturers confirmed the involvement of isolates from genera *Micrococcus, Enterococcus, Staphylococcus* and *Bacillus* in strong corrosion damage in aluminium alloy, commonly used in aircraft construction. These bacteria may create a microacidic environment (acid producing bacteria), which favours the development of other bacteria, or produce EPS, favouring the formation of biofilm (EPS-producing bacteria). Thus, in an embodiment of the invention, the surface (eg, steel surface) of the system to be treated is in contact with diesel or naptha, or the fluid to be treated is diesel or naptha (and optionally the first host cells are of one or more species defined in this paragraph). In an embodiment of the invention, the surface (eg, aluminium-containing surface, eg, an aircraft surface) of the system to be treated is in contact with one, two, three or all genera: *Micrococcus, Enterococcus, Staphylococcus* and *Bacillus* (first host cells). In an example of any embodiment in this paragraph, the surface is a surface of a steel or aluminium component of the system.

[0612] Acid-producing bacteria: Aerobic bacteria are able to produce short-chain organic acids such as acetic, formic,

lactic, propionic and butyric acids as products of their metabolism from the fermentative metabolism of organic materials . They are also initial colonizers due to aerobic metabolism. These microorganisms are present in a variety of environments, including gas stands and oils. Organic acids serve as substrates for the SRB, accelerating the corrosion process, besides reducing the pH of the surrounding medium. Furthermore, the large amount of organic acid produced acts in metal depolarisation, starting the local corrosive process.

[0613] Sulphur-oxidising bacteria: The sulphur-oxidising bacteria are aerobic and facultative anaerobic microorganisms which obtain the energy necessary for growth from the oxidation of inorganic sulphur compounds such as sulphide, sulphite, thiosulphate and, in some cases the sulphur. Oxidative metabolism results in the production of sulphuric acid which promotes environment acidification. This group encompasses many genera, the *Acidithiobacillus* genus being the most studied. The group also includes bacterial species from the genera *Sulfolobus, Thiomicrospira, Beggiatoa, Acidithiobacillus, and Thiothrix* as well as the *species Thiosphaera pantotropha* and *Paracoccus denitrificans.* In an example, the first host cells are cells of any one of these species.

[0614] Iron-oxidising bacteria: Iron oxidising bacteria are aerobic microorganisms, belonging to a large and diverse group, that get energy necessary for their metabolism from iron oxidation. Consequently, there is the formation of iron hydroxides that generally form insoluble precipitate on substrate surfaces, promoting regions with different oxygen levels. They are widely found in water from rivers, lakes and oil production. They have mostly a locomotor sheath and their presence can be detected by a large accumulation of ferric precipitated as corrosion product. This accumulation or inorganic fouling leads to problems to industrial equipment such as blockages in oil pipelines. Among the most common are: *Thiobacillus ferrooxidans* and the genera *Crenothrix, Gallionella, Leptothrix* and *Spherotillus.* In an example, the first host cells are cells of any one of these species.

[0615] Sulphur- or sulphate-reducing bacteria (SRB): The SRB form a morphological- and phylogenetically heterogenous group that includes bacteria and restricted anaerobic archaebacteria, although some species have significant tolerance to oxygen. They are mainly gram- negative bacteria, mesophilic and some thermophilic generally spore-forming. These microorganisms are capable of oxidising various organic compounds of low molecular weight, including mono- or dicarboxylic aliphatic acids, alcohols, hydrocarbons and aromatic compounds, using sulphate ions or other sulphur compounds (thiosulphate, sulphite, etc.) as electron acceptors. Acetate, lactate, pyruvate and ethanol are among the most commonly used substrates by SRB. The stimulation of SRB growth is due to existing anaerobic conditions in biofilms explained by the deposition of corrosion products combined with microorganisms and, during oil recovery, where there is injection of aqueous media such as sea water, rich in sulphate. Large amounts of biogenic hydrogen sulphide can be produced; most of the H_2S formed in pipelines and other oil, gas or petrochemicals recovery, processing, storage or transportation equipment originates from the metabolic activity of SRB. Another economic impact on the oil industry is the acidification of oil and gas by H_2S.

[0616] Considering the numerous economic losses related to metabolic activity of SRB , efforts have been directed to the use of environmentally-harmful and toxic metabolic inhibitors such as molybdate, nitrate and nitrite, and application of biocides, which help the control of metabolic activity of SRB and subsequent inhibition of biogenic H_2S production.

[0617] Several mechanisms contribute to contain the formation process of biogenic H_2S by using metabolic inhibitors: I- competition between SRB and heterotrophic bacteria that are reducers of nitrite or nitrate by ordinary electron donors, resulting in competitive SRB exclusion; II- increased redox potential due to the presence of intermediaries of nitrate reduction (nitrous oxide and nitric oxide), since the biological production of H_2S occurs only at low redox potential (below -100 mV); III- Change of energy metabolism of some SRB, reducing nitrate instead of sulphate; IV- sulphide oxidising bacteria and nitrate or nitrite reducing bacateria that use the nitrate or nitrite to re-oxidise H_2S, resulting in H_2S removal; V- inhibition of the dissimilatory sulphite reductase by nitrite to inhibit the final enzymatic step *via* sulphate reduction in SRB.

[0618] In certain embodiments of the present invention, the host cell population in contact with the substrate to be treated or comprised by the fluid to be treated is also contacted with one or more nitrate and/or one or more nitrite in the presence of the vectors of the invention. For example, in step (i) simultaneously or sequentially with the vectors, the nitrate/nitrite and vectors are combined with (eg, injected into) oil, gas, petrochemical, water or other fluid comprised by the industrial or domestic system. Similarly, additionally or alternatively, molybdates also may also be used in these systems as a control mechanism for SRB. Thus, in one embodiment, the host cell population in contact with the substrate to be treated or comprised by the fluid to be treated is also contacted with one or more molybdate in the presence of the vectors of the invention. For example,in step (i) simultaneously or sequentially with the vectors, the molybdate(s) and vectors are combined with (eg, injected into) oil, gas, petrochemical or other fluid comprised by the industrial or domestic system.

[0619] In other embodiments, the population is contacted with nitrate-reducing bacteria and/or nitrate reducing sulphide oxidising bacteria (NRSOB) (herein collectively, "NRB") in the presence of the vectors of the invention. For example, simultaneously or sequentially with the vectors, the NRB are combined with (eg, injected into) oil, gas, petrochemical, water or other fluid comprised by the industrial or domestic system. In an example, the NRB comprise vectors of the invention, wherein the vectors are capable of transfer from the NRB cells to the first host cells (SRB cells); and following combining the NRB and SRB cells, the vectors are introduced into the SRB cells. In an example, the SRB cells are

comprised by a mixture of microbial cells (eg, comprised by a microbial biofilm) before contact with said vectors, wherein the mixture comprises cells of the NRB species. Thus, in this case the invention involves contacting the SRB cells with NRB cells (containing vectors) where the NRB cell species are already coexisting with the SRB in the biofilm to be targeted, which thus increases compatibility and chance of uptake of the vector-containing NRB into the biofilm cell population. This is useful for increasing the chances of the vectors being taken into the biofilm, thereby increasing chances of efficacy to modify SRB cells and chances of propagation of the CRISPR arrays of the invention within the biofilm (especially when the arrays are comprised by mobile genetic elements, such as transposons or comprised by phage, as herein described).

[0620] SRB and NRB typically compete for the same non-polymer carbon source (such as acetates) present in certain oilfield and industrial water systems needed for growth of bacteria. By increasing the growth rate of the NRB in comparison to the SRB, the NRB may out compete the SRB in consumption of the available non-polymer carbon source, depriving the SRB of its ability to grow and create the undesirable sulphides and reduce corrosion rates. Further, by inhibiting the growth rate of the SRB, the NRB may predominate, again out competing the SRB for the available non-polymer carbon in the system, eg, oilfield or industrial water system. Thus, contacting the SRB cells in the population with NRB can help to reduce SRB cell viability by increasing the ratio of NRB to SRB in the population.

[0621] In an embodiment, the invention comprises contacting the population comprising the first host cell (eg, SRB) with organic and/or inorganic nitrates and nitrite. These serve to stimulate the growth of the NRB present, thus helping the NRB to outcompete SRB. Organic and inorganic nitrates or inorganic nitrites may be used injected into the certain oilfield and industrial water systems. Inorganic nitrates and inorganic nitrites available for use in the present disclosure include, for instance, potassium nitrate, potassium nitrite, sodium nitrate, sodium nitrite, ammonium nitrate, and mixtures thereof. These organic and inorganic nitrates and inorganic nitrites are commonly available, but are non-limiting and any appropriate nitrate or nitrite may be used.

[0622] The amount of organic or inorganic nitrate or nitrite used is dependent upon a number of factors, including the amount of sulphate and/or organic acids present in the population in the system, and the expected amount of NRB needed to counteract the SRB. In certain embodiment, for treating MIC of a substrate in contact with a liquid, or for treating biofouling of a liquid according to the invention, the concentration of organic or inorganic nitrate or nitrite used is less than 2000 ppm by weight of the liquid, alternatively 500 to 1600 ppm by weight or alternatively between about 900 and 1100 ppm by weight when applied using a batch application method. When applied through continuous operation, the concentration of the organic or inorganic nitrate or nitrite may be less than 500 ppm by weight, alternatively between 10 and 500 ppm, or alternatively between 10 and 100 ppm of the liquid.

[0623] In an embodiment, the population is contacted with the vectors of the invention and simultaneously or sequentially with NRB (eg, that comprise the vectors) and nitrate and/or nitrite.

[0624] Suitable NRB include any type of bacteria capable of performing anaerobic nitrate reduction, such as heterotrophic nitrate-reducing bacteria, and nitrate-reducing sulphide-oxidising bacteria. In an example, the NRB comprises one, two, three or more (eg, one or more) NRB selected from the group consisting of *Campylobacter sp. Nitrobacter sp., Thiobacillus sp., Nitrosomonas sp., Thiomicrospira sp., Sulfurospirillum sp., Thauera sp., Paracoccus sp., Pseudomonas sp. and Rhodobacter sp.* For example, the NRB is selected from one or more of *Nitrobacter vulgaris, Nitrosomonas europea, Pseudomonas stutzeri, Pseudomonas aeruginosa, Paracoccus denitrificans, Sulfurospirillum deleyianum, and Rhodobacter sphaeroides.*

[0625] In certain embodiments, the NRB is a NRB strain that is found in a crude oil, gas, petrochemical or water recovery, processing, transportation or storage system (eg, in equipment thereof), or is found in a subterranean formation, such as a water or oil well. The NRB may be optimized to metabolize under the system conditions. The NRB are, for example, selected from a library of NRB strains or may be cultured from the system to be treated or a similar system.

[0626] The amount of NRB contacted with the SRB cells in the system may depend upon a number of factors including the amount of SRB expected, as well as any biocide that may be present. When injected into subterranean formation, the permeability and porosity of the subterranean formation may be considered as well. In certain embodiments of the present disclosure, the amount of NRB injected into the liquid is between 1 and 10^8 bacteria count/ml of the liquid, or alternatively between 10 and 10^4 bacteria count/ml of the liquid.

[0627] In addition to stimulating the NRB to out compete the SRB, it may be desirable to introduce additional SRB inhibitors in certain embodiments of the present disclosure together with the nitrates. In an example, the SRB are contacted with one or more SRB inhibitors selected from the group consisting of 9,10-anthraquinone, molybdates (such as sodium molybdate and/or lithium molybdate) and mixtures thereof. In certain embodiments of the present disclosure, molybdate is added to the liquid in the range of 5 to 100 ppm by weight of liquid.

[0628] In an example, vectors of the invention and one or more biocides (ie, biocides of the first host cells, such as SRB biocides) are mixed prior to contacting the first cells with the mixture, eg, by injection of the mixture into liquid that is in contact with the surface to be treated or injection of the mixture into the fluid to be treated.

[0629] Additionally or alternatively to NRB cells containing vectors, the invention contemplates use of a species of Bacillus cells comprising vectors of the invention.

[0630] In an embodiment, the vectors are bacteriophage that are capable of infecting the SRB and the phage are contacted with the first host cells (eg, SRB), whereby CRISPR arrays comprised by the phage are introduced into first host cells for modification thereof according to the invention. In an embodiment, when the first cells are SRB, the SRB are also contacted with the phage vectors of the invention and simultaneously or sequentially with NRB. Instead of, or in addition to contacting with NRB, the SRB are contacted with nitrate and/or nitrite.

[0631] Example mechanisms involved in MIC are as follows; in an embodiment, the "controlling" using the method comprises reducing a mechanism selected from:

- Microbial (eg, bacterial) promotion of bio-mineralisation due to deposition of iron hydroxides on the metal surface, modifying the electrochemical processes at the interface metal/solution, inducing corrosion;
- Production of EPS that favours the formation of biofilm;
- Microbial (eg, bacterial) promotion of the degradation of petroleum products due to the release of the enzyme aryl hydrocarbon hydroxylase (AHH) that acts on the corrosion of metals;
- Production of sulphuric acid, which increases the corrosion process; and
- Oxidation of sulphur.

[0632] In an embodiment, the method comprises reducing a mechanism selected from:

- Bacterial promotion of bio-mineralisation due to deposition of iron hydroxides on the surface, wherein the surface is a metallic surface;
- Production of EPS;
- Bacterial promotion of the degradation of petroleum products in the system due to the release of the enzyme aryl hydrocarbon hydroxylase (AHH), wherein the surface is a metallic surface;
- Production of sulphuric acid; and
- Oxidation of sulphur.

EXAMPLES APPLICABLE TO MIC OR BIFOULING CONTROL

[0633] There are specific example applications envisioned by the present invention to reduce the corrosion and/or biofouling associated with bacteria. The applications described below are not intended to limit the concept of the present invention, and are merely illustrative of how the invention may be used to control bacterially induced corrosion or to reduce environmental pollution.

[0634] Acid Mine Drainage: In acid mine drainage, bacterial growth can increase acidity in the environment . A reaction scheme exists for the creation of acid and, therefore, potential environmental damage. The problem of acid mine drainage is recognised throughout the world as a severe environmental problem. The origin of acid mine drainage is the weathering and oxidation of pyritic and other sulphide containing minerals. Mine drainage is formed when pyrite, an iron sulfide, is exposed and reacts with air and water to form sulphuric acid and dissolved iron. Some or all of this iron can precipitate to form the red, orange, or yellow sediments in the bottom of streams containing mine drainage. The acid run-off further dissolves heavy metals such as copper, lead, mercury into ground or surface water. The rate and degree by which acid-mine drainage proceeds can be increased by the action of certain bacteria.

[0635] In an example, the system is therefore a mine or comprised by a mine. The fluid is mine drainage fluid and the method reduces sulphuric acid caused by mine drainage. In an example, the surface is in contact with mine drainage fluid.

[0636] In an example, the first host cells are *Acidithiobacillus ferrooxidans, Acidithiobacillus thiooxidans, Acidithiobacillus denitrificans,* Leptospirillum *ferrooxidans* or *Sulfobacillus thermosulfidooxidons* cells or a mixture of two or more of these.

[0637] Hydraulic Fracturing: Hydraulic fracturing is a method to fracture rock formations to facilitate the extraction of gas and other hydrocarbons. Essentially, once a gas bearing formation is identified, wells are bored into the earth in both vertical and horizontal directions to access the gas. The wells are then used to fracture the shale using high pressure water, sand and a plethora of chemicals to maintain the fractures and fissures from being closed by the intense pressure of the overburden once the hydrofracturing is completed. Millions of gallons of water are used to frac a well. Between 30% and 70% of the frac fluid returns to the surface as "flowback". Flowback contains any matter that is dissolved in the frac water, including salt. What is dissolved depends on the location. The flowback is held in plastic lined pits at the well site until it is trucked and treated prior to disposal. At some point in time the high flow and relatively low salinity water converts to a lower flow, but much higher salinity "produced water" to distinguish it from "flowback" water.

[0638] In either case the problem of microbially induced corrosion (MIC) exists. Of particular interest are the SRB. In an example, therefore, the system is hydraulic fracturing system and the fluid is a hydraulic fracturing liquid (eg, flowback water or produced water) or the surface to be treated is in contact with such a liquid. The method, for example, reduces SRB viability (eg, kills SRB and/or reduces SRB proliferation in the liquid) and the first host cells are SRB.

[0639] For example, the first host cells are *Acidithiobacillus* bacteria, *Acidithiobacillus thiooxidans, Ferrobacillus ferrooxidans, Thiobacillus thiooxidans, Thiobacillus thioparus, Thiobacillus concretivorus, Desulphovibrio* (eg, *solixigens, vulgaris, desulphuricans* or *africanus*) or *Desulphotomaculum* (eg, *orientis* or *nigrificans*) cells, or a mixture of two or more of these species.

[0640] Cooling Equipment (eg, Cooling Towers): The presence of bacteria in cooling equipment, such as cooling towers, can adversely affect the functioning of the cooling in several ways. For example, SRB support the creation of acid conditions on the walls of cooling towers, heat exchangers, etc., which leads to corrosion and potential shutdown of the cooling system while repairs are made. Additionally, biofilms on the walls of, for example, the heat exchangers, reduce the heat transfer coefficient of the heat exchangers, resulting in decreased operational efficiency of the cooling system.

[0641] Additionally, the corrosion of iron-containing components can be especially detrimental. Oxidation of iron to iron(II) and reduction of sulphate to sulphide ion with resulting precipitation of iron sulphide and generation of corrosive hydrogen ions *in situ* may take place *via* the SRB. The corrosion of iron by sulphate reducing bacteria is rapid and, unlike ordinary rusting, it is not self-limiting. Tubercles produced by *Desulphovibrio* consist of an outer shell of red ferric oxide mixed ' with black magnetic iron oxide, containing a soft, black center of ferrous sulphide.

[0642] In an example, therefore, the system is a cooling system and the fluid is a fluid (eg, water or an aqueous liquid) comprised by the system or the surface to be treated is a surface of cooling equipment in contact with such a fluid. In an example, the first host cells are SRB (eg, any SRB disclosed herein). In an example, the surface is an iron-containing surface.

[0643] In an example, the first host cells are *Legionella* cells. Such species are detrimental to human health and propagated in water cooling, heating, processing or storage equipment. In an example, therefore, the system is such an equipment.

[0644] Pipeline Corrosion: Hydrocarbon and petrochemical pipelines often include sufficient moisture to permit bacterial growth, resulting in MIC eg, caused by SRB. The MIC is often caused by biofilms of aerobic bacteria which protect SRB which is anaerobic and in direct contact with the pipeline's inner surface. This creates acid conditions and other metal-corroding conditions, which will result in localised corrosion and eventual failure of the pipe.

[0645] In an example, the system comprises an equipment surface (eg, pipeline or drilling equipment) comprising a surface in contact with the first host cells (eg, SRB). For example, the system is a crude oil, hydrocarbon, petrochemical (eg, diesel or petroleum), gas or water recovery, processing, storage or transportation system. For example, the pipeline is a petrochemicals pipeline. For example, the pipeline is comprised by an oil or gas rig. For example, the pipeline surface is in contact with sea water. For example, the pipeline surface is in contact with a petrochemical fluid, crude oil or natural gas.

[0646] Wastewater Treatment: Wastewater treatment involves adding activated sludge downstream of a wastewater treatment plant in order to remove organic pollutants. Thus, after water is treated in a waste treatment facility, many organic pollutants are present which can be "digested" by bacteria. Thus, the activated sludge is added to the treated water in a tank/container to treat the effluent from the wastewater treatment facility.

[0647] However, sometimes bacteria in the tank/container (whether originating from the activated sludge, the wastewater itself, or the surrounding environment), will dominate and grow very rapidly. Such rapid growth can result in a filamentous-shaped bacterial growth. Filaments can form up to 20-30% of the bacterial population in the tank or container, and they float. This filamentous growth results in what is known as bulking sludge. The present invention can be utilised for bulking sludge control, which is an important Aspect in wastewater treatment.

[0648] Thus, in one example, the system is a water treatment system and the surface is a surface of a container of the system, wherein the surface is in contact with water and the first host cells; or the fluid to be treated comprises said water and cells. In an example, the method controls bacterial growth in sludge of a wastewater system.

[0649] Shipping & Transportation: Ships and boats can experience MIC on their outer surfaces (eg, hulls) in contact with sea water or waterways (eg, rivers, lakes, ponds or fresh water). Inner hull surfaces can also be subject to MIC since they are typically in contact with moisture or liquids that can harbour MIC-mediating microbes such as bacteria, for example in contact with ballast water. For example, sea water is often carried in the hulls of ships (such as oil tankers) to provide stability at sea; such sea water harbours bacteria that can mediate SRB. Other transportation vehicles, such as motor-driven vehicles (cars, trucks, vans or lorries), trains, spacecraft and aircraft can also be susceptible.

[0650] Thus, in one example, the system is a transportation vehicle (eg, for transporting goods and/or people or livestock, eg, a cars, truck, van or lorry, train, spacecraft or aircraft). For example, the vehicle is a ship or boat, eg, an oil, gas or petrochemicals sea vessel (eg, an oil tanker). In an example, the surface to be treated is in contact with sea water. In an example, the surface is an outer surface of a ship or boat hull. In an example, the surface is an inner surface of a ship or boat hull.

[0651] Bacterial Persistence or Growth (Biofouling) in Ballast Water: A specific application of the invention is the treatment of marine vehicle (eg, ship or boat) ballast water to reduce undesirable bacteria, such as *Vibrio cholerae, E coli* and/or *Enterococci* sp.

[0652] Shipping moves over 90% of the world's commodities and is responsible for the global transfer of approximately 2—3 billion tons of ballast water, which is routinely carried by ships to maintain their stability. A similar volume of ballast water may also be transferred domestically within countries and regions each year (GloBallast Partnerships, www.globallast.imo.org). Ballast water has been recognized as the main source of invasive marine organisms that threaten naturally evolved biodiversity, the consequences of which are increasingly being realized (Anil et al. 2002). The unintentional introduction of disease-causing pathogenic bacteria, which are transported from the place of origin or formed during transportation, can have direct impact on society and human health. Ship ballast tanks hold different non-indigenous vertebrates, invertebrates, plants, microscopic algae, bacteria, etc. (Williams et al. 1988; Carlton and Geller 1993; Smith et al. 1996; Ruiz et al. 2000; Drake et al. 2002, 2005, 2007; Mimura et al. 2005). Microorganisms, such as bacteria are introduced into alien environments in larger numbers than other organism owing to their high natural abundance, capability to form resting stages, and capability to withstand a wide range of environmental conditions. Although all the organisms taken onboard into ballast tanks may not survive, bacteria and micro-algae are well capable of surviving prolonged periods of unfavorable conditions by forming cysts, spores, or other physiological resting stages (Roszak et al. 1983; Hallegraeff and Bolch 1992; Anil et al. 2002; Carney et al. 2011). Once released these microorganisms are well suited to be invasive owing to their small size which facilitates their passive dispersal and simpler requirements for survival than metazoans (Deming 1997). The concentration of cells of *Vibrio* species in ballast samples examined from ships in Singapore Harbour were in the range of 1.1-3.9 $\times$ $10^4 ml^{-1}$ (Joachimsthal et al. 2004). The unintentional introduction of disease-causing pathogenic bacteria can have direct societal impacts, including effects on human health. In an earlier study it was found that most of the pathogens introduced to Chesapeake Bay originated from bacteria associated with plankton rather than the water column itself (Ruiz et al. 2000). Thus, ballast water microorganisms such as bacteria and archaea are of major concern in ballast water treatment/management programs.

[0653] The International Maritime Organization (IMO) has developed a convention aimed at preventing these harmful effects, adopting the International Convention for the Control and Management of Ships' Ballast Water and Sediments (the Ballast Water Management Convention) in 2004. In the US, the United States Coast Guard's Final Rule on Ballast Water Management entered into force in June 2012, applying to ballast water discharge in US waters.

[0654] Ballast-water exchange at sea is not considered an ideal method of ballast-water management, and considerable efforts are being made to develop treatment methods. These methods must be in accordance with Standard D-2 of the IMO's Ballast Water Management Convention. Standard D2 specifies that treated and discharged ballast water must have:

- fewer than ten viable organisms greater than or equal to 50 micrometers in minimum dimension per cubic metre
- fewer than ten viable organisms less than 50 micrometres in minimum dimension and greater than or equal to 10 micrometers in minimum dimension per millilitre.

[0655] In addition, Standard D2 specifies that the discharge of the indicator microbes shall not exceed specified concentrations as follows:

- toxicogenic *Vibrio cholerae* (O1 and 0139) with less than one colony-forming unit (cfu) per 100 millilitres or less than 1 cfu per 1 gram (wet weight) zooplankton samples
- *Escherichia coli* less than 250 cfu per 100 millilitres
- intestinal *Enterococci* less than 100 cfu per 100 millilitres.

[0656] These are the indicator microbes, as a human health standard, but they are not limited to these types. Indeed, it has been suggested that in fact, in some cases the ballast water treatment used may make things worse. By removing small organisms that eat bacteria, some treatment systems have turned ballast tanks into bacteria incubators, so that the treated discharges consistently contained higher concentrations of bacteria, in some trials, thousands of times higher, than discharges that were left untreated. The increased bacteria may include human pathogens.

[0657] In an example of the disclosure (eg, according to Aspect 70 or an Aspect dependent from Aspect 70), therefore, the system is a ship or boat or marine vehicle (eg, a ship or boat, eg, an oil tanker in a harbour, dock or at sea). In an example, the fluid comprising the first host cells is ballast water of ship or boat a marine vehicle (eg, ship or boat ballast water, eg, oil tanker ballast water). In an example, the system is a sea container or a platform or rig (eg, oil or gas rig), eg at sea. In an example, the fluid is ballast water of such a container, platform or rig.

[0658] In an embodiment, the detrimental bacteria (first host cells according to the invention, eg, according to Aspect 70 or an Aspect dependent from Aspect 70) are of a species selected from the group consisting of *Vibrio cholerae; Vibrio rumoiensis; Vibrio sp.; E coli; Enterococcus sp.; Pseudomonas synxantha; Pseudomonas stutzeri; Vibrio lentus; Pseudoolteromonos marina Pseudoalteromonas tetraodonis; Pseudoalteromonas sp.; Pseudomonas putida; Pseudomonas oleovorans; Vibrio splendidus; Vibrio cyclitrophicus; Enterococcus hirae; Enterococcus faecium Vibrio rotiferianus; Pseudoalteromonas undina; Serratia plymuthica; Pseudomonas fulva; Pseudomonas tolaasii; Pseudomonas stutzeri; Pseu-*

domonas stutzeri; Vibrio tubiashii; Holomonos venusta; Idiomarina loihiensis; Vibrio cyclitrophicus; Vibrio tubiashii; Serratia plymuthica; Pseudoalteromonas sp.; Pseudoalteromonas atlantica; Pseudomonas synxantha; Pseudomonas stutzeri; Pseudoalteromonas carrageenovora; Tenacibaculum sp.; Bacillus mycoides; Vibrio natriegens; Bacillus baekryungensis; Enterococcus hirae; Lactobacillus pentosus; Pseudoalteromonas carrageenovora; and *Pseudomonas aeruginosa.*

[0659] In an example, the first host cells are aerobic heterotrophic bacteria. In an example, the first host cells are *Vibrio cholerae* cells (eg, strain O1 and/or O139). In an example, the first host cells are *E coli* cells. In an example, the first host cells are *Enterococcus* sp. cells.

[0660] "Characterization of Bacteria in Ballast Water Using MALDI-TOF Mass Spectrometry", Kaveh E et al, PLoS One. 2012; 7(6): e38515; Published online 2012 Jun 7. doi: 10.1371/journal.pone.0038515 (incorporated herein by reference) discloses a suitable rapid and cost-effective method for monitoring bacteria in ballast water.

[0661] A specific example of the invention is as follows:-

[0662] A method of controlling bacterial biofouling in ballast water of a ship or boat, wherein the water comprises a population of first host cells of a first microbial species (such as *Cholera, E coli* or *Enterococci* sp) that mediates said biofouling, the method comprising

(i) contacting the population with a plurality of vectors that are capable of transforming or transducing the cells, each vector comprising a CRISPR array whereby CRISPR arrays are introduced into the host cells, wherein

(a) each CRISPR array comprises one or more sequences for expression of a crRNA and a promoter for transcription of the sequence(s) in a host cell; and
(b) each crRNA is capable of hybridising to a target sequence of a host cell to guide Cas (eg, a Cas nuclease) in the host cell to modify the target sequence (eg, to cut the target sequence); the target sequence being a gene sequence for mediating host cell viability; and

(ii) allowing expression of said cRNAs in the presence of Cas in host cells, thereby modifying target sequences in host cells, resulting in reduction of host cell viability and control of said biofouling.

[0663] In an example, step (i) comprises mixing the ballast water with the vectors, eg, in the hull of a ship or boat.

[0664] In an example, the ship or boat is a marine vehicle and the water is sea water. Instead of a ship or boat, in an alternative the ballast water is comprised by a container or a drilling platform at sea, eg, an oil platform or oil rig.

[0665] The disclosure also comprises a method of discharging ballast water from a ship or boat, wherein the discharged ballast water comprises water treated by the method of the specific example above. In an example, the water is discharged into a body of water, eg, a sea, ocean or waterway (eg, a river, canal, lake or reservoir).

[0666] The disclosure also comprises ship or boat ballast water comprising CRISPR arrays, wherein the ballast water is obtained or obtainable by the specific example above. The disclosure also comprises a sea container ballast water comprising CRISPR arrays, wherein the ballast water is obtained or obtainable by the specific example above. The disclosure also comprises ballast water of a platform or rig (eg, oil or gas rig) at sea, the water comprising CRISPR arrays, wherein the ballast water is obtained or obtainable by the specific example above. The arrays are as recited in (a) and (b) of the specific example.

References:

[0667]

1. *Anil AC, Venkat K, Sawant SS, Dileepkumor M, Dhargalkar VK, Ramaiah N, Harkantra SN and* Ansari ZA (2002) Marine bioinvasion: Concern for Ecology and Shipping. Current Science 83(3): 214-218*;*

2. Azam F and Malfatti F (2007) Microbial structuring of marine ecosystems. Nature Reviews Microbiology 5: 782-791*;*

3. Belkin S and Colwell RR (2005) Ocean and health: pathogens in the marine environment. New York, NY:Springer;

4. Carlton JT and Geller JB (1993) Ecological roulette: the global transfer of non-indigenous marine organisms. Science 261: 78-82*;*

5. Carman KR and Dobbs FC (1997) Epibiotic microorganisms in copepods and other marine crustaceans. Microscopy Research and Technique 37: 116-135*;*

6. Carney KJ, Delany JE, Sawant SS, Mesbahi E (2011) The effects of prolonged darkness on temperate and tropical marine phytoplankton, and their implications for ballast water risk management. Marine Pollution Bulletin 62(6):1233-1244*;*

7. Colwell RR (1996) Global climate and infectious disease: the cholera paradigm. Science 274: 2025-2031*;*

8. Conway DVP, White RG, Hugues-Dit-Ciles J, Gallienne CP, Robins DB (2003) Guide to the coastal and surface

zooplankton of the southwestern Indian Ocean. In: Marine Biological Association of the United Kingdom Occasional Publication. UKDEFRA Darwin Initiative Project 162/09/004 Zooplankton of the Mascarene Plateau, vol 15, pp 1-354;
9. Daley RJ and Hobbie JE (1975) Direct counts of aquatic bacteria by a modified epifluorescence technique. Limnology and Oceanography 20: 875-882;
10. Deming JW (1997) Unusual or extreme high-pressure marine environments. In: ASM Manual of Environmental Microbiology, Hurst CJ, Knudsen GR, Mclnerney MJ, Stetzenbach LD, Walter MV (editors), Washington, DC: ASM Press, pp 366-376;
11. Drake LA, Ruiz GM, Galil BS, Mullady TL, Friedmann DO and Dobbs FC (2002) Microbial ecology of ballast water during a transoceanic voyage and the effects of open-ocean exchange. Marine Ecology Progress Series 233: 13-20;
12. Drake LA, Meyer AE, Forsberg RL, Baier RE, Doblin MA, Heinemann S, Johnson WP, Koch M, Rublee PA and Dobbs FC (2005) Potential invasion of microorganisms and pathogens via 'interior hull fouling': biofilms inside ballast water tanks. Biological Invasions 7: 969-982;
13. Drake LA, Doblin MA and Dobbs FC (2007) Potential microbial bioinvasions via ship's ballast water, sediment and biofilm. Marine Pollution Bulletin 55: 333-341;
14. Dawson MP, Humphrey BA and Marshall KC (1981) Adhesion: A tactic in the survival strategy of a marine Vibrio during starvation. Current Microbiology 6: 195-199;
15. Foladori P, Bruni L, Andreottola G and Ziglio G (2007) Effects of sonication on bacteria viability in wastewater treatment plants evaluated by flow cytometry-fecal indictors, wastewater and activated sludge. Water Research 41: 235-243;
16. Hallegraeff GM and Bolch CJ (1992) Transport of diatoms and dinoflagellate resting spores in ships ballast water: implications for plankton biogeography and aquaculture. Journal of Plankton Research 14(8): 1067-1084;
17. Harris JM (1993) The presence, nature and role of gut microflora in aquatic invertebrates: a synthesis. Microbial Ecology 25: 195-231;
18. Heidelberg JF, Heidelberg KB and Colwell RR (2002) Bacteria of the gamma-subclass Proteobacteria associated with zooplankton in Chesapeake Bay. Applied and Environmental Microbiology 68: 5498-5507;
19. Hood MA, Ness GE, Rodrick GE, Blake NJ (1984) The ecology of Vibrio cholerae in two Florida estuaries. In: Vibrios in the Environment, Colwell R (editor), New York, NY: Wiley, pp 399-409;
20. Huq A, West PA, Small EB and Colwell RR (1984) Influence of water temperature, salinity and pH on survival and growth of toxigenic Vibrio cholerae serovar O1 associated with live copepods in laboratory microcosms. Applied and Environmental Microbiology 48: 420-424;
21. International Maritime Organization (IMO) (2004) International convention for the control and management of ships' ballast water and sediments. London: International Maritime Organization;
22. Joachimsthal EL, Ivanov V, Toy ST-L and Tay J-H (2004) Bacteriological examination of ballast water in Singapore Harbour by flow cytometry with FISH. Marine Pollution Bulletin 49: 334—343;
23. Jyoti KK and Pandit AB (2001) Water disinfection by acoustic and hydrodynamic cavitation. Biochemical Engineering Journal 7: 201-212;
24. Kasturirangan LR (1963) A key for the identification of the more common planktonic copepoda of Indian coastal waters. In: Indian National Committee on Oceanic Research, Panikkar NK (editor), New Delhi: Council of Scientific and Industrial Research, p 87;
25. Khandeparker L & Anil AC; Ecohealth. 2013 Sep;10(3):268-76. doi: 10.1007/s10393-013-0857-z. Epub 2013 Jul 12, "Association of bacteria with marine invertebrates: implications for ballast water management";
26. Krieg NR (1984) Bergey's manual of systematic bacteriology, Vol 1. Williams & Wilkins, Baltimore;
27. Lee BG and Fisher NS (1992) Decomposition and release of elements from zooplankton debris. Marine Ecology Progress Series 88: 117-128;
28. Lloyd's Register (2010) Ballast water treatment technology guide;
29. McFall-Ngai MJ and Ruby EG (1991) Symbiont recognition and subsequent morphogenesis as early events in an animal-bacterial mutualism. Science 254: 1491-1494;
30. Mimura H, Katakura R and Ishida H (2005) Changes of microbial populations in a ship's ballast water and sediments on a voyage from Japan to Qatar. Marine Pollution Bulletin 50: 751-757;
31. Munro PM and Colwell RR (1996) Fate of Vibrio cholerae O1 in sea water microcosms. Water Research 30(1): 47-50;
32. Pfeffer C and Oliver DJ (2003) A comparison of thiosulphate-citrate-bile salts-sucrose (TCBS) agar and thiosulphate-chloride-iodide (TCI) agar for the isolation of Vibrio species from estuarine environments. Letters in Applied Microbiology 36: 150-151;
33. Peter H and Sommaruga R (2008) An evaluation of methods to study the gut bacterial community composition of fresh water zooplankton. Journal of Plankton Research 30: 997-1006;
34. Polz MF, Distel DL, Zarda B, Amann R, Felbeck H, OttJA and Cavanaugh CM (1994) Phylogenetic analysis of

a highly specific association between ectosymbiotic, sulfuroxidizing bacteria and a marine nematode. Applied and Environmental Microbiology. 60: 4461-4467;
35. Pruzzo C, Vezzulli L and Colwell RR (2008) Global impact of Vibrio cholerae interactions with chitin. Environmental Microbiology 10(6):1400-1410;
36. Rehnstam-Holm A-S, Godhe A, Härnström K, Raghunath P, Saravanan V, Collin B, Karunasagar I and Karunasagar I (2010) Association between phytoplankton and Vibrio spp. along the southwest coast of India: a mesocosm experiment. Aquatic Microbial Ecology 58: 127-139;
37. Richard C (1993) Chromobacterium violaceum, opportunistic pathogenic bacteria in tropical and subtropical regions. Bulletin de la Société de Pothologie exotique 86: 169-173;
38. Roszak DB, Grimes DJ and Colwell RR (1983) Viable but nonrecoverable stage of Salmonella enteritidis aquatic systems. Canadian Journal of Microbiology 30: 334-338;
39. Ruiz GM, Rawlings TK, Dobbs FC, Drake LA, Mullady T, Huq A and Colwell RR (2000) Global spread of microorganisms by ships—Ballast water discharged from vessels harbours a cocktail of potential pathogens. Nature 408: 49-50;
40. Sawant SS, Anil AC, Krishnamurthy V, Gaonkar C, Kolwalkar J, Khandeparker L, Desai DV, Mahulkar AV, Ronade VV and Pandit AB (2008) Effect of hydrodynamic cavitation on zooplankton: A tool for disinfection. Biochemical Engineering Journal 42: 320-328;
41. Seth N, Chakravarty P, Khandeparker L, Anil AC and Pandit AB (2010) Quantification of the energy required for the destruction of Balanus amphitrite larva by ultrasonic treatment. Journal of the Marine Biological Association of the United Kingdom 90 (7):1475-1482;
42. Smith LD, Wonham MJ, McCann LD, Reid DM, Carlton JT, Ruiz GM (1996) Biological Invasions by Nonindigenous Species in United States Waters: Quantifying the Role of Ballast Water and Sediments. Parts I and II. Report Number CG-D-02-97, Groton, CT: US Coast Guard Research and Development Center;
43. Tang KW (2005) Copepods as microbial hotspots in the ocean: effects of host feeding activities on attached bacteria. Aquatic Microbial Ecology 38: 31-40;
44. Tang KW, Freund CS and Schweitzer CL (2006) Occurrence of copepod carcasses in the lower Chesapeake Bay and their decomposition by ambient microbes. Estuarine, Coastal and Shelf Science 68: 499-508;
45. Tang KW, Dziallas C, Hutalle-Schmelzer K and Grossart HP (2009) Effects of food on bacterial community composition associated with the copepod Acartia tonsa Dana. Biological Letters 5: 549-553;
46. Thompson CC, Thompson FL, Vandemeulebroecke K, Hoste B, Dawyndt P and Swings J (2004) Use of recA as an alternative phylogenetic marker in the family Vibrionaceae. International Journal of Systematic and Evolutionary Microbiology 54:919-929;
47. Todd CD, Laverack MS, Boxshall GA (1966) Coastal marine zooplankton: A practical manual for students, London, UK: The Natural History Museums, p 106;
48. Viñas MD, Diovisolvi NR and Cepeda GD (2010) Individual biovolume of some dominant copepod species in coastal waters off Buenos Aires Province, Argentine sea. Brazilian Journal of Oceanography 58(2): 177-181;
49. Williams RJ, Griffiths FB, Van der Wal EJ and Kelly J (1988) Cargo vessel ballast water as a vector for the transport of nonindigenous marine species. Estuarine, Coastal and Shelf Science 26: 409-420.

[0668] The invention also provides vectors and CRISPR arrays as follows.

[0669] An eighty-fifth aspect relates to a vector comprising a CRISPR array for introduction into a bacterial host cell, wherein the bacterium is capable of water-borne transmission, wherein

(a) the CRISPR array comprises a sequence for expression of a crRNA and a promoter for transcription of the sequence in a said host cell;
(b) the crRNA is capable of hybridising to a host cell target sequence to guide a Cas (eg, a Cas nuclease) in the host cell to modify the target sequence (eg, to cut the target sequence); the target sequence being a nucleotide sequence (eg, a gene or regulatory sequence) for mediating host cell viability;
(c) wherein the sequence of (a) comprises a sequence R1-S1-R1' for expression and production of the crRNA, wherein R1 is a first CRISPR repeat, R1' is a second CRISPR repeat, and R1 or R1' is optional; and S1 is a first CRISPR spacer that comprises or consists of a nucleotide sequence that is 80% or more identical (eg, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identical) to the host cell target sequence. By "water-borne transmission" is meant that cells of said bacterium are capable of being spread in water or an aqueous liquid between different organisms, within an organism, between different environments, or between an organism and an environment. Examples are *Vibro cholera, Enterococcus spp* and *E coli. Vibrio cholerae* is a gram negative comma-shaped bacterium with a polar flagellum. It belongs to the class of the Gamma Proteobacteria. There are two major biotypes of *V. cholerae,* classical and El Tor, and numerous serogroups. *V. cholerae* is the etiological agent of cholera, a severe bacterial infection of the small intestine, and a major cause of death in developing countries. The pathogenicity

genes of V. cholerae are interesting targets to detect and to study *V. cholerae* infections. Most of these genes are located in two pathogenicity islands, named TCP (Toxin-Coregulated Pilus) and CTX (Cholera ToXins), organized as prophages 1,2. TCP contains a cluster of genes involved in host adhesion via pili, while CTX genes are involved in the synthesis of the cholera toxin3.

[0670] In an embodiment, the vector is an isolated vector (ie, a vector not in a said host cell). In an example, the vector is an engineered or synthetic vector (ie, a non-naturally occurring vector).

[0671] In an example, the array is an ICP1 array, ie, an array of an ICP1 *V cholerae* phage, eg, wherein the phage is ICP1_2003_A, ICP1_2004_A, ICP1_2005_A, ICP1_2006_E or ICP1_20011_A. In an example the array is a CR1 or CR2 ICP1 phage array, eg, an engineered or non-naturally occurring derivative of such an array. In an example, the CRISPR array and Cas are type 1-E or type 1-F, eg, subtype system 17.

[0672] In an example, the the CRISPR array comprises a plurality of sequences, each for expression of a respective crRNA and a associated with a promoter for transcription of the sequence in a said host cell. In an example, the vector or each vector comprises a plurality (eg, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) of said CRISPR arrays.

[0673] In an example, the vector comprises a nucleotide sequence encoding said Cas.

[0674] In another example, the vector is devoid of such a sequence. For example, in this case, the array(s) are operable with one or more Cas produced by the host cell.

[0675] An eighty-sixth aspect relates to the vector of the eighty-fifth Aspect, wherein the cell is a *Vibrio cholerae, Enterococcus* or *E coli* cell.

[0676] An eighty-seventh aspect relates to the vector of the eighty-fifth or eighty-sixth Aspect, wherein the vector is devoid of a nucleotide sequence that is capable of expressing said Cas.

[0677] In an example, therefore the vector does not encode a Cas nuclease. In an alternatiave the vector encodes a said Cas.

[0678] An eighty-eighth aspect relates to the vector of the eighty-fifth, eighty-sixth or eighty-seventh Aspect, wherein the target sequence is a protospacer sequence of 17-45 contiguous nucleotides, eg, 18, 19, 20 or 21 contiguous nucleotides.

[0679] In an example, each spacer (S1) is a nucleotide sequence of 17-45 contiguous nucleotides, eg, 18, 19, 20, 21, 30, 31, 32 or 33 contiguous nucleotides. The protospacer sequence for *V cholerae* PLE is, for example, 32 contiguous nucleotides and a vector targeting this can, for example, have a spacer sequence of 32 contiguous nucleotides that is 100% or at least 80, 90 or 95% identical to the 32 nucleotide PLE sequence. Where the vector comprises a plurality of spacers, the spacers can be a mixture of different spacers, or can be identical spacers. For example, the array comprises a plurality of spacers, wherein a sub-set of spacers are identical. The identical spacers can be homologous to protospacer sequence of a gene encoding a pathogenicity factor of the host cell, for example. Using multiple spacers may be advantageous if the host cuts one or more of the spacers once the vector is inside the cell - uncut spacers are still able to form crRNAs and home to target sequences. Using a mixture of different spacers in the vector or in an array is advantageous to minimise the risk of adaptation of the host to the invading vector, thereby minimising resistance.

[0680] An eighty-ninth aspect relates to the vector of any one of the eighty-fifth to eighty-eighth Aspects, wherein the target sequence is a virulence, resistance or essential gene sequence. In an example, the target sequence is a sequence of a PICI-like element (PLE), *eg, a V. cholerae* PLE. Eg, PLE1.

[0681] A ninetieth aspect relates to the vector of any one of the eighty-fifth to eighty-ninth Aspects, wherein the target sequence is a pathogenicity island sequence, *optionally* wherein the host cell is a *Vibrio cholera* cell and the target sequence is a TCP, CTX or VPI sequence. In an example (eg, wherein the host is *Vibrio*) pathogenicity island is TCP (Toxin-Coregulated Pilus) or CTX (Cholera ToXins). The *Vibrio* pathogenicity island (VPI) contains genes primarily involved in the production of toxin coregulated pilus (TCP). It is a large genetic element (about 40 kb) flanked by two repetitive regions (*att*-like sites), resembling a phage genome in structure. The VPI contains two gene clusters, the TCP cluster, and the ACF cluster, along with several other genes. The *acf* cluster is composed of four genes: *acfABCD.* The *tcp* cluster is composed of 15 genes: *tcpABCDEFHIJPQRST* and regulatory gene *toxT.*

[0682] A ninety-first aspect relates to the vector of any one of the eighty-fifth to nintieth Aspects, wherein the host cell is *Vibrio cholera* and the target sequence is a CTXϕ gene sequence. The genes for cholera toxin are carried by CTXphi (CTXϕ), a temperate bacteriophage inserted into the *V. cholerae* genome. CTXϕ can transmit cholera toxin genes from one *V. cholerae* strain to another, one form of horizontal gene transfer. The genes for toxin coregulated pilus are coded by the VPI pathogenicity island (VPI).

[0683] A ninety-second aspect relates to *the vector of any* one of the eighty-fifht to ninetieth Aspects, wherein the host cell is *Vibrio cholera* and the target sequence is a ctxB, tcpA, ctxA, tcpB, wbet, hlyA, hapR, rstR, mshA or tcpP sequence.

[0684] A ninety-third aspect relates to the vector of any one of the eighty-fifth to ninety-second Aspects, wherein the target sequence is 17-45 contiguous nucleotides (eg, 18, 19, 20 or 21 contiguous nucleotides) and at least 80% (eg, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identical) identical to a sequence of a phage inducible chromosomal island (PICI) of a Gram-positive bacterium, eg, a *Staphylococcus aureus* pathogenicity island (SaPI). Examples of ICP1

phage (aka ICP1-related phage) spacer sequences are provided below. These spacer sequences are homologous to target sequences (protospacer sequences) in *V cholera.*

[0685] A ninety-fourth aspect relates to the vector of any one of the eighty-fifth to ninety-third Aspects, wherein the host cell is *Vibrio cholera* and the crRNA is capable of hybridising to a target sequence within 5, 4 or 3 nucleotides of a protospacer adjacent motif (PAM) in a *Vibrio cholerae* cell, wherein the PAM is GA.

[0686] A ninety-fifth aspect relates to the vector of any one of the eight-fifth to ninety-fourth Aspects, wherein the host cell is *Vibrio cholera* and the cell is an EI Tor, O1 or 0139 *Vibrio cholerae* cell. In an example, the *V. cholerae* is serotype O1 EI Tor N16961; EI Tor biotype18; or EI Tor strain MJ-1236. In an example, the host cell is a E. coli O157:H7 cell.

[0687] A ninety-sixth aspect relates to the vector of any one of the eighty-fifth to ninety-fifth Aspects, wherein the vector is a bacteriophage that is capable of infecting a said host cell. In an example, the host cell is *E coli,* and the phage is a lambda or T4 phage. In an example, the host cell is an *Enterococcus* cell and the phage is a *Enterococcus* phage IME-EF1, phiEF24C, φEf1 or EFDG1 (see Appl Environ Microbiol. 2015 Apr;81(8):2696-705. doi: 10.1128/AEM.00096-15. Epub 2015 Feb 6, "Targeting Enterococcus faecalis biofilms with phage therapy", Khalifa L *et al*).

[0688] A ninety-seventh aspect relates to the vector of the ninety-sixth Aspect, wherein the host cell is *Vibrio cholera* and the vector is a bacteriophage capable of infecting a *Vibrio cholerae* cell.

[0689] A ninety-eighth aspect relates to the vector of the ninety-seventh Aspect, wherein the bacteriophage is selected from CTXϕ, an ICPI phage and a myovirus, eg, wherein the phage is ICP1_2003_A, ICP1_2004_A, ICP1_2005_A, ICP1_2006_E or ICP1_20011_A, optionally an engineered and non-naturally occurring phage.

[0690] A ninety-ninth aspect relates to the vector of any one of the eighty-fifth to ninety-eighth Aspects, wherein the vector is or comprises an ICE, eg, a transposon. The ICE can comprise any of the features of an ICE described herein.

[0691] A one-hundreth aspect relates to the vector of the ninety-ninth Aspect, wherein the transposon is a conjugative transposon capable of transfer from a first to a second said host cell.

[0692] A one-hundredth and first aspect relates to the vector of the ninety-ninth or one hundredth Aspect, wherein the transposon leaves a copy of the CRISPR array in the first cell.

[0693] 85. A one-hundreth and second aspect relates to the vector of any one of the eighty-fifth to one hundredth and first Aspects, wherein the or each array is comprised by a respective mobile genetic element (MGE), wherein the MGE comprises an origin of transfer (*ori*T) operable in the host cell. The MGE can be according to any MGE described herein.

[0694] A one-hundredth and third aspect relates to the vector of any one of the eighty-fifht to one hundredth and second Aspects, wherein the vector is an engineered vector.

[0695] A one-hundredth and fourth aspect relates to a water or food treatment composition comprising a plurality of vectors according to any one of the eighty-fifth to one hundredth and third Aspects.

[0696] In an example, the water is ballast water, sea water, brackish water, fresh water, drinking water, waterway water (eg, estuary water) or industrial water. In an example, the water is water in human GI tract fluid.

[0697] In an example, the host cell is comprised by shellfish, fish, rice or grains. In an example, the composition is for treating food and the host cell is a *E. coli* O157:H7 cell. In an example, the target sequence is a sequence encoding a Shiga toxin in an *E coli* (eg, O157:H7) host cell. In an alternative to the water-borne species described so far, the host cell is a *Salmonella* or *Listeria* cell.

[0698] A one-hundredth and fifth aspect relates to a medicament for treatment or prevention of *Vibrio cholerae* infection in a human, the medicament comprising a plurality of vectors according to any one of the eighty-fifth and one hundredth and third Aspects. In an alternative, the invention provides a medicament for treatment or prevention of *E coli* infection in a human, the medicament comprising a plurality of vectors of the invention. In an alternative, the invention provides a medicament for treatment or prevention of *Enterococcus* infection in a human, the medicament comprising a plurality of vectors of the invention.

[0699] A one-hundredth and sixth aspect relates to the composition or medicament of the one-hundredth and fourth or one-hundredth and fifth Aspect, further comprising an anti-host cell antibiotic or an anti-host cell biocide.

[0700] Example 5 below is an example relating to cholera.

[0701] Any of the general features (see below) also may apply to the present configuration (sixth configuration). Any configuration below is combinable with the present configuration, eg, to provide combinations of features for inclusion in one or more claims herein.

REGULATING CAS ACTIVITY

[0702] These aspects of the invention are useful for regulating Cas activity, eg, in a cell or *in vitro.* The invention involves targeting a Cas-encoding gene to restrict Cas activity, which is advantageous for temporal regulation of Cas. The invention may also be useful in settings where increased stringency of Cas activity is desirable, eg, to reduce the chances for off-target Cas cutting in when modifying the genome of a cell. Applications are, for example, in modifying human, animal or plant cells where off-target effects should be minimised or avoided, eg, for gene therapy or gene targeting of the cell or a tissue or an organism comprising the cell. For example, very high stringency is required when

using Cas modification to make desired changes in a human cell (eg, iPS cell) that is to be administered to a patient for gene therapy or for treating or preventing a disease or condition in the human. The disclosure provides these applications as part of the methods and products of the invention.

Disclosed herein are the following clauses:-

[0703] In a first clause, there is provided a method of modifying an expressible gene encoding a first Cas, the method comprising

(a) combining a guide RNA (gRNA1) with the Cas gene in the presence of first Cas that is expressed from said gene; and
(b) allowing gRNA1 to hybridise to a sequence of said Cas gene (eg, a promoter or a first Cas-encoding DNA sequence thereof) and to guide first Cas to the gene, whereby the Cas modifies the Cas gene.

[0704] In an example, the method is a cell-free method (eg, recombineering method) *in vitro.* In another example, the method is carried out in a cell, eg, wherein the gene is cut by Cas that it encodes itself (ie, endogenous Cas is used to cut the gene).

[0705] A second clause relates to the method of the first clause, wherein the Cas is a nuclease and the Cas gene is cut.

[0706] A third clause relates to the method of the first and second clause, wherein the Cas gene is mutated, down-regulated or inactivated.

[0707] A fourth clause relates to the method of any one of the first to third clauses, wherein the first Cas is a Cas9.

[0708] A fifth clause relates to the method of any one of the first to fourth clauses, wherein gRNA1 is a single guide RNA.

[0709] A sixth clause relates to the method of any one of the first to fifth clauses, wherein the method is carried out in a host cell.

[0710] A seventh clause relates to the method of the sixth clause, wherein the cell is a prokaryotic cell, eg, a bacterial or archaeal cell (eg, an *E coli* cell).

[0711] An eighth clause relates to the method of the sixth clause, wherein the method is a recombineering method.

[0712] A ninth clause relates to the method of the sixth, seventh or eighth clause, wherein the cell is of a human or non-human animal pathogen species or strain (eg, *S aureus*).

[0713] A tenth clause relates to the method of any one of the sixth to ninth clauses, wherein the cell is a cell of a human microbiome species, eg, a human gut microbiome species.

[0714] An eleventh clause relates to the method of any one of the sixth to tenth clauses, wherein the Cas gene is comprised by a host CRISPR/Cas system.

[0715] Optionally an exogenous first Cas-encoding sequence is not used in the method, for example when the host cell comprises a wild-type endogenous Cas nuclease that is cognage to gRNA1.

[0716] A twelfth clause relates to the method of the sixth clause, wherein the cell is a eukaryotic cell (eg, a human, non-human animal, yeast or plant cell).

[0717] A thirteenth clause relates to the method of the twelfth clause, wherein the method is carried out in a non-human embryo; non-human zygote; non-human germ cell; or a human or animal (eg, wherein the method is a cosmetic method); optionally wherein the method is not a method for treatment of the human or animal body by surgery or therapy or diagnosis.

[0718] A fourteenth clause relates to the method of any one of the sixth to third clauses, wherein the Cas gene is comprised by a nucleic acid that is introduced into the cell in step (a).

[0719] A fifteenth clause relates to tThe method of any one of the sixth to fourth clauses for reducing the development of host cell resistance to transformation by a nucleic acid vector or maintenance of a nucleic acid vector in the host cell.

[0720] A sixteenth clause relates to the method any one of the first to fifteenth clauses, or a nucleic acid or cell product thereof for human or animal medical therapy, prophylaxis or diagnosis (eg, for gene therapy of a human or animal, human or animal cell when the method is carried out in a human or animal cell; or for treating or preventing a bacterial infection in a human or animal when the method is carried out in a bacterial cell).

[0721] A seventeenth clause relates to the method any one of the first to sixth clauses, wherein the method is carried out *in vitro.*

[0722] An eighteenth clause relates to the method any one of the first to sixteenth clauses, wherein the method is carried out *in vivo,* optionally not in a human embryo and optionally wherein the method is not a method for treatment of the human or animal body by surgery or therapy or diagnosis.

[0723] A nineteenth clause relates to tThe method any one of the first to eighteenth clauses, wherein gRNA1 is produced by transcription from a first nucleic acid that is combined with the Cas gene in step (a).

[0724] A twentieth clause relates to the method of the nineteenth clause, wherein the method is carried out in a cell and the first nucleic acid encoding gRNA1 is introduced into the cell in step (a); or a first nucleic acid encoding a crRNA

is introduced into the cell in step (a) wherein the crRNA forms gRNA1 with a tracrRNA in the cell.

[0725] A twenty-first clause relates to the method of the nineteenth or twentieth clause, wherein the Cas gene is combined with a target nucleic acid comprising a target site (CS-t) to be modified by first Cas, and wherein

I. the Cas gene comprises a first protospacer (PS1) adjacent a PAM (P1) that is cognate to the first Cas, wherein PS1 is modified (eg, cut) at a first site (CS1) by first Cas;
II. gRNA1 comprises a sequence that is complementary to PS1 for guiding first Cas wherein PS1 is modified at CS1 in step (b);
III. the target nucleic acid comprises a protospacer sequence (PS-t) adjacent a PAM (P-t), wherein P-t is cognate to the first Cas;
IV. before or during step (b) the method comprises combining a guide RNA (gRNA-t) with the target nucleic acid and first Cas expressed from said gene, wherein gRNA-t hybridises to PS-t and guides first Cas to modify CS-t; and
V. the method optionally comprises isolating or sequencing the modified target nucleic acid.

[0726] A twenty-second clause relates to the method of the twenty-first clause, wherein gRNA-t is produced by transcription from a nucleic acid (eg, said first nucleic acid) that is combined with the Cas in step IV.

[0727] A twenty-third clause relates to the method of the twenty-second clause, wherein the method is carried out in a cell and the nucleic acid encodes a crRNA, wherein the crRNA forms gRNA-t with a tracrRNA in the cell.

[0728] A twenty-fourth clause relates to the method of any one of the twenty-first to twenty-third clauses, wherein the production of gRNA1 is commenced after the production of gRNA-t, whereby PS-t is modified (eg, cut) in copies of the target nucleic acid before PS1 is modified (eg, cut) to down-regulate or inactivate first Cas expression.

[0729] A twenty-fifth clause relates to the method of any one of the first to twenty-fourth clauses, further comprising combining the cut target nucleic acid with a further nucleic acid, whereby homologous recombination between the nucleic acids takes place and

(i) a nucleotide sequence of the target nucleic acid is deleted;
(ii) a nucleotide sequence of the further nucleic acid is deleted;
(iii) a nucleotide sequence of the target nucleic acid is inserted into the further nucleic acid; and/or
(iv) a nucleotide sequence of the further nucleic acid is inserted into the target nucleic acid.

[0730] A twenty-sixth clause relates to the method of the twenty-fifth clause, wherein (i) takes place, thereby inactivating a nucleotide sequence or regulatory element of the target nucleic acid.

[0731] A twenty-sevneth clause relates to the method of the twenty-fifth clause, wherein (i) takes place, thereby activating a nucleotide sequence or regulatory element of the target nucleic acid.

[0732] A twenty-eighth clause relates to the method of the twenty-fifth, twenty-sixth or twenty-seventh clause, wherein (ii) takes place, thereby inactivating a nucleotide sequence or regulatory element of the further nucleic acid.

[0733] A twenty-ninth clause relates to the method of the twenty-fifth, twenty-sixth or twenty-seventh clause, wherein (ii) takes place, thereby activating a nucleotide sequence or regulatory element of the further nucleic acid.

[0734] A thirtieth clause relates to the method of any one of the twenty-fifth to twenty-ninth clauses, wherein (iii) takes place, optionally placing the inserted sequence in functional relationship with a regulatory element of the further nucleic acid and/or creating a new marker sequence.

[0735] A thirty-first clause relates to the method of any one of the twenty-fifth to thirtieth clauses, wherein (iv) takes place, optionally placing the inserted sequence in functional relationship with a regulatory element of the target nucleic acid and/or creating a new marker sequence.

[0736] A thirty-second clause relates to the method of the thirtieth or thirty-first clause, further comprising detecting the new marker sequence or an expression product thereof to determine that homologous recombination has taken place.

[0737] A thirty-second clause relates to the method of any one of the twenty-first to thirty-second clauses, further comprising isolating or sequencing the target nucleic acid product, the further nucleic acid product and/or the first vector product.

[0738] A thirty-fourth clause relates to the method of any one of the first to thiry-third clauses, wherein the first vector is as defined in any one of the thirty-fifth to fifty-fifth clauses.

[0739] A thirty-fifth clause relates to a first (eg, isolated) nucleic acid vector or combination of vectors, eg, for use in the method of the first clause, wherein

(a) the first vector or a vector of said combination comprises an expressible nucleotide sequence that encodes a guide RNA (gRNA1, eg, a single gRNA) that is complementary to a predetermined protospacer sequence (PS1) for guiding a first Cas to modify PS1 at a first site (CS1), wherein PS1 is adjacent a PAM (P1) that is cognate to the first Cas; or the expressible sequence encodes a crRNA that forms gRNA1 with a tracrRNA; and

(b) PS1 and P1 are sequences of an expressible first Cas-encoding gene and PS1 is capable of being modified at CS1 by the first Cas.

[0740] Each vector herein in any configuration can be a linear or circular (eg, closed circular, optionally supercoiled) DNA carrying the specified sequence(s).

[0741] A thirty-sixth clause relates to the vector or combination of the thirty-fifth clause, wherein the first Cas is a nuclease, wherein CS1 is capable of being cut by the nuclease.

[0742] A thirty-seventh clause relates to the vector or combination of the thirty-fifth or thirty-sixth clause, wherein the first Cas is a Cas9.

[0743] A thirty-eight clause relates to the vector or combination of any one of the thirty-fifth to thirty-seventh clauses, wherein gRNA1 is a single guide RNA.

[0744] A thirty-ninth clause relates to the vector or combination of any one of the thirty-fifth to thirty-eighth clauses, wherein the nucleotide sequence is expressible in a prokaryotic cell (eg, a bacterial or archaeal cell) for producing gRNA1.

[0745] A fortieth clause relates to a recombineering kit comprising the vector or combination of clause 39 (eg, wherein the cell is a recombineering-permissive *E coli* cell).

[0746] A forty-first clause relates to the vector or combination of any one of the thirty-fifth to thirty-eighth clauses, wherein the nucleotide sequence is expressible in a eukaryotic cell (eg, a human, animal, plant or yeast cell) for producing gRNA1.

[0747] A forty-second clause relates to the vector or combination of any one of the thirty-fifth to forty-first clauses, wherein the first vector or a vector of said combination (eg, the second vector) comprises an expressible nucleotide sequence that encodes a guide RNA (gRNA-t, eg, a single gRNA) that is complementary to a predetermined protospacer sequence (PS-t) of a target nucleic acid for guiding first Cas to modify (eg, cut) PS-t; or the expressible sequence encodes a crRNA that forms gRNA-t with a tracrRNA; the target nucleic acid comprises PS-t adjacent a PAM (P-t), wherein P-t is cognate to the first Cas for modifying PS-t.

[0748] A forty-third clause relates to the vector or combination of the forty-second clause, further in combination with said target nucleic acid.

[0749] A forty-fourth clause relates to the vector or combination of the forty-third clause, wherein said target nucleic acid is a chromosomal or episomal nucleic acid of a cell.

[0750] A forty-fifth clause relates to the vector or combination of the forty-fourth clause, wherein the cell is the cell is of a human or non-human animal pathogen species or strain (eg, *S aureus*).

[0751] A forty-sixth clause relates to the vector or combination of the forty-fouth or forty-fifth clause, wherein the cell is a cell of a human microbiome species, eg, a human gut microbiome species.

[0752] A forty-seventh clause relates to the vector or combination of any one of the forty-fourth to forty-sixth clauses, wherein PS-t is comprised by an essential gene, virulence gene or antibiotic resistance gene sequence of the cell (eg, a prokaryotic cell).

[0753] A forty-eighth clause relates to the vector or combination of the forty-seventh clause, wherein the gene is down-regulated or inactivated when first Cas modifies (eg, cuts) PS-t.

[0754] A forty-ninth clause relates to the vector or combination of the forty-seventh clause, wherein the gene is up-regulated or activated when first Cas modifies PS-t.

[0755] A fiftieth clause relates to the vector or combination of any one of the thirty-fifth to forty-ninth clauses in combination with said gene encoding the first Cas (eg, comprised by the first vector).

[0756] A fifty-first clause relates to the vector or combination of any one of the thirty-fifth to fiftieth clauses when inside a cell, wherein the cell comprises a CRISPR/Cas system comprising said gene encoding the first Cas.

[0757] A fifty-second clause relates to the vector or combination of any one of the thirty-fifth to fifty-first clauses for treating, preventing or diagnosing a disease or condition in a human or non-human animal, eg, for gene therapy of a human or animal, human or animal cell when the method is carried out in a human or animal cell; or for treating or preventing a bacterial infection in a human or animal when the method is carried out in a bacterial cell.

[0758] A fifty-third clause relates to a foodstuff, food ingredient or precursor ingredient, beverage, water (eg, intended for human consumption), an industrial or environmental substance(eg, oil, petroleum product, soil or a waterway or reservoir; or equipment for recovering or processing oil, petroleum product, soil, water, a foodstuff, foodstuff ingredient or precursor, or a beverage or beverage ingredient of precursor) comprising a first vector or combination according to any one of the thirty-fifth to fifty-second clauses . A fifty-fourth clause relates to an antibiotic (eg, anti-bacterial or anti-archaeal) composition a first vector or combination according to any one of the thirty-fifth to fifty-second clauses.

[0759] A fifty-fifth clause relates to a medicament for treating or preventing a disease or condition (eg, a bacterial infection or obesity) in a human or animal, the medicament comprising a first vector or combination according to any one of the thirty-fifth to fifty-second clauses.

[0760] In an example, the vector, combination, medicament or antibiotic is comprised by a medical device or medical container (eg, a syringe, inhaler or IV bag).

[0761] Any of the general features (see below) also may apply to the present configuration. Any configuration below is combinable with the present configuration, eg, to provide combinations of features for inclusion in one or more claims herein.

GENERALLY APPLICABLE FEATURES

[0762] The following features apply to any configuration (eg, in any of its aspects, embodiments, concepts, paragraphs or examples) of the invention:-

[0763] In an example, the target sequence is a chromosomal sequence, an endogenous host cell sequence, a wild-type host cell sequence, a non-viral chromosomal host cell sequence, not an exogenous sequence and/or a non-phage sequence (ie, one more or all of these), eg, the sequence is a a wild-type host chromosomal cell sequence such as as antibiotic resistance gene or essential gene sequence comprised by a host cell chromosome. In an example, the sequence is a host cell plasmid sequence, eg, an antibiotic resistance gene sequence.

[0764] In an example, at least two target sequences are modified by Cas, for example an antibiotic resistance gene and an essential gene. Multiple targeting in this way may be useful to reduce evolution of escape mutant host cells.

[0765] In an example, the Cas is a wild-type endogenous host cell Cas nuclease and/or each host cell is a wild-type host cell. Thus, in an embodiment the invention uses host cells without the need to de-repress endogenous Cas first to provide relevant Cas activity. In an example, each host cell has constitutive Cas nuclease activity, eg, constitutive wild-type Cas nuclease activity. In an example, the host cell is a bacterial cell; in an other example the host cell is an archael cell. Use of an endogenous Cas is advantageous as this enables space to be freed in vectors encoding HM- or PM-cRNA or gRNA. For example, Type II Cas9 nucleotide sequence is large and the use of endogenous Cas of the host cell instead is advantageous in that instance when a Type II CRISPR/Cas system is used for host cell modification in the present invention. The most commonly employed Cas9, measuring in at 4.2 kilobases (kb), comes from *S pyogenes.* While it is an efficient nuclease, the molecule's length pushes the limit of how much genetic material a vector can accommodate, creating a barrier to using CRISPR in the tissues of living animals and other settings described herein (see F.A. Ran et al., "In vivo genome editing using Staphylococcus aureus Cas9," Nature, doi:10.1038/nature14299, 2015). Thus, in an embodiment, the vector of the invention is a AAV vector or has an exogenous DNA insertion capacity no more than an AAV vector, and the Cas is an endogenous Cas of the host cell, wherein the cell is a bacterial or archaeal cell.

[0766] *S thermophilus* Cas9 (UniProtKB - G3ECR1 (CAS9_STRTR)) nucleotide sequence has a size of 1.4kb.

[0767] In an embodiment, therefore, the invention provides

[0768] A nucleic acid vector comprising more than 1.4kb or more than 4.2kb of exogenous DNA sequence encoding components of a CRISPR/Cas system, wherein the sequence comprises an engineered array or engineered sequence (optionally as described herein) for expressing one or more HM- or PM-crRNAs or gRNAs in host cells (any cell herein, eg, human, anial or bacterial or archael host cells), wherein the array or engineered sequence does not comprise a nucleotide sequence encoding a Cas nuclease that is cognate to the cRNA(s) or gRNA(s); optionally wherin at least 2, 3 or 4 cRNAs or gRNAs are encoded by the exogenous DNA. In an embodiment, the host cell is a bacterial or archael cell that expresses a Cas nuclease that is cognate to the crRNAs or gRNAs. In another example, such as for use with human or animal (eg, rodent, rat or mouse) cells the Cas nuclease is encoded by a different nucleic acid vector. In an example, wherein the cell is a human or animal cell, the vector is an AAV or lentiviral vector. In an example, the invention comprises a host cell comprising such a vector, wherein the host cell expresses said Cas. In an example, the host cell is a human or animal cell *ex vivo.*

[0769] The invention also provides

[0770] A nucleic acid vector comprising more than 1.4kb or more than 4.2kb of exogenous DNA sequence, wherein the exogenous DNA encodes one or more components of a CRISPR/Cas system and comprises an engineered array or sequence (eg, any such one described herein) for expressing one or more HM-crRNAs or gRNAs in host cells, wherein the exogenous sequence is devoid of a nucleotide sequence encoding a Cas nuclease that is cognate to the cRNA(s) or gRNA(s); optionally wherein at least 2 different cRNAs or gRNAs are encoded by the exogenous DNA. In an example, the invention comprises a host cell comprising such a vector, wherein the host cell expresses said Cas. In an example, the cRNAs or gRNAs are capable of hybridising in host cells to respective target protospacer sequences, wherein each protospacer sequence is comprised by an antibiotic resistance or essential host gene. This is exemplified by the worked examples herein where we show selective host cell growth inhibition by at least 10-fold in a mixed and non-mixed cell population. The mixture simulates a combination of species and strains found in human microbiota.

[0771] By "exogenous DNA sequence encoding components of a CRISPR/Cas system" is meant DNA that is inserted into a vector backbone, or such DNA in a progeny of a vector into which said insertion has previously taken place (eg, using recombinant DNA technology, such as recombineering). In an example, the exogenous DNA is 95, 90, 80, 85, 70 or 60% of the insertion capacity of the vector.

[0772] In an example, the vector is a viral vector. Viral vectors have a particularly limited capacity for exogenous DNA

insertion, thus virus packaging capacity needs to be considered. Room needs to be left for sequences encoding vital viral functions, such as for expressing coat proteins and polymerase. In an example, the vector is a phage vector or an AAV or lentiviral vector. Phage vectors are useful where the host is a bacterial cell.

[0773] The disclosure provides a combination product kit (eg, for treating or preventing a disease or condition in a human or animal subject as described herein), wherein the kit comprises an array, vector, system, cell, engineered cRNA or gRNA-encoding sequence or the cRNA or gRNA, which is in combination with an antibiotic (first antibiotic), wherein the cRNA or gRNA is capable of hybridising to a protospacer sequence comprised by a bacterial host cell antibiotic resistance gene wherein the antibiotic is said first antibiotic. The antibiotic can be be any antibiotic disclosed herein. In an embodiment, the antibiotic is combined in a formulation with the array, vector, system, cell, engineered cRNA or gRNA-encoding sequence or the cRNA or gRNA. In an example, the kit comprises the antibiotic in a container separate from a container comprising the array, vector, system, cell, engineered cRNA or gRNA-encoding sequence or the cRNA or gRNA.

[0774] In an embodiment, unless otherwise specified the or each cell is a bacterial cell, archaeal cell, algal cell, fungal cell, protozoan cell, invertebrate cell, vertebrate cell, fish cell, bird cell, mammal cell, companion animal cell, dog cell, cat cell, horse cell, mouse cell, rat cell, rabbit cell, eukaryotic cell, prokaryotic cell, human cell, animal cell, rodent cell, insect cell or plant cell. Additionally, in this case preferably the cells are of the same phylum, order, family or genus.

[0775] By use of the term "engineered" it will be readily apparent to the skilled addressee that the array, sequence, vector, MGE or any other configuration,concept, aspect, embodiment, paragraph or example etc of the invention is non-natrually occurring. For example, the MGE, vector or array comprises one or more sequences or components not naturally found together with other sequences or components of the MGE, vector or array. For example, the array is recombinant, artificial, synthetic or exogenous (ie, non-endogenous or not wild-type) to the or each host cell.

[0776] In an example, the array or vector of the invention is isolated, for example isolated from a host cell. In an example, the array or vector is not in combination with a Cas endonuclease-encoding sequence that is naturally found in a cell together with repeat sequences of the array.

[0777] In an example, the vector, MGE or array is not in combination with a Cas endonuclease-encoding sequence when not in a host cell. In an example, the vector or MGE does not comprise a Cas endonuclease-encoding sequence.

[0778] In an example, the target modification or cutting is carried out by a dsDNA Cas nuclease (eg, a Cas9, eg, a spCas9 or saCas9), whereby repair of the cut is by non-homologous end joining (NHEJ). This typically introduces mutation (indels) at the repair site, which is useful for inactivation of the target site (eg, phage gene or regulatory element, such as an essential gene or regulatory element thereof). In another example, the cutting is carried out by a ssDNA Cas nuclease (eg, a Cas9 nuclease) that cuts in a single strand (but does not do double stranded DNA cuts). This is useful for favouring HDR repair of the cut, which reduces the chances of indels. This may be useful where the target site (or gene or regulatory element comprising it) is desired, eg, where a HM- or PM-DNA is inserted at the target site for desired modification of the site. For example, in this case the modified gene produces a fusion protein comprising HM-DNA-encoded amino acid fused to host DNA-encoded sequence, or PM-DNA-encoded amino acid sequence fused to phage DNA-encoded sequence. The disclosure also provides a sequence encoding a fusion protein obtained or obtainable by such a method. In another example, the HM- or PM-DNA comprises a regulatory element (eg a promoter, enhancer, repressor or inducible switch for regulating gene expression), such that the fusion product comprises said DNA fused to host or phage gene or regulatory element DNA, thereby producing a fusion gene. The disclosure also provides a fusion gene obtained or obtainable by such a method. In an embodiment, the disclosure provides a vector (eg, a virus, virion, phage, phagemid, prophage or plasmid) comprising such a fusion gene, optionally wherein the vector is contained in a bacterial cell (eg, a prokaryotic, eukaryotic, bacterial, archaeal or yeast cell). In an example, the cell is *in vitro.*

[0779] In an example, the HM- or PM-DNA is vector DNA inside the cell. For example, the HM-or PM-DNA in the vector can be flanked by site-specific recombination sites (eg, frt or lox sites) which are cut by the action of a site-specific recombinase which is encoded by either a host cell or vector sequence. In another example, the vector comprises DNA that is transcribed into RNA copies of the HM- or PM-DNA and a reverse transcriptase (eg, encoded by the vector nucleic acid sequence) for producing HM- or PM-DNA from the RNA. This is useful for producing many copies of the desired HM- or PM-DNA to increase the chances of efficient and effective introduction at one or more of the target sites. In another embodiment, the HM- or PM-DNA is, or is encoded by, nucleic acid of a second vector (eg, a second phage or plasmid) that has transduced or transformed the host cell. For example, this could be a helper phage (which may also encode one or more coat proteins required for packaging of the first page vector). In another example, the DNA is provided in vector DNA and flanked by arms, wherein the 5' arm comprises a PAM that is recognised by a Cas nuclease when the vector is contained in the host cell and the 3' arm is flanked immediately downstream (3') by such a PAM, whereby in the host cell Cas cleavage liberates the HM- or PM-DNA with its flanked arms that can be designed to be homologous to host sequences flanking the cut in the host target sequence, whereby the HM-DNA is integrated into the host genome,or the PM-DNA is integrated into the phage genome. In one aspect, the disclosure provides a nucleic acid comprising such a HM- or PM-DNA and arms or a vector (eg, a phage or packaged phage) comprising such a nucleic acid, optionally wherein the vector is contained by a host cell. Optionally, the HM-DNA is in combination with a HM-array

as herein defined. Optionally, the PM-DNA is in combination with a PM-array as herein defined.

[0780] A particular application of the invention is the alteration of the proportion of *Bacteroidetes* (eg, *Bacteroides*) bacteria in a mixed bacterial population *ex-* or *in vivo.* As discussed above, this may be useful for environmental treatment such as treatment of waterways or drinking water infected with undesired *Bacteroidetes,* or for favouring useful commensal or symbiotic *Bacteroidetes* in humans or animals, eg, for producing bacterial cultures for administration to humans or animals for such purpose. In an example of the latter, the invention is useful for increasing the relative ratio of *Bacteroidetes versus Firmicutes,* which has been associated with lower of body mass and thus finds utility in treating or preventing obesity for medical or cosmetic purposes.

[0781] Studies suggest that *Bacteroides* have a role in preventing infection with *Clostridium difficile.* The development of the immune response that limits entry and proliferation of potential pathogens is profoundly dependent upon *B fragilis.* Also, Paneth cell proteins may produce antibacterial peptides in response to stimulation by *B thetaiotomicron,* and these molecules may prevent pathogens from colonizing the gut. In addition, *B thetaiotomicron* can induce Paneth cells to produce a bactericidal lectin, RegIII, which exerts its antimicrobial effect by binding to the peptidoglycan of gram-positive organisms. Thus, the use of the invention in any of its configurations for increasing the proportion of *Bacteroides* (eg, *thetaiotomicron* and/or *fragalis*) in a mixed population of gut bacteria is useful for limiting pathogenic bacterial colonisation of the population or a gut of a human or non-human animal.

[0782] Hooper *et al* demonstrated that *B thetaiotomicron* can modify intestinal fucosylation in a complex interaction mediated by a fucose repressor gene and a signaling system. Using transcriptional analysis it was demonstrated that *B thetaiotaomicron* can modulate expression of a variety of host genes, including those involved in nutrient absorption, mucosal barrier fortification, and production of angiogenic factors.

[0783] In an embodiment, the mixed population consists of the first and second bacteria (ie, and no further bacterial population).

[0784] In an example, the or each array is recombinant array in a vector and/or an isolated array in a vector. In an example, the array is contained in a host cell (eg, a microbial, bacterial or archaeal cell). In an example, said Cas is an endogenous Cas nuclease (eg Cas9) of the host cell. By harnessing the Cas of the host, this enables efficient use of host-type repeats in the array and possibility of using endogenous crRNA too - freeing up capacity which is otherwise limited in vectors, such as viruses or phage (noting that the Cas gene sequence such as Type II Cas9 is large).

[0785] In an example, the host CRISPR/Cas system is a Type I system. In an example, the host CRISPR/Cas system is a Type II system. In an example, the host CRISPR/Cas system is a Type III system.

[0786] The cas guided by the HM-crRNA or gRNA of the disclosure is a host endogenous Cas or a vector-encoded Cas compatible with the PAM in the target sequence.

[0787] Optionally, the host (or first and/or second bacteria) is a gram negative bacterium (eg, a spirilla or vibrio). Optionally, the host (or first and/or second bacteria) is a gram positive bacterium. Optionally, the host (or first and/or second bacteria) is a mycoplasma, chlamydiae, spirochete or mycobacterium. Optionally, the host (or first and/or second bacteria) is a *Streptococcus* (eg, *pyogenes* or *thermophilus)* host. Optionally, the host (or first and/or second bacteria) is a *Staphylococcus* (eg, *aureus,* eg, MRSA) host. Optionally, the host (or first and/or second bacteria) is an *E. coli (eg,* O157: H7) host, eg, wherein the Cas is encoded by the vecor or an endogenous host Cas nuclease activity is de-repressed. Optionally, the host (or first and/or second bacteria) is a *Pseudomonas* (eg, aeruginosa) host. Optionally, the host (or first and/or second bacteria) is a *Vibro* (eg, *cholerae* (eg, 0139) or *vulnificus*) host. Optionally, the host (or first and/or second bacteria) is a *Neisseria* (eg, *gonnorrhoeae* or *meningitidis)* host. Optionally, the host (or first and/or second bacteria) is a *Bordetello* (eg, *pertussis)* host. Optionally, the host (or first and/or second bacteria) is a *Haemophilus* (eg, *influenzae)* host. Optionally, the host (or first and/or second bacteria) is a *Shigella* (eg, *dysenteriae*) host. Optionally, the host (or first and/or second bacteria) is a *Brucella* (eg, *abortus)* host. Optionally, the host (or first and/or second bacteria) is a *Francisella* host. Optionally, the host (or first and/or second bacteria) is a *Xanthomonas* host. Optionally, the host (or first and/or second bacteria) is a *Agrobacterium* host. Optionally, the host (or first and/or second bacteria) is a *Erwinia* host. Optionally, the host (or first and/or second bacteria) is a *Legionella* (eg, *pneumophila*) host. Optionally, the host (or first and/or second bacteria) is a *Listeria* (eg, *monocytogenes)* host. Optionally, the host (or first and/or second bacteria) is a *Campylobacter* (eg, *jejuni*) host. Optionally, the host (or first and/or second bacteria) is a *Yersinia* (eg, *pestis)* host. Optionally, the host (or first and/or second bacteria) is a *Borelia*(eg, *burgdorferi*) host. Optionally, the host (or first and/or second bacteria) is a *Helicobacter* (eg, *pylori)* host. Optionally, the host (or first and/or second bacteria) is a *Clostridium* (eg, *dificile* or *botulinum)* host. Optionally, the host (or first and/or second bacteria) is a *Erlichia* (eg, *chaffeensis*) host. Optionally, the host (or first and/or second bacteria) is a *Salmonella* (eg, *typhi* or *enterica,* eg, serotype typhimurium, eg, DT 104) host. Optionally, the host (or first and/or second bacteria) is a *Chlamydia* (eg, *pneumoniae*) host. Optionally, the host (or first and/or second bacteria) is a *Parachlamydia* host. Optionally, the host (or first and/or second bacteria) is a *Corynebacterium* (eg, *amycolatum*) host. Optionally, the host (or first and/or second bacteria) is a *Klebsiella* (eg, *pneumoniae*) host. Optionally, the host (or first and/or second bacteria) is a *Enterococcus* (eg, *faecalis* or *faecim,* eg, linezolid-resistant) host. Optionally, the host (or first and/or second bacteria) is a *Acinetobacter* (eg, *baumannii,* eg, multiple drug resistant) host.

[0788] In an example, the cell is a prokaryotic cell. In an example, the cell is a bacterial cell. In an example, the cell is a archaeal cell. In an example, the cell is a microbe cell. In an example, the cell is a protozoan cell. In an example, the cell is a fish cell. In an example, the cell is a bird cell. In an example, the cell is a reptilian cell. In an example, the cell is an arachnid cell. In an example, the cell is a yeast cell (eg, a *Saccharomyces* cell). In an example, the host cell is a plant cell. In an example, the host cell is an animal cell (eg, not a human cell, eg, not a rodent cell). In an example, the host cell is a human cell (eg, not a cell in an embryo or in a human), for example a host cell *in vitro.* In an example, the cell is a livestock or companion pet animal cell (eg, a cow, pig, goat, sheep, horse, dog, cat or rabbit cell). In an example, the host cell is an insect cell (an insect at any stage of its lifecycle, eg, egg, larva or pupa). In an example, the host cell is a protozoan cell. In an example, the cell is a cephalopod cell.

[0789] Optionally the array, system, engineered nucleotide sequence or vector nucleic acid further comprises a (eg, one, tow or more) nuclear localisation signal (NLS), eg, for targeting to the nucleus when the host cell is a eukaryotic cell, eg, a plant or animal. In an example, a NLS flanks each end of a Cas-encoding nucleic acid sequence of the invention and/or an array of the invention - particularly for use in targeting in a eukaryotic host cell.

[0790] A tracrRNA sequence may be omitted from a array or vector of the invention, for example for Cas systems of a Type that does not use tracrRNA.

[0791] In an example, the Cas guided to the target is an exonuclease. Optionally a nickase as mentioned herein is a doube nickase.

[0792] An example of a nickase is a Cas9 nickase, ie, a Cas9 that has one of the two nuclease domains inactivated - either the RuvC and/or HNH domain.

[0793] Optionally the host system is a Type I system (and optionally the array, HM-crRNA or gRNA is of a different CRISPR system, eg, Type II or III). Optionally the array or engineered sequence is in combination in a virus or plasmid with a nucleotide sequence encoding a Cas of the same system as the array, HM-crRNA or gRNA, eg, where the Cas does not operate or operate efficiently with the host system. Optionally the host system is a Type II system (and optionally the array, HM-crRNA or gRNA is of a different CRISPR system, eg, Type I or III). Optionally the array or engineered sequence is in combination in a virus or plasmid with a nucleotide sequence encoding a Cas of the same system as the array, HM-crRNA or gRNA, eg, where the Cas does not operate or operate efficiently with the host system. Optionally the host system is a Type III system (and optionally the array, HM-crRNA or gRNA is of a different CRISPR system, eg, Type I or II). Optionally the array of engineered sequence is in combination in a virus or plasmid with a nucleotide sequence encoding a Cas of the same system as the array, eg, where the Cas does not operate or operate efficiently with the host system.

[0794] Mention herein of using vector DNA can also in an alternative embodiment apply *mutatis mutandis* to vector RNA where the context allows. For example, where the vector is an RNA vector. All features of the invention are therefore in the alternative disclosed and to be read as "RNA" instead of "DNA" when referring to vector DNA herein when the context allows. In an example, the or each vector also encodes a reverse transcriptase.

[0795] In an example, the or each array or engineered nucleotide sequence is provided by a nanoparticle vector or in liposomes.

[0796] In an example, the Cas is a Cas nuclease for cutting, dead Cas (dCas) for interrupting or a dCas conjugated to a transcription activator for activating the target.

[0797] In an example, the host CRISPR/Cas system comprises a host CRISPR array and a cognate host Cas for nucleotide sequence targeting in the host. In an example, the host target sequence comprises at lest 5, 6, 7, 8, 9, 10, 20, 30 or 40 contiguous nucleotides. In an example, the target sequence is cut by Cas, eg, a Cas9. In an embodiment, the sequence is not in a spacer.

[0798] In an example, the or each array or engineered sequence comprises an exogenous promoter functional for transcription of the crRNA or gRNA in the host.

[0799] In an example, the or each array repeats are identical to repeats in the host array, wherein the CRISPR array does not comprise a PAM recognised by a Cas (eg, a Cas nuclease, eg, Cas9) of the host CRISPR/Cas system. This applies *mutatis mutandis* to repeat sequence of the HM-crRNA and gRNA. This embodiment is advantageous since it simply enables the CRISPR array to use the endogenous host Cas to target the host target sequence. This then is efficient as the array is tailored for use by the host machinery, and thus aids functioning in the host cell. Additionally, or alternatively (eg where the array is provided in combination with an exogenous (non-host endogenous) Cas-encoding sequence) this embodiment enables the CRISPR array to use the endogenously-encoded tracrRNA, since the CRISPR array repeats will hybridise to the endogenous tracrRNA for the production of pre-crRNA and processing into mature crRNA that hybridises with the host target sequence. The latter complex can then guide the endogeous Cas nuclease (eg, Cas9) or guide Cas produced from the sequence comprised by the CRISPR array. This embodiment therefore provides the flexibility of simply constructing a vector (eg, packaged virus or phage) containing the CRISPR array but not comprising a tracrRNA- and/or Cas nuclease-encoding sequence. This is more straightforward for vector construction and also it frees up valuable space in the vector (eg, virus or phage) which is useful bearing in mind the capacity limitation for vectors, particularly viral vectors (eg, phage). The additional space can be useful, for example, to enable inclusion

of many more spacers in the array, eg, to target the host genome for modification, such as to inactivate host genes or bring in desired non-host sequences for expression in the host. Additionally or alternatively, the space can be used to include a plurality of CRISPR arrays in the vector. These could, for example, be an arrangement where a first array is of a first CRISPR/Cas type (eg, Type II or Type II-A) and the second array could be of a second type (eg, Type I or III or Type II-B). Additionally or alternatively, the arrays could use different Cas nucleases in the host (eg, one array is operable with the host Cas nuclease and the second array is operable with an exogenous Cas nuclease (ie, a vector-encoded nuclease)). These aspects provide machinery for targeting in the host once the vector has been introduced, which is beneficial for reducing host resistance to the vector, as the host would then need to target a greater range of elements. For example, if the host were able to acquire a new spacer based on the first CRISPR array sequence, the second CRISPR array could still function in the host to target a respective target sequence in the host cell. Thus, this embodiment is useful to reduce host adaptation to the vector.

[0800] Another benefit is that it is possible (for example, with this arrangement) to include in the CRISPR array (or distributed over a plurality of such arrays in the vector) multiple copies of the same spacer (eg, a spacer used to target a target site in the host cell). This is beneficial since it has been proposed that adaptation of hosts, such as bacteria and archaea, may involve loss of spacers from their arrays where the spacers target beneficial host DNA (PLoS Genet. 2013;9(9):e1003844. doi: 10.1371/journal.pgen.1003844. Epub 2013 Sep 26, "Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids", Jiang W *et al*). It is thought that the removal of spacer-repeat units occurs through recombination of repeat sequences. Thus, according to the present aspect of the invention, there is provided one, two, three, four, five, six or more CRISPR arrays or engineered sequences of the invention comprising a plurality (eg, 2, 3, 4 5, 10,15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100 or more) copies of a spacer for hybridising to a host target sequence. This reduces the chances of all of these spacers being lost by recombination in the host cell. In a further application of this aspect, the CRISPR arrays comprise a first array comprising one or more (eg, 2, 3, 4 5, 10,15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90 or more) of the spacer copies and a second array comprising one or more (eg, 2, 3, 4 5, 10,15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90 or more) of the identical spacer copies, wherein spacer copies in the first array are each flanked by first repeats and the identical spacer copies in the second array are each flanked by second repeats, wherein the first repeats are different from the second repeats. This has the benefit that at least one of the first and second repeats can be selected not to be recognised by a host Cas nuclease (or the same host Cas nuclease), to reduce the chances of host adaptation involving more than one of the arrays. In an example, the first array is in combination with a Cas nuclease sequence that is not encoded by the host cell and which is a cognate Cas for the first repeats. Optionally, also the second array is in combination with a Cas nuclease sequence (eg, the same or different from that for the first array) that is not encoded by the host cell and which is a cognate Cas for the second repeats.

[0801] An embodiment provides a first array contained in a first vector and a second array contained in a second vector which does not contain the first array (eg, wherein the vectors are plasmids or virions (eg, of the same virus type) or packaged phage (eg, of the same phage type). This is useful since the vectors can be simultaneously or sequentially introduced into the same host cell. Thus, when the host gains resistance to the first array, the second is introduced to provide a second array with which the resistant host (eg, bacterium or archaeon) has not previously co-evolved, thereby providing a second modification (eg, knock-down) wave against the host cell. This also provides flexibility since a third such vector, comprising a spacer or array that is different from the first and second arrays and spacers, can be introduced into the host cell simultaneously or sequentially with the second vector to provide a further route to host cell modification that has not previously been present during evolution of the hosts that are resistant to a spacer in the first array. Instead of arrays, engineered nucleotide sequences of the disclosure can be used.

[0802] Thus, in one embodiment, the invention provides a composition for modifying a host cell, wherein the composition provides any array or engineered sequence as described herein. Thus, in one embodiment, the invention provides a composition for modifying a host cell, wherein the composition provides a first array as described herein in a first vector (eg, virion or packaged phage) and a second first such array as described herein in a second vector (eg, virion or packaged phage respectively), wherein the second array comprises one or more spacers that target one or more host target sequences and which is/are not comprised by the first array. Instead of arrays, engineered nucleotide sequences of the disclosure can be used.

[0803] In an embodiment the array or engineered sequence is contained in a virophage vector and the host is alternatively a virus which can infect a cell. For example, the host is a large virus that may have infected an amoeba cell. For example, the host is a Sputnik virus, Pithovirus, mimivirus, mamavirus, Megavirus or Pandoravirus, eg, wherein the host virus is in water. In an example of this embodiment, the invention is for water or sewage treatment (eg, purification, eg, waterway, river, lake, pond or sea treatment).

[0804] In an embodiment the or each vector or engineered sequence is or is comprised by a ΦNM1 phage, eg, wherein the host cell(s) is a *S. aureus* (eg, MRSA) cell.

[0805] The general features also provide the following clauses:-

[0806] In a first clause, there is provided an antimicrobial composition (eg, an antibiotic, eg, a medicine, disinfectant

or mouthwash), comprising an array, engineered sequence, virus, virion, phage, phagemid, prophage, population or collection according to any aspect of the disclosure.

[0807] A second clause relates to the composition of the first clause for medical or dental or opthalmic use (eg, for treating or preventing an infection in an organism or limiting spread of the infection in an organism.

[0808] In an example, the organism is a plant or animal, eg, vertebrate (eg, any mammal or human disclosed herein) or crop or food plant.

[0809] A third clause relates to a composition comprising an array, engineered sequence, system, collection, virus, virion, phage, phagemid, prophage, composition, population or collection according to the disclosure for cosmetic use (eg, use in a cosmetic product, eg, make-up), or for hygiene use (eg, use in a hygiene product, eg, soap).

[0810] A fourth clause relates to use of a composition comprising an array, engineered sequence, collection, virus, virion, phage, phagemid, prophage, population or collection according to any one of the first to thirty-fifth clauses, in medicine or for dental therapeutic or prophylacticc use.

[0811] A fifth clause relates to use of a composition comprising an array, engineered sequence, collection, system, virus, virion, phage, phagemid, prophage, composition, population or collection according to the disclosure, in cosmetic use (eg, use in a cosmetic product, eg, make-up), or for hygiene use (eg, use in a hygiene product, eg, a soap).

[0812] A sixth clause relates to use of an array, engineered sequence, system, collection, virus, virion, phage, phagemid, prophage, composition, population or collection according to the disclosure in a host modifying (HM) CRISPR/Cas9 system (eg, Type I, II or III) that is capable of modifying a target nucleotide sequence of a host cell, wherein the array, engineered sequence, system, virus, virion, phage, phagemid, prophage, population or collection is according to the present disclosure.

[0813] A seventh clause relates to the use of the fourth, fifth or sixth clause, wherein the array, engineered sequence, system, collection, virus, virion, phage, phagemid, prophage, population or collection is not in a host cell.

[0814] An eighth clause relates to the use of the fifth or sixth clause, wherein the array, engineered sequence, collection, system, virus, virion, phage, phagemid, prophage, population or collection is in a host cell (eg, a microbe, bacterium or archaeon cell).

[0815] A ninth clause relates to the use of any one of the fourth to sixth clauses for modifying a microbial cell (eg, for killing or reducing growth of the cell or a culture of microbe cells).

[0816] A tenth clause relates to a method of modifying a target nucleotide sequence in a host cell (eg a microbe bacterium or archaeon), the method comprising transforming the host cell with the array, engineered sequence, system, collection, virus, virion, phage, phagemid, population or collection according to the present disclosure, whereby the target nucleotide sequence is Cas modified, wherein the host target sequence is a nucleotide sequence of a host CRISPR/Cas system of the cell.

[0817] An eleventh clause relates to a method of reducing the development of host cell resistance to transformation by a nucleic acid vector or maintenance of a nucleic acid vector in the host cell, wherein the host cell comprises a target nucleotide sequence, the method comprising transforming the host cell with the array, engineered sequence, collection, system, virus, virion, phage, phagemid, population or collection according to the disclosure, whereby the target nucleotide sequence is Cas modified (eg, cut, mutated or knocked-down).

[0818] A twelfth clause relates to the method of the eleventh clause, wherein the vector is a virus that is capable of infecting the host cell and the transforming step comprises infecting the host cell with the vector.

[0819] A thirteenth clause relates to the method of the eleventh ro twelfth clause, wherein the host cell is a bacterial or archaeal cell and the vector is a phage or phagemid.

[0820] A fourteenth clause relates to the method of any one of the eleventh to thirteenth clauses, wherein the host target sequence is essential to host CRISPR/Cas-mediated acquisition of vector sequence spacers. A fifteenth clause relates to the array, engineered sequence, system, vector, cell, collection, composition, use or method of any preceding clause, wherein at least component (ii) is contained in a virus (eg, a phage) that is capable of expressing an endolysin for host cell lysis, optionally wherein the endolysin is a phage phi11, phage Twort, phage P68, phage phiWMY or phage K endolysin (eg, MV-L endolysin or P-27/HP endolysin).

[0821] A sixteenth clause relates to the array, engineered sequence, system, vector, collection, cell, composition, use or method of the fifteenth clause in combination with an endolysin for host cell lysis, eg, in combination with MV-L endolysin or P-27/HP endolysin or a functional homologue thereof.

[0822] A seventeenth clause relates to the array, engineered sequence, system, vector, collection, cell, composition, use or method of any preceding clause in combination with an antimicrobial, eg, antibiotic agent, eg, a beta-lactam antibiotic.

[0823] An eighteenth clause relates to the array, engineered sequence, system, vector, collection, cell, composition, use or method of any preceding clause, wherein the host cell is a *Staphylococcus, Streptococcus, Pseudomonas, Salmonella, Listeria, E coli, Desulfovibrio* or *Clostridium* host cell.

[0824] A nineteenth clause relates to the array, engineered sequence, system, vector, collection, cell, composition, use or method of any preceding clause, wherein the host cell is a *Staphylococcus* (eg, *S aureus*) host cell and at least

component (ii) is contained in a Class I, II or III *Staphylococcus* phage (eg, a packaged phage), optionally a *Caudovirales* or *Myoviridae* phage.

[0825] A twentieth clause relates to the array, engineered sequence, system, vector, cell, collection, composition, use or method of any preceding clause, wherein the host cell is a beta-lactam antibiotic-resistant *Streptococcus aureus,* methicillin-resistant *Streptococcus aureus* (MRSA), vancomycin-resistant *Streptococcus aureus* or teicoplanin-resistant *Streptococcus aureus* and optionally the target sequence is a sequence of a host beta-lactam antibiotic-resistance gene, methicillin-resistance gene, vancomycin-resistance gene or teicoplanin-resistance gene respectively.

[0826] Suitable methods for producing and testing phage vectors of the invention are, for example, general methods disclosed in WO2014/124226.

Mobile Genetic Elements, Transposons & Carriers (for any configuration)

[0827] Plasmids are very common in *Bacteroides* species and are found in 20 to 50% of strains. Many plasmids possess oriT and a transacting mobilisation gene, which allow them to be transferred by conjugation. Thus, in an example, the vector is a plasmid comprising oriT and/or a mobilisation gene, eg, wherein the first or second bacteria are *Bacteroides.* In an example, the engineered sequence is comprised by such a vector.

[0828] In an example, the host cells, or the first or second bacteria naturally comprise transposons. Transposons, both mobilisable and conjugative, do not replicate independently; rather, they excise from and integrate into chromosomal DNA and are copied along with the chromosomal DNA. Conjugative transposons have a mechanism of excision and integration that resemble some features of both plasmids and bacteriophage. Conjugative transposons are practically ubiquitous among the *Bacteroides:* over 80% of *Bacteroides* strains contain at least one conjugative transposon. The conjugative transposons of *Bacteroides* belong to at least two families; CTnDot is the best described . Often, the name of the strain in which they are found is added to the designation (e.g., CTnDot, found in the DOT strain of *B thetaiotaomicron*). In addition to being able to insert into the chromosome, *Bacteroides* conjugative transposons can insert into coresident plasmids and mobilise them in *cis* (i.e., they can act on entities that are physically adjacent) by integrating themselves into the plasmid and facilitating transfer of the plasmid-conjugative transposon hybrid into another cell. They can also mobilise coresident plasmids "in *trans"* by supplying factors needed to facilitate transfer of the plasmid, while remaining physically separate from the plasmid.

[0829] Conjugative transposons do not exclude each other as do plasmids, so a strain can accumulate more than one conjugative transposon. Furthermore, there is some evidence that the presence of more than one copy of the conjugative transposon in the strain results in a stimulation of transposition (transactivation). Theoretically, this suggests that as more conjugative transposons accumulate in the environment, the transfer of the transposon genes to other bacteria will also increase, and there will be a significant upward spiraling of distribution of the genes. Many of the *Bacteroides* transposons carry the tetQ gene and thus confer tetracycline resistance. Further, self-transfer and other activities are significantly stimulated by low levels of tetracycline, regulated by the tetQ-rteA-rteB operon. Tetracycline increases transcription of rteA and -B, which code for the sensor and activator components of a two-component regulatory system. In turn, RteB activates expression of rteC, which is necessary for self-transfer.

[0830] In an example, the vector (eg, the vector comprising the engineered sequence) comprises a transposon, wherein the transposon comprises the engineered sequence, HM- or PM-array of the invention, wherein the transposon is operable in the host cell(s) or in the first or second bacteria host cell species. In an embodiment, the transposon is a *Bactroides* transposon (eg, a CTnDot transposon, eg, a *B thetaiotaomicron* or *B fragalis* CTnDot tranposon) and thehost cells, or the first or second bacteria are or comprise *Bacteroides* (eg, of the same species as said CTnDot transposon). In an example, the transposon is a conjugative transposon. In an example, the transposon is a mobilisable transposon. In an example, the transposon is tranferrable between *Bacteroides* cells. In an example, the transposon comprises an *intDOT* sequence. In an example, the transposon comprises an *ori*T. In an example, the transpsoson encodes one or more mating pore proteins for conjugative transfer of the transposon between host cells.

[0831] In an example, the disclosure provides a transposon that comprises a *Bacteroides tetQ* gene. In an example, the transposon further comprises a *Bacteroides* tetQ-rteA-rteB operon. In an example, the first or second bacteria are *Bacteroides.* In an example, the transposon is a *Bacteroides* CTnDot transposon that encodes one or more mating pore proteins for conjugative transfer of the transposon between host cells and comprises one or more arrays or engineered sequences of the invention, an *ori*T, an *intDOT* sequence and a tetQ-rteA-rteB operon, and is optionally for administration or is administered to said human or non-human animal as mentioned herein in combination with tetracycline. Transfer of most *Bacteroides* CTns is stimulated by tetracycline. The transposon is operable with an integrase in a host cell or is operable with an exogenous integrase carried on the same or a different vector to the transposon. In an embodiment, the vector is a phage (eg, a packaged phage) comprising the transposon and a nucleotide sequence encoding a cognate integrase. The phage is capable of infecting a host bacterial cell for replication and excision of the transposon, eg, for conjugative transfer to neighbouring host cells in a mixed bacterial population (eg, a gut microbiota population).

[0832] In an embodiment, the transposon is comprised by a vector that carries one or more gene sequences necessary

for transposon transfer between host cells, wherein said gene sequences are outside of the transposon on the vector nucleic acid. For example, the vector is a packaged phage that is capable of infecting a host cell (eg, a *Bacteroides* host cell), wherein the phage nucleic acid comprises a said transposon comprising a array of the invention and upstream or downstream of the transposon one or more genes operable for conjugative tranfer of the transposon (eg, one or more genes encoding relaxases, coupling proteins and/or mating bridge proteins for transposon conjugative transfer; and/or one or both of *mob* and *tra* operons), wherein one, more or all these genes is not comprised by the transposon. In an example, these genes are genes for excision of the transposon from chromosomal DNA inside a first host cell. Thus, the transposon is able to mobilise inside that cell and carries with it genes necessary for the subsequent conjugative transfer into a second host cell. By providing some of the transposon genes in this way on the vector outside the transposon, this frees up room in the transposon for inclusion of engineered sequence or array DNA of the invention so that this can be accommodated and carried by mobilised transposons. The invention provides such a vector comprising one or more such transposons for use in the use of the invention or for use in the metjod or system of the disclosure or generally for introduction into bacterial cells (in this case instead of targting a phage sequence, the array included in the transposon can target a bacterial target sequence to modify the sequence, eg, cut it using Cas in a cell harbouring the transposon).

[0833] Molecular mechanisms of CTnDot excision and integration more closely resemble that of bacteriophage rather than transposition. The CTnDOT integrase and excision proteins themselves are quite similar to those from bacteriophage. Thus, in one embodiment the function of one or more integrase and/or excision proteins of the transposon of the disclosure are provided by the phage integrase and/or excision proteins respectively, and the transposon does not comprise corresponding gene(s) encoding such integrase or excision proteins whose functions are provided by phage proteins.

[0834] In an example, the transpsoson comprises *rteC* and the operon *xis2c-xis2d-orf3-exc.* Optionally, additionally the vector comprises *mob* and *tra* operons outside of the tranpsoson (eg, upstream or downstream of the transposon in the vector nucleic acid). Thus, this frees up space in the transposon for providing CRISPR array sequence or engineered sequence of the invention.

[0835] Many conjugative transposons are able to mobilise other elements. For example, many coresident plasmids are mobilized by a conjugative transposon in *trans.* This occurs when a plasmid containing an *ori*T utilizes the CTn-provided mating pore proteins for transfer to a recipient cell. The *Bacteroides* CTns have also been shown to mobilize elements when in *cis,* a feature that is not typical for CTns. For example, if CTnDOT excises from the chromosome and integrates on a plasmid, it can provide the mating pore, an *ori*T, and the mobilization (relaxase/coupling) proteins, allowing it to transfer the entire plasmid by acting "in cis." This ability to use both trans and cis mechanisms of mobilization is unusual and suggests that the *Bacteroides* CTns have a greater capacity to mobilize other elements.

[0836] In an example, the vector of the invention is a plasmid comprisng one or more engineered sequences or arrays of the invention and an *ori*T that is cognate to a host cell species CTnDot transposon that encodes mating pore proteins, whereby the plasmid is mobilisable in a host cell comprising a said CTnDot transposon. Thus, the plasmid is capable of horizontal transfer between host cells, thereby spreading arrays of the invention in a population of such host cells (eg, *Bacteroides* cells). In an example, the invention provides a composition comprising a population of carrier bacteria, wherein the carrier bacteria are compatible with such a plasmid vector of the invention, whereby the vector is capable of horizontal transfer to recipient host bacteria cells (eg, *Bacteroides* or *Firmicutes,* eg, *Streptococcus* cells) comprising cognate CTnDot transposons when the carrier and recipient bacteria are mixed. In an example, the carrier bacteria are comprised by a beverage (eg, probiotic drink, such as one described herein) or foodstuff for human or non-human animal consumption, whereby the carrier bacteria can mix with recipient bacteria harboured by the human or animal (eg, in the oral cavity or in the gut). Other transposons within the CTnDOT-like family include CTnERL and CTn341, although these elements differ from CTnDOT, and thus instead of a CTnDot transposon, the transposon of the general aspect of the disclosure can be a CTnERL or CTn341 transposon carrying one or more desired CRISPR arrays or engineered sequences for targeting one or more bacterial or phage nucleotide target sites when the transposon is comprised by a bacterial or archaeal host cell.

[0837] In order for transfer of the conjugative transposon to occur, there are three main steps that take place. The first step is excision from the chromosome to form a covalently closed circular intermediate. Second, a single-stranded copy is then transferred through the mating pore to a recipient cell, after which the copy becomes double stranded. Third, the intact double-stranded CTn integrates into the chromosome of the recipient. Conjugative transposition is replicative, as a copy of the CTn is retained in the donor cell. Because the element resides within the chromosome, it is also transferred vertically to progeny cells. This is important because when desired CRISPR arrays or engineered sequences (and optionally Cas sequence) are present on CTns, they are not only transferred readily within the population, but they are also very stably maintained from generation to generation. This is as seen, for example, with retained antibotic resistance determiniants. Further, it is believed that *Bacteroides* may serve as a reservoir of antibiotic resistance determinants which disseminates these genes to other organisms outside the *Bacteroides* genus, possibly even transferring these elements to organisms that are transiently passing through the gut. Similarly, a reservoir of arrays or engineered se-

quences of the invention can be created using vectors of the invention that are administered to a human or non-human animal, eg, for treating or preventing obesity, diabetes or IBD or any other disease or condition disclosed herein.

[0838] In an example, one can exploit the reservoir of desired CRISPR arrays or engineered sequences by using one or more arrays or sequences comprised by a transposon (eg, a CTnDot) that is capable of being harboured by *Bacteroides* cells (eg, in the gut or oral cavity of a human or non-human animal), wherein the array(s)/sequence(s) do not target a sequence of the host *Bacteroides* cell, but do target a nucleotide sequence comprised by a gut microbiota cell (eg, bacterial cell) of a different species (eg, a *Firmicutes* cell or pathogenic bacterial cell, eg, *Streptococcus, C dificile, H pylori, Salmonella, Listeria, Yersinia, Shigella* or *Campylobacter* cell). Thus, in this way transfer of the arrays or sequences of the invention to neighbouring recipient pathogenic or undesired bacteria can take place, and once inside the recipient cell the array(s) of the invention are operable to guide Cas to the respective target site in the host cell to modify (eg, cut) the site. In this case, the array/sequence can comprise repeat sequences that are found in the recipient cell of interest so that the array/sequence can operate with an endogenous CRISPR/Cas system inside the recipient cell. This avoids the need to include Cas and/or tracrRNA-encoding sequences in the vector, engineered sequence or transposon of the disclosure, thereby freeing up space and simplifying construction. Increased space is useful for enabling inclusion of more spacers to target more target sites in the recipient cell. In an alternative, the transposon array(s) or sequence(s) comprises a Type II Cas9-encoding sequence and cognate repeat sequences. For example, the Cas9 (any Cas9 mentioned herein) is a *S pyogenes, S thermophilus* or *S aureus* Cas9 and may optionally be a nickase or dCas9 ("dead Cas9"). As *Bacteroides* are obligate anaerobes (or have a strong preference for anaerobic environments) and typically are pathogenic outside the gut environment, it may not be desirable to use *Bacteroides* cells as carriers for the vectors of the invention or transposons of the disclosure, eg, when administering to the gut or oral cavity of a human or animal. To address this, the invention provides a carrier population of bacteria harbouring vectors, engineered sequence(s) or transposons of the disclosure, wherein the carrier bacteria are compatible with such a vector, sequence or transposon, whereby the vector, sequence or transposon is capable of horizontal transfer to recipient host bacteria cells (eg, *Bacteroides*) in gut microbiota when the carrier and recipient bacteria are mixed. In an example, the carrier bacteria are comprised by a beverage (eg, probiotic drink, such as one described herein) or foodstuff for human or non-human animal consumption, whereby the carrier bacteria can mix with recipient bacteria harboured by the human or animal (eg, in the oral cavity or in the gut). In an embodiment, the vectors, sequences or transposons comprise CRISPR arrays of the invention, wherein the arrays target nucleotide sequences comprised by the recipient cells to modify the target sequences, eg, by cutting the sequences to inactivate genes comprising the target sequences. In an alternative, the vectors, sequences or transposons are capable of horizontal transfer (eg, conjugative transposon transfer) to a second recipient population of bacteria, which are of a different species to the first recipient bacteria, wherein the nucleotide sequence target sites are comprised by the second recipient bacteria but not comprised by the first recipient bacteria, whereby the target sites are modified by Cas in the second recipient bacteria (host cells).

[0839] In an example, the first recipient bacteria are *Bacteroides* bacteria and the second recipient bacteria are *Firmicutes* or pathogenic bacteria, eg, gut bacteria. In an example, the carrier bacteria comprise vectors of the invention (phage or plasmids) comprising one or more conjugative transposons (eg, CTnDot trasposons) that are capable of being harboured by the carrier bacteria, first bacteria and second bacteria, eg, wherein the transposons comprise *ori*T and the carrier bacteria, first bacteria and second bacteria are compatible with *ori*T.

[0840] In an alternative, the carrier bacteria are capable of transferring the vector, engineered sequence or transposon of the disclosure directly to *Firmicutes* or pathogenic bacteria, eg, in an animal or non-human animal, eg, in the gut, oral cavity or systemically (eg, in the blood). In an example, the pathogenic bacteria are *C dificile, H pylori,* pathogenic *E coli, Salmonella, Listeria, Yersinia, Shigella, S aureus, Streptococcus* or *Campylobacter* bacteria.

[0841] In an example, the carrier bacteria are bacteria of one or more species selected from the group consisting of a *Lactobacillus* species (eg, *acidophilus* (eg, La-5, La-14 or NCFM), *brevis, bulgaricus, plantarum, rhammosus, fermentum, caucasicus, helveticus, lactis, reuteri* or *casei eg, casei* Shirota), a *Bifidobacterium* species (eg, *bifidum, breve, longum* or *infantis), Streptococcus thermophilus* and *Enterococcus faecium.* For example, the bacteria are *L acidophilus* bacteria.

[0842] Mobilisable transposons, like mobilisable plasmids, cannot self-transfer but can transfer between cells in the presence of the TcR helper element. The most commonly discussed *Bacteroides* transposons of this class include Tn4399,Tn4555, and the nonreplicating *Bacteroides* units. The mobilisable transposon Tn4555, for example, was first detected during studies of transmissible cefoxitin resistance in a clinical isolate of *Bacteroides vulgatus.* In an embodiment, therefore, the transpsoson of the disclosure is a mobilisable transposon (eg, a *Bacteroides* mobilisable transposon), eg, a Tn4399 or Tn4555 comprising one or more arrays or sequences of the invention. The transposon is in combination with a TcR helper element.

[0843] In an example, the transposon of the disclosure is *Enterococcus* Tn916 or Gram-positive Tn1546 transposon. A transposon (eg, as a CTnDot, Tn4399 or Tn4555 transposon) can be characterised for example according to its terminal repeats and/or transposase- or resolvase-encoding sequence(s). In an alternative example, the vector or transposon comprises an origin of replication selected from pMB1, pBR322, ColE1, R6K (in combination with a *pir* gene),

p15A, pSC101, F1 and pUC. In an example, the transposon is in combination with a factor (eg, an antibiotic, eg, tetracycline) that is required for transposon mobilisation or transfer. In an example, the transposon comprises an antibiotic resistance gene (eg, tetracycline resistance gene) and the transposon is in combination with said antibotic (eg, administered simultaneously or sequentially to the human with said antibiotic). In an example, the transposon is a piggyBac, Mariner or Sleeping Beauty transposon in combination with a cognate tranpsosase. In an example, the transposon is a Class I transposon. In an example, the transposon is a Class II transposon. In an example, the transposon is a Tn family transposon.

TARGETING ANTIBOTIC RESISTANCE IN BACTERIAL HOSTS

[0844] Antibiotic resistance is a worldwide problem. New forms of antibiotic resistance can cross international boundaries and spread between continents with ease. Many forms of resistance spread with remarkable speed. World health leaders have described antibioticresistant microorganisms as "nightmare bacteria" that "pose a catastrophic threat" to people in every country in the world. Each year in the United States, at least 2 million people acquire serious infections with bacteria that are resistant to one or more of the antibiotics designed to treat those infections. At least 23,000 people die each year as a direct result of these antibiotic-resistant infections. Many more die from other conditions that were complicated by an antibioticresistant infection. In addition, almost 250,000 people each year require hospital care for Clostridium difficile (C. difficile) infections. In most of these infections, the use of antibiotics was a major contributing factor leading to the illness. At least 14,000 people die each year in the United States from C. difficile infections. Many of these infections could have been prevented. Antibiotic-resistant infections add considerable and avoidable costs to the already overburdened U.S. and other healthcare systems. In most cases, antibiotic-resistant infections require prolonged and/or costlier treatments, extend hospital stays, necessitate additional doctor visits and healthcare use, and result in greater disability and death compared with infections that are easily treatable with antibiotics. The total economic cost of antibiotic resistance to the U.S. economy has been difficult to calculate. Estimates vary but have ranged as high as $20 billion in excess direct healthcare costs, with additional costs to society for lost productivity as high as $35 billion a year (2008 dollars). The use of antibiotics is the single most important factor leading to antibiotic resistance around the world. Antibiotics are among the most commonly prescribed drugs used in human medicine. However, up to 50% of all the antibiotics prescribed for people are not needed or are not optimally effective as prescribed. Antibiotics are also commonly used in food animals to prevent, control, and treat disease, and to promote the growth of food-producing animals. The use of antibiotics for promoting growth is not necessary, and the practice should be phased out. Recent guidance from the U.S. Food and Drug Administration (FDA) describes a pathway toward this goal. It is difficult to directly compare the amount of drugs used in food animals with the amount used in humans, but there is evidence that more antibiotics are used in food production.

[0845] The other major factor in the growth of antibiotic resistance is spread of the resistant strains of bacteria from person to person, or from the non-human sources in the environment, including food. There are four core actions that will help fight these deadly infections: 1. preventing infections and preventing the spread of resistance; 2. tracking resistant bacteria; 3. improving the use of today's antibiotics; and 4. promoting the development of new antibiotics and developing new diagnostic tests for resistant bacteria. Bacteria will inevitably find ways of resisting the antibiotics we develop, which is why aggressive action is needed now to keep new resistance from developing and to prevent the resistance that already exists from spreading.

[0846] The invention provides improved means for targeting antibiotic-resistant hosts and for reducing the likelihood of hosts developing further resistance to the compositions of the invention.

[0847] Further examples of host cells and targeting of antibiotic resistance in such cells using the present invention are as follows:-

[0848] In a first example, optionally the host cell(s) are Staphylococcus *aureus* cells, eg, resistant to an antibiotic selected from methicillin, vancomycin, linezolid, daptomycin, quinupristin, dalfopristin and teicoplanin and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0849] In a second example, optionally the host cell(s) are *Pseudomonas aeuroginosa* cells, eg, resistant to an antibiotic selected from cephalosporins (eg, ceftazidime), carbapenems (eg, imipenem or meropenem), fluoroquinolones, aminoglycosides (eg, gentamicin or tobramycin) and colistin and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic. In a third example, optionally the host cell(s) are *Klebsiella* (eg, *pneumoniae)* cells, eg, resistant to carbapenem and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0850] In a fourth example, optionally the host cell(s) are *Streptoccocus (eg, thermophilus, pneumoniae* or pyogenes) cells, eg, resistant to an antibiotic selected from erythromycin, clindamycin, beta-lactam, macrolide, amoxicillin, azithromycin and penicillin and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0851] In a fifth example, optionally the host cell(s) are *Salmonella* (eg, serotype Typhi) cells, eg, resistant to an

antibiotic selected from ceftriaxone, azithromycin and ciprofloxacin and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0852] In a sixth example, optionally the host cell(s) are *Shigella* cells, eg, resistant to an antibiotic selected from ciprofloxacin and azithromycin and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0853] In a seventh example, optionally the host cell(s) are *mycobacterium tuberculosis* cells, eg, resistant to an antibiotic selected from Resistance to isoniazid (INH), rifampicin (RMP), fluoroquinolone, amikacin, kanamycin and capreomycin and azithromycin and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0854] In an eighth example, optionally the host cell(s) are *Enterococcus* cells, eg, resistant to vancomycin and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0855] In a ninth example, optionally the host cell(s) are *Enterobacteriaceae* cells, eg, resistant to an antibiotic selected from a cephalosporin and carbapenem and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0856] In a tenth example, optionally the host cell(s) are *E. coli* cells, eg, resistant to an antibiotic selected from trimethoprim, itrofurantoin, cefalexin and amoxicillin and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0857] In an eleventh example, optionally the host cell(s) are *Clostridium* (eg, *dificile)* cells, eg, resistant to an antibiotic selected from fluoroquinolone antibiotic and carbapenem and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0858] In a twelfth example, optionally the host cell(s) are Neisseria *gonnorrhoea* cells, eg, resistant to an antibiotic selected from cefixime (eg, an oral cephalosporin), ceftriaxone (an injectable cephalosporin), azithromycin and tetracycline and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0859] In a thirteenth example, optionally the host cell(s) are *Acinetoebacter baumannii* cells, eg, resistant to an antibiotic selected from beta-lactam, meropenem and a carbapenem and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0860] In a fourteenth example, optionally the host cell(s) are Campylobacter cells, eg, resistant to an antibiotic selected from ciprofloxacin and azithromycin and the host target site (or one or more of the target sites) is comprised by a gene conferring host resistance to said antibiotic.

[0861] In a fifteenth example, optionally, the host cell(s) produce Beta (β)-lactamase.

[0862] In a sixteenth example, optionally, the host cell(s) are bacterial host cells that are resistant to an antibiotic recited in any one of the first to fourteenth examples.

[0863] In an embodiment, the host cell is a USA300 *S aureus* strain cell.

[0864] In an example, the or each host target sequence is comprised by a plasmid of the host cell, eg, a *S aureus* plasmid (eg, of a USA300 strain), eg, a target comprised by the pUSA01, pUSA02 or pUSA03 plasmid of a *S aureus* cell. In an example, the first and/or second target is comprised by a host *mecA, mecA2* or *sek* gene sequence (eg, of a *S aureus* strain cell). In an example, the first and/or second target is comprised by a host pathogenicity island nucleotide (eg, DNA) sequence. In example, a spacer of the disclosure comprises or consists of a spacer disclosed in Table 1 on page 26 of WO2014/124226, which spacer sequences are incorporated herein by reference. In an example, the engineered sequence, HM-crRNA or gRNA comprises such a spacer.

[0865] The composition, use, method system, vector, collection, array, engineered sequence, virus, phage, phagemid, prophage or virion of the disclosure which is effective to reduce or kill or inhibit growth of an antibotic-resistant bacterial host in a mouse skin colonisation assay (eg, as disclosed in WO2014/124226, Kugelberg E, et al. Establishment of a superficial skin infection model in mice by using Staphylococcus aureus and Streptococcus pyogenes. Antimicrob Agents Chemother. 2005;49:3435-3441 or Pastagia M, et al. A novel chimeric lysin shows superiority to mupirocin for skin decolonization of methicillin-resistant and -sensitive Staphylococcus aureus strains. Antimicrob Agents Chemother. 2011;55:738-744) wherein the first and/or second target is comprised by a host gene that confers resistance to said antibiotic, eg, wherein the host is a *S aureus* (eg, USA300 strain) host.

[0866] Reference *S pyogenes* sequence is available under Genbank accession number NC_002737. with the cas9 gene at position 854757-858863. The *S pyogenes* Cas9 amino acid sequence is available under number NP_269215. These sequences are incorporated herein by reference for use in the present invnention. Further sequences as disclosed in 20150079680, whether explicitly or incorporated by reference therein, are also incorporated herein by reference for use in the present invention. Reference is also made to the disclosure of sequences and methods in WO2013/176772, which is incorporated herein by reference. Example tracrRNA sequences are those disclosed on page 15 of WO2014/124226, which are incorporated herein by reference for use in the present invention.

[0867] In an example, the or each repeat comprises or consists of from 20 to 50 (eg, from 24 to 47, eg, 30, 29, 28, 27, 26, 25 or 24) contiguous nucleotides in length.

[0868] In an example, the or each spacer comprises or consists of from 18 to 50 (eg, from 24 to 47, or 20 to 40, eg, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18), eg, 19 or 20 contiguous nucleotides in length.

[0869] In an example, the first repeat (most 5' in the HM-array of the disclosure) is immediately 5' of a spacer sequence that is complementary to a sequence comprising the first host target. This is useful, in view of the observation that newly acquired spacers (eg of invading phage sequence) are commonly incorporated at this position in bacteria, and thus positioning of the first spacer of the disclosure in this way is useful to promote its use.

[0870] In an example, the virus (eg, phage) nucleic acid comprises an origin of replication (*ori*) and a packaging site. In an example, the nucleic acid of the virus also comprises one, more or all genes encoding essential capsid proteins, eg, *rinA, terS* and *terL* genes. In an example, one, more or all of these is instead comprised by a companio helper virus (eg, helper phage) that is for con-infection with the virus of the disclosure - this frees up space in the the latter for including more HM-array nucleic acid and/or more Cas-encoding nucleic acid operable in the host. In an example, the virus nucleic acid comprises a fragment of a wild-type phage genome, wherein the fragment consists of consecutive nucleotides of the genome comprising at least the *rinA, terS* and *terL* genes or equivalent genes encoding phage proteins.

[0871] In an example, the host cell is of a strain or species found in human microbiota.

[0872] In an example, the or each target site is comprised by a gene that mediates host pathogenic adhesion, colonisation, invasion, immune response inhibition, virulence, essential protein or function expression or toxin generation. In an example, the gene is a gene encoding a cytotoxin, alpha-haemolysin, beta-haemolysin, gamma-haemolysin, leukocidin, Panton-Valentine lekocidin (PVL), exotoxin, TSST-1, enterotoxin, SEA, SEB, SECn, SED, SEE, SEG, SEH, SEI, exfolative toxin, ETA or ETB, optionally wherein the host is *S aureus,* eg, MRSA.

[0873] In an example, the or each CRISPR array is an array according to any of the configurations, embodiments, examples, concepts, aspects, paragraphs or clauses disclosed herein. In an example, the or each engineered nucleotide sequence is an engineered nucleotide sequence according to any of the configurations, embodiments, examples, concepts, aspects, paragraphs or clauses disclosed herein.

[0874] In an example, the or each vector is a vector according to any of the configurations, embodiments, examples, concepts, aspects, paragraphs or clauses disclosed herein.

[0875] In an example according to any of the configurations, embodiments, examples, Aspects, paragraphs or clauses disclosed herein, the vector or MGE is or comprises a casposon. MGEs are described further below. In an example, the casposon is a family 1, 2 or 3 casposon. In an example, an MGE of the disclosure comprises casposon terminal inverted repeats and optionally a casposon Cas1-encoding sequence. In an example, an MGE of the disclosure is or comprises a casposon minus Cas1 and operable for mobilisation with Cas1 of a host cell. See BMC Biol. 2014 May 19;12:36. doi: 10.1186/1741-7007-12-36, "Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity", Krupovic M *et al* for details of casposons.

FURTHER EXAMPLE APPLICATIONS OF THE PRESENT INVENTION

[0876] In an example, the composition (eg, HM-compostion or engineered sequence in combination with antibiotic) is as any of the following: In an example, the composition is a medical, opthalmic, dental or pharmaceutical composition (eg, comprised by a an anti-host vaccine). In an example, the composition is a an antimicrobial composition, eg, an antibiotic or antiviral, eg, a medicine, disinfectant or mouthwash. In an example, the composition is a cosmetic composition (eg, face or body make-up composition). In an example, the composition is a herbicide. In an example, the composition is a pesticide (eg, when the host is a *Bacillus* (eg, *thuringiensis)* host). In an example, the composition is a beverage (eg, beer, wine or alcoholic beverage) additive. In an example, the composition is a food additive (eg, where the host is an *E coli, Salmonella, Listeria* or *Clostridium* (eg, *botulinum)* host). In an example, the composition is a water additive. In an example, the composition is a additive for acquatic animal environments (eg, in a fish tank). In an example, the composition is an oil or petrochemical industry composition or comprised in such a composition (eg, when the host is a sulphate-reducing bacterium, eg, a *Desulfovibrio* host). In an example, the composition is a oil or petrochemical additive. In an example, the composition is a chemical additive. In an example, the composition is a disinfectant (eg, for sterilizing equipment for human or animal use, eg, for surgical or medical use, or for baby feeding). In an example, the composition is a personal hygiene composition for human or animal use.

[0877] In an example, the composition is a composition for environmental use, eg, for soil treatment or environmental decontamination (eg, from sewage, or from oil, a petrochemical or a chemical, eg, when the host is a sulphate-reducing bacterium, eg, a *Desulfovibrio* host). In an example, the composition is a plant growth stimulator. In an example, the composition is a composition for use in oil, petrochemical, metal or mineral extraction. In an example, the composition is a fabric treatment or additive. In an example, the composition is an animal hide, leather or suede treatment or additive. In an example, the composition is a dye additive. In an example, the composition is a beverage (eg, beer or wine) brewing or fermentation additive (eg, when the host is a *Lactobacillus* host). In an example, the composition is a paper additive. In an example, the composition is an ink additive. In an example, the composition is a glue additive. In an example, the composition is an anti-human or animal or plant parasitic composition. In an example, the composition is an air additive

(eg, for air in or produced by air conditoning equipment, eg, where the host is a *Legionella* host). In an example, the composition is an anti-freeze additive (eg, where the host is a *Legionella* host). In an example, the composition is an eyewash or opthalmic composition (eg, a contact lens fluid). In an example, the composition is comprised by a dairy food (eg, the composition is in or is a milk or milk product; eg, wherein the host is a *Lactobacillus, Streptococcus, Lactococcus* or *Listeria* host). In an example, the composition is or is comprised by a domestic or industrial cleaning product (eg, where the host is an *E coli, Salmonella, Listeria* or *Clostridium (eg, botulinum)* host). In an example, the composition is comprised by a fuel. In an example, the composition is comprised by a solvent (eg, other than water). In an example, the composition is a baking additive (eg, a food baking additive). In an example, the composition is a laboratory reagent (eg, for use in biotechnology or recombinant DNA or RNA technology). In an example, the composition is comprised by a fibre retting agent. In an example, the composition is for use in a vitamin synthesis process. In an example, the composition is an anti-crop or plant spoiling composition (eg, when the host is a saprotrophic bacterium). In an example, the composition is an anti-corrosion compound, eg, for preventing or reducing metal corrosion (eg, when the host is a sulphate-reducing bacterium, eg, a *Desulfovibrio* host, eg for use in reducing or preventing corrosion of oil extraction, treatment or containment equipment; metal extraction, treatment or containment equipment; or mineral extraction, treatment or containment equipment). In an example, the composition is an agricultural or farming composition or comprised in such a composition. In an example, the composition is a silage additive. The disclosure provides a HM-CRISPR array, HM-CRISPR/Cas system, HM-crRNA, HM-spacer, HM-DNA, HM-Cas, HM-composition or gRNAas described herein for use in any of the compositions described in this paragraph or for use in any application described in this paragraph, eg, wherein the host cell is a mircrobial cell or a bacterial or archaeal cell. The disclosure provides a method for any application described in this paragraph, wherein the method comprises combining a HM-CRISPR array, HM-CRISPR/Cas system, HM-crRNA, HM-spacer, HM-DNA, HM-Cas, gRNA or HM-composition of the disclosure with a host cell (eg, mircrobial, bacterial or archaeal cell). In an embodiment, the host cell is not present in or on a human (or human embryo) or animal.

[0878] Any aspect of the present invention is for an industrial or domestic use, or is used in a method for such use. For example, it is for or used in agriculture, oil or petroleum industry, food or drink industry, clothing industry, packaging industry, electronics industry, computer industry, environmental industry, chemical industry, aeorspace industry, automotive industry, biotechnology industry, medical industry, healthcare industry, dentistry industry, energy industry, consumer products industry, pharmaceutical industry, mining industry, cleaning industry, forestry industry, fishing industry, leisure industry, recycling industry, cosmetics industry, plastics industry, pulp or paper industry, textile industry, clothing industry, leather or suede or animal hide industry, tobacco industry or steel industry.

[0879] Herein, where there is mention of a *Desulfovibrio* host, the host can be instead a *Desulfobulbus, Desulfobacter, Desulfobacterium, Desulfococcus, Desulfomonile, Desulfonema, Desulfobotulus or Desulfoarculus* host or any other sulphur-reducing bacterium disclosed herein. In an embodiment for oil, water, sewage or environmental application, the host is a *Desulfovibrio capillatus* host.

[0880] Extensive microbiological analysis and 16S rRNA sequencing have indicated that the genus *Desulfovibrio* is but one of about eight different groups of sulfate-reducing eubacteria that can be isolated from the environment. Seven of these groups are gram-negative, while one represents the gram-positive bacteria (*Desulfotomaculum*). The genus Desulfovibrio has a rather small genome. Initial estimates were 1.7 Mbp and 1.6 Mbp for the genomes of *D. vulgaris* and *D. gigas* (which may be hosts according to the invention), respectively. This aids indentification of desired target sequences (eg, a sequence in an essential or reistance gene) for use in the invention. Characterization of an indigenous plasmid of *D. desulfuricans* (which may be ahost according to the invention) G200 has allowed the construction of a shuttle vector (Wall 1993, which vector may be used as a vector for the present invention), and the isolation and characterization of two bacteriophages from *D. vulgaris* Hildenborough (which may be ahost according to the invention) (Seyedirashti, 1992) may provide other ways to efficiently genetically manipulate *Desulfovibrio* spp. In an example, the vector is a mu or mu-like bacteriophage.

[0881] An example host is Desulfovibrio vulgaris subsp. vulgaris Postgate and Campbell (ATCC® 29579™)strain designation: NCIB 8303 [DSM 644, Hildenborough].

[0882] Treatment of the bacteria with mitomycin C or UV has previously been used to induce phage from the bacteria (Driggers & Schmidt), and this is a suitable method for obtaining suitable host-matched phage for generating a vector for use in any example or aspect of the present invention.

[0883] An application of the invention is in the dairy industry (eg, cheese or butter or milk products manufacture) or fermenting (eg, wine or vinegar or soy) or beer brewing or bread making industries. For example, for dairy industry application, a method of the disclosure is a method for producing a dairy food, comprising fermenting a culture of lactic acid-producing bacteria (eg, *Lactobacillus* host cells) for a period of time to produce lactic acid from the culture, and thereafter inhibiting growth of the bacteria by causing expression of crRNA from one or more arrays, systems, vectors, populations or collections of the disclosure mixed with the bacteria, whereby lactic acid production by the bacteria is reduced or inhibited. This is useful for reducing food/drink spoiling or undesirable food/ drink taste and/or odour. On an example there is included an inducible HM-array in the bacteria, wherein the method comprises adding an inducer agent

after the first period.

References

[0884]

Wall, J. D., B. J. Rapp-Giles, and M. Rousset. 1993. "Characterization of a small plasmid from Desulfovibrio desulfuricans and its use for shuttle vector construction". J. Bacteriol. 175:4121-4128;
Seyedirashti S et al; J Gen Microbiol. 1992 Jul;138(7):1393-7, "Molecular characterization of two bacteriophages isolated from Desulfovibrio vulgaris NCIMB 8303 (Hildenborough)";
Driggers & Schmidt, J. gen. Virol. (1970), 6, 421-427, "Induction of Defective and Temperate Bacteriophages in Caulobacter".

CONCEPTS:

Altering the Relative Ratio of Sub-Populations of First **and Second Bacteria in** a **Mixed Population of Bacteria,** eg, **in Microbiota**

[0885] In a first concept, there is provided use of a host modifying (HM) CRISPR/Cas system for altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria, the second bacteria comprising host cells, for each host cell the system comprising components according to (i) to (iv):-

(i) at least one nucleic acid sequence encoding a Cas nuclease;
(ii) a host cell target sequence and an engineered host modifying (HM) CRISPR array comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that hybridises to the host cell target sequence to guide said Cas to the target in the host cell to modify the target sequence;
(iii) an optional tracrRNA sequence or a DNA sequence expressing a tracrRNA sequence;
(iv) wherein said components of the system are split between the host cell and at least one nucleic acid vector that transforms the host cell, whereby the HM-crRNA guides Cas to the target to modify the host CRISPR/Cas system in the host cell; and

wherein the target sequence is modified by the Cas whereby the host cell is killed or host cell growth is reduced.

[0886] A second concept relates to a host modifying (HM) CRISPR/Cas system for the use of the first concept for modifying a target nucleotide sequence of a bacterial host cell, the system comprising components according to (i) to (iv):-

(i) at least one nucleic acid sequence encoding a Cas nuclease;
(ii) a host cell target sequence and an engineered host modifying (HM) CRISPR array comprising a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to the host target sequence to guide said Cas to the target in the host cell to modify the target sequence;
(iii) an optional tracrRNA sequence or a DNA sequence for expressing a tracrRNA sequence;
(iv) wherein said components of the system are split between the host cell and at least one nucleic acid vector that can transform the host cell, whereby the HM-crRNA guides Cas to the target to modify the host CRISPR/Cas system in the host cell.

[0887] A third concept relates to the system of the second concept, wherein the vector or vectors lack a Cas (eg, a Cas9) nuclease-encoding sequence.

[0888] A fourth concept relates to the use or system of any preceding concept, wherein each host cell is of a strain or species found in human microbiota.

[0889] A fifth concept relates to the use of the first to fourth concept for (a) the alteration of the proportion of *Bacteroidetes* bacteria in a mixed bacterial population; (b) reducing the proportion of a *Firmicutes* sub-population (host cells) in a mixed bacterial population; (c) reducing the proportion of a first *Firmicutes* species (host cells) in a mixed population, wherein the mixed population comprises a second *Firmicutes* species whose growth is not inhibited by said cRNA; (d) reducing the proportion of a first gram positive bacterial species (host cells) in a mixed bacterial population, wherein the mixed population comprises a second gram positive bacterial species whose growth is not inhibited by said cRNA; (e) reducing the proportion of a bacterial species (host cells) in a mixed bacterial population, wherein the mixed population comprises a different bacterial species whose growth is not inhibited by said cRNA, wherein the first species has 16s ribosomal RNA-encoding DNA sequence that is at least 80, 82, 83, 84, 85, 90 or 95% identical to an 16s ribosomal RNA-encoding DNA sequence of the other species;(f) reducing the proportion of a first bacterial human gut microbiota species (host

cells, eg, a *Firmicutes)* in a mixed bacterial population, wherein the mixed population comprises a different bacterial species, wherein the different species is a human gut probiotic species whose growth is not inhibited by said cRNA; or (g) reducing the proportion of a bacterial human gut microbiota species ((host cells, eg, a Firmicutes) in a mixed bacterial population, wherein the mixed population comprises a different bacterial species, wherein the different species is a human gut commensal species whose growth is not inhibited by said cRNA.

[0890] A sixth concept relates to the system of the second or third concept for (a) the alteration of the proportion of *Bacteroidetes* bacteria in a mixed bacterial population; (b) reducing the proportion of a *Firmicutes* sub-population (host cells) in a mixed bacterial population; (c) reducing the proportion of a first *Firmicutes* species (host cells) in a mixed population, wherein the mixed population comprises a second *Firmicutes* species whose growth is not inhibited by said cRNA; (d) reducing the proportion of a first gram positive bacterial species (host cells) in a mixed bacterial population, wherein the mixed population comprises a second gram positive bacterial species whose growth is not inhibited by said cRNA; (e) reducing the proportion of a bacterial species (host cells) in a mixed bacterial population, wherein the mixed population comprises a different bacterial species whose growth is not inhibited by said cRNA, wherein the first species has 16s ribosomal RNA-encoding DNA sequence that is at least 80, 82, 83, 84, 85, 90 or 95% identical to an 16s ribosomal RNA-encoding DNA sequence of the other species;(f) reducing the proportion of a first bacterial human gut microbiota species (host cells, eg, a *Firmicutes)* in a mixed bacterial population, wherein the mixed population comprises a different bacterial species, wherein the different species is a human gut probiotic species whose growth is not inhibited by said cRNA; or (g) reducing the proportion of a bacterial human gut microbiota species (host cells, eg, a Firmicutes) in a mixed bacterial population, wherein the mixed population comprises a different bacterial species, wherein the different species is a human gut commensal species whose growth is not inhibited by said cRNA; wherein (a) to (g) are for treating or preventing in a human or animal subject (i) a microbiota infection by said bacterial species whose proportion is reduced; or (ii) a disease or condition mediated by said bacterial species whose proportion is reduced.

[0891] A seventh concept relates to the use or system of the fifth or sixth concept for increasing the relative ratio of Bacteroidetes versus Firmicutes.

[0892] An eighth concept relates to the use or system of any preceding concept, wherein said Cas nuclease is provided by an endogenous Type II CRISPR/Cas system of the cell.

a ninth concept relates to the use or system of any preceding concept, wherein component (i) is endogenous to the host cell.

[0893] A tenth concept relates to the use or system of any preceding concept, wherein the target sequence is comprised by an antibiotic resistance gene, virulence gene or essential gene of the host cell.

[0894] An eleventh concept relates to the use or system of any preceding concept, wherein the target sequence is a host chromosomal sequence.

[0895] A twelfth concept relates to the use or system of any preceding concept, wherein alternatively HM-crRNA and tracrRNA are comprised by a single guide RNA (gRNA), eg provided by the vector.

[0896] A thirteenth concept relates to the use or system of any preceding concept, wherein the host cell comprises a deoxyribonucleic acid strand with a free end (HM-DNA) encoding a HM-sequence of interest and/or wherein the system comprises a sequence encoding the HM-DNA, wherein the HM-DNA comprises a sequence or sequences that are homologous respectively to a sequence or sequences in or flanking the target sequence for inserting the HM-DNA into the host genome (eg, into a chromosomal or episomal site).

[0897] A fourteenth concept relates to an engineered nucleic acid vector for the use of the first concept for modifying a bacterial host cell comprising an endogenous CRISPR/Cas system, the vector

) comprising nucleic acid sequences for expressing a plurality of different crRNAs (eg, comprised by gRNAs) for use in a CRISPR/Cas system or use according to any preceding concept; and) optionally lacking a nucleic acid sequence encoding a Cas nuclease,

wherein a first of said crRNAs is capable of hybridising to a first nucleic acid sequence in said host cell; and a second of said crRNAs is capable of hybridising to a second nucleic acid sequence in said host cell, wherein said second sequence is different from said first sequence; and

the first sequence is comprised by an antibiotic resistance gene (or RNA thereof) and the second sequence is comprised by an antibiotic resistance gene (or RNA thereof); optionally wherein the genes are different;

the first sequence is comprised by an antibiotic resistance gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof);

the first sequence is comprised by an essential gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof); or

the first sequence is comprised by a virulence gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof).

[0898] A fifteenth concept relates to the vector of the fourteenth concept inside a host cell comprising one or more

Cas that are operable with cRNA (eg, single guide RNA) encoded by the vector.

[0899] A sixteenth concept relates to the use, system or vector of any preceding concept, wherein the HM-CRISPR array comprises multiple copies of the same spacer.

[0900] A seventh concept relates to the use, system or vector of any preceding concept, wherein the vector(s) comprises a plurality of HM-CRISPR arrays.

[0901] An eighteenth concept relates to the use, system or vector of any preceding concept, wherein each vector is a virus or phage.

[0902] A nineteenth concept relates to the use, system or vector of any preceding concept, wherein the system or vector comprises two, three or more of copies of nucleic acid sequences encoding crRNAs (eg, gRNAs), wherein the copies comprise the same spacer sequence for targeting a host cell sequence (eg, a virulence, resistance or essential gene sequence).

[0903] A twentieth concept relates to the use, system or vector of the nineteenth concept, wherein the copies are split between two or more vector CRISPR arrays.

[0904] A twenth-first concept relates to a bacterial host cell comprising a system or vector recited in any preceding concept.

[0905] A twenthy-second concept relates to the use, system, vector or cell of any preceding concept, wherein the array is in combination with an antibiotic agent; or the use comprising exposing the host cells to a first antibiotic, wherein the target sequence is comprised by an antibiotic resistance gene for resistance to said first antibiotic.

[0906] A twenty-third concept relates to the use, system, vector or cell of any preceding concept, wherein the host cell is a *Staphylococcus, Streptococcus, Pseudomonas, Salmonella, Listeria, E coli, Desulfovibrio,* or *Clostridium* host cell.

[0907] A twenthy-fourth concept relates to the use, system or cell of any one of the first to thirteenth or sixteenth or twenty-third concepts, wherein each vector is according to the fourteenth or fifteenth concept.

[0908] A twenty-fifth concept relates to the use, system, vector or cell of any preceding concept wherein host cell population growth is reduced by of at least 5-fold compared to the growth of a population of said host cells not transformed with said HM-array or a nucleotide sequence encoding said gRNA.

[0909] A twenty-sixth concept relates to the use, system, vector or cell of any preceding concept wherein host cell population growth on a surface is inhibited. In an example, the population is in contact with a human tissue surface (eg, a gut tissue surface, eg, *in vivo* or *ex vivo.).*

[0910] A twenty-seventh concept relates to the use, system, vector or cell of any preceding concept wherein the first bacteria are probiotic, commensal or symbiotic with humans (eg, in the human gut).

[0911] A twenty-eighth concept relates to the use, system, vector or cell of any preceding concept wherein the first and second bacteria are both *Firmicutes* and are bacteria of different species or strains; or wherein the first bacteria are *Enterobacteriaceae* and the second bacteria are *Firmicutes.*

[0912] A twenty-ninth concept relates to the use, system, vector or cell of any preceding concept wherein the host cells are archaeal cells instead of bacterial cells or each population is an archaeal population instead of a bacterial population.

[0913] A thirtieth concept relates to the use of any one of the first, foruth, fifth, seventh to thirteenth, sixteenth to twentieth and twenty-second to twenty-ninth concepts for treating an industrial or an *ex vivo* medical fluid, surface, apparatus or container; or for treating a waterway, water, a beverage, a foodstuff or a cosmetic, wherein the host cell(s) are comprised by or on the fluid, surface, apparatus, container, waterway, water, beverage, foodstuff or cosmetic.

[0914] A thirty-first concept relates to the use, system, vector or cell of any preceding concept, wherein the HM-cRNA or gRNA comprises a sequence that is capable of hybridising to a host cell target protospacer sequence that is a adjacent a NNAGAAW or NGGNG protospacer adjacent motif (PAM).

[0915] A thirty-second concept relates to a nucleic acid vector according to, or for use in, the use, system or cell of any preceding concept, the vector comprising more than 1.4kb of exogenous DNA sequence, wherein the exogenous DNA encodes one or more components of a CRISPR/Cas system and comprises an engineered array for expressing HM-crRNAs or gRNAs in host cells, wherein the exogenous sequence is devoid of a nucleotide sequence encoding a Cas nuclease that is cognate to the cRNA(s) or gRNA(s); wherein at least 2 different cRNAs or gRNAs are encoded by the exogenous DNA (eg, by at least 2 HM-CRISPR arrays).

[0916] A thirty-third concept relates to the vector of the thirty-second concept, wherein the vector is a viral vector capable of transforming host cells.

[0917] A thirty-foruth concept relates to the vector of the thirty-second to thirty-third concept, wherein the cRNAs or gRNAs are capable of hybridising in host cells to respective target protospacer sequences, wherein each protospacer sequence is comprised by an antibiotic resistance or essential host gene.

[0918] A thirty-fifth concept relates to the vector of any one of the thirty-fourth to thirty-sixth concepts, wherein the host cells are cells of a human microbiota species.

CHARACTERISATIONS:

Harnessing Wild-Type Endogenous Cas For Population Growth Inhibition & Treatment of Bacteria on Surfaces

[0919] In a first characterisation, there is provided use of wild-type endogenous Cas nuclease activity of a bacterial host cell population to inhibit growth of the population, wherein the population comprises a plurality of host cells and each host cell has an endogenous CRISPR/Cas system having wild-type Cas nuclease activity, the use comprising transforming host cells of the population, wherein each transformed host cell is transformed with an engineered nucleotide sequence for providing host modifying (HM) cRNA or guide RNA (gRNA) in the host cell, the HM-cRNA or gRNA comprising a sequence that is capable of hybridising to a host cell target protospacer sequence for guiding endogenous Cas to the target, wherein the cRNA or gRNA is cognate to an endogenous Cas nuclease of the host cell that has said wild-type nuclease activity and following said transformation of the host cells growth of the population is inhibited.

[0920] The host cells may be of the same species or strain.

[0921] A second characterisation relates to the use of the first characterisation, wherein the inhibition of host cell population growth is a reduction in growth of at least 5-fold compared to the growth of a population of said host cells not transformed with said engineered nucleotide sequence.

[0922] A third characterisation relates to the use of the first characterisation, wherein population growth on a surface is inhibited.

[0923] A fourth characterisation relates to the use of the second characterisation, wherein population growth on a surface is inhibited.

[0924] A fifth characterisation relates to the use of the first characterisation, said inhibiting comprising using a HM-CRISPR/Cas system for killing or reducing the growth of said host cells, for each host cell the system comprising components according to (i) to (iv):-

(i) at least one nucleic acid sequence encoding said Cas nuclease;
(ii) an engineered host modifying HM-CRISPR array comprising a spacer sequence (HM-spacer) and repeats encoding said HM-crRNA;
(iii) an optional tracrRNA sequence or a DNA sequence expressing a tracrRNA sequence;
(iv) wherein said components of the system are split between the host cell and at least one nucleic acid vector that transforms the host cell, whereby the HM-crRNA guides said Cas to the target to modify the target sequence;

wherein the target sequence is modified in host cells by the Cas whereby the host cells are killed or host cell growth is reduced.

[0925] A sixth characterisation relates to the use of any preceding characterisation, for altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria, the second bacteria comprising said host cells.

[0926] A seventh characterisation relates to the use of the sixth characterisation, wherein the host cells are of a strain or species found in human microbiota.

[0927] An eighth characterisation relates to the use of the sixth or seventh characterisation, wherein the host cells are mixed with cells of a different strain or species, wherein the different cells are *Enterobacteriaceae* or bacteria that are probiotic, commensal or symbiotic with humans (eg, in the human gut).

[0928] A ninth characterisation relates to the use of any preceding characterisation for the alteration of the proportion of *Bacteroidetes* bacteria in a mixed bacterial population comprising *Bacteroidetes* bacteria and other bacteria, optionally for increasing the relative ratio of *Bacteroidetes* versus one, more or all *Firmicutes* species (eg, versus *Streptococcus)* in the population.

[0929] A tenth characterisation relates to the use of any preceding characterisation for altering the relative ratio of first bacteria versus second bacteria in a mixed population, wherein the first and second bacteria are both *Firmicutes* and are bacteria of different species or strains, the second bacteria comprising host cells. In an example, the use increases the proportion of first to versus second bacteria. An eleventh characterisation relates to the use of the first characterisation, wherein the engineered nucleotide sequence is not in combination with an exogenous Cas nuclease-encoding sequence.

[0930] A twelfth characterisation relates to the use of the fifth characterisation, wherein the vector or vectors lack a Cas nuclease-encoding sequence.

[0931] A thirteenth characterisation relates to the use of the first characterisation, wherein each host cell is of a strain or species found in human microbiota.

[0932] A fourteenth characterisation relates to the use of the sixth characterisation, wherein each host cell is of a strain or species found in human microbiota.

[0933] A fifteenth characterisation relates to the use of the thirteenth characterisation, wherein each host cell is mixed with cells of a different strain or species, wherein the different cells are *Enterobacteriaceae* or bacteria that are probiotic,

commensal or symbiotic with humans (eg, in the human gut).

[0934] A sixteenth characterisation relates to the use of the first characterisation , wherein the use alters the proportion of Bacteroidetes bacteria in a mixed bacterial population comprising Bacteroidetes bacteria and other bacteria, optionally wherein the use alters the relative ratio of Bacteroidetes versus one, more or all Firmicutes (eg, Streptococcus) species in the population.

[0935] A seventeenth characterisation relates to the use of the first characterisation, wherein the first and second bacteria are both Firmicutes and the use alters the relative ratio of the first versus the second bacteria in the mixed population. In an example, the use increases the proportion of first to versus second bacteria.

[0936] An eighteenth characterisation relates to the use of the first characterisation, wherein said Cas nuclease is provided by a host cell endogenous Type II CRISPR/Cas system and/or the HM-cRNA or gRNA comprises a sequence that is capable of hybridising to a host cell target protospacer sequence that is a adjacent a 5'-NNAGAAW-3' protospacer adjacent motif (PAM).

[0937] A nineteenth characterisation relates to the use of the fifth characterisation, wherein said Cas nuclease is provided by a host cell endogenous Type II CRISPR/Cas system.

[0938] A twentieth characterisation relates to the use of the fifth characterisation, wherein component (iii) is endogenous to the host cell.

[0939] A twenty-first characterisation relates to the use of the fifth characterisation, wherein each transformed host cell comprises an endogenous RNase III that is operable with component (ii) in the production of said HM-crRNA in the cell.

[0940] A twenty-second characterisation relates to the use of the first characterisation, wherein the target sequence is comprised by an antibiotic resistance gene, virulence gene or essential gene of the host cell. A twenty-third characterisation relates to the use of the first characterisation, wherein the engineered nucleotide sequence is in combination with an antibiotic agent.

[0941] A twenty-fourth characterisation relates to the use of the fifth characterisation, wherein the HM-crRNA and tracrRNA are comprised by a single guide RNA (gRNA).

[0942] A twenty-fifth characterisation relates to the use of the first characterisation, wherein transformed host cells each comprise a deoxyribonucleic acid strand with a free end (HM-DNA) encoding a HM-sequence of interest, wherein the HM-DNA comprises a sequence or sequences that are homologous respectively to a sequence or sequences in or flanking the target sequence for inserting the HM-DNA into the host genome, wherein HM-DNA sequences are inserted into host cell genomes.

[0943] A twenty-sixth characterisation relates to the use of the first characterisation, comprising expressing in host cells a plurality of different crRNAs (or gRNAs) for hybridising to host cell protospacer target sequences; wherein a first of said crRNAs (or gRNAs) is capable of hybridising to a first protospacer nucleic acid sequence; and a second of said crRNAs (or gRNAs) is capable of hybridising to a second protospacer nucleic acid sequence, wherein said second sequence is different from said first sequence; and

(a) the first sequence is comprised by an antibiotic resistance gene (or RNA thereof) and the second sequence is comprised by an antibiotic resistance gene (or RNA thereof); optionally wherein the genes are different;
(b) the first sequence is comprised by an antibiotic resistance gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof);
(c) the first sequence is comprised by an essential gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof); or
(d) the first sequence is comprised by a virulence gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof).

[0944] A twenth-seventh characterisation relates to the use of the sixth characterisation, wherein the host cells are comprised by a mixed bacterial population comprised by a human or animal subject and the use (i) treats in the subject an infection by said host cells comprised by the mixed population; (ii) treats or reduces the risk in the subject of a condition or disease mediated by said host cells; (iii) reduces body odour of the human that is caused or mediated by said host cells; or (iv) is a personal hygiene treatment of the human.

[0945] A twenty-eighth characterisation relates to the use of the first characterisation, wherein the use treats or reduces the risk of an infection by said host cells in a human or animal subject, wherein host cells each comprise an antibiotic resistance gene (for resistance to a first antibiotic) which comprises said target protospacer sequence, wherein the use comprises administering the engineered nucleotide sequence and the first antibiotic to the subject, wherein the infection is reduced or prevented in the subject.

[0946] A twenty-ninth characterisation relates to the use of the first characterisation, wherein each engineered nucleotide sequence further comprises an antibiotic resistance gene, wherein the HM-crRNA or gRNA does not target the antibiotic resistance gene and the use comprises exposing the population to said antibiotic and a plurality of said engineered sequences, thereby promoting maintenance of HM-crRNA or gRNA-encoding sequences in host cells.

[0947] A thirtieth characterisation relates to the use of the first characterisation, wherein the host cells are gram positive cells or *Streptococcus, Staphylococcus, Pseudomonas, Salmonella, Listeria, E coli, Desulfovibrio, V cholerae* or *Clostridium* cells.

[0948] A thirty-first characterisation relates to the use of the first characterisation for treating an industrial or medical fluid, surface, apparatus or container; or for treating a waterway, water, a beverage, a foodstuff or a cosmetic, wherein the host cells are comprised by or on the fluid, surface, apparatus, container, waterway, water, beverage, foodstuff or cosmetic, and wherein growth of the host cell population is inhibited thereby carrying out said treatment.

[0949] In an alternative, any characterisation is dependent from any preceding characterisation.

ASPECTS:

Horizontal Transfer Between Carrier & Host Cells in Mixed Populations

[0950] In a first aspect, there is provided a method for producing a mixed bacterial population comprising carrier bacteria, wherein the population comprises first and second sub-populations of first and second bacteria respectively, wherein the sub-populations are bacteria of first and second species that are different from each other and the second bacteria comprise a plurality of host cells, wherein the carrier bacteria are first bacteria cells each comprising an engineered nucleotide sequence for providing host cell modifying (HM) cRNA or guide RNA (gRNA) in the host cells, the HM-cRNA or gRNA comprising a sequence that is capable of hybridising to a host cell target protospacer sequence for guiding a first Cas nuclease to the target to modify the target, wherein the carrier bacteria do not comprise the target sequence, the method comprising

a. Providing a plurality of nucleic acids, each comprising a said engineered nucleotide sequence;
b. Combining said plurality of nucleic acids with a first mixed population comprising first and second sub-populations of the first and second bacterial species respectively, the second sub-population comprising host cells;
c.Allowing the nucleic acids to transform cells of said first sub-population in the presence of the host cells, thereby producing a second mixed population comprising said carrier cells and said host cells, wherein said engineered nucleotide sequence comprised by carrier cells is capable of horizontal transfer to host cells to transform host cells for production of said HM-cRNA or gRNA in transformed host cells.

[0951] A second aspect relates to the method of the first aspect, further comprising obtaining the second mixed population.

[0952] A third aspect relates to the method of the first aspect, further comprising isolating a plurality of carrier cells from the second mixed population.

[0953] A fourth aspect relates to the method of the first aspect, further comprising producing said HM-cRNA or gRNA in the transformed host cells, wherein said HM-crRNA or gRNA sequence hybridises to target protospacer sequence in said transformed host cells and guides the first Cas nuclease to the target, thereby modifying the target with the first Cas nuclease.

[0954] A fifth aspect relates to the method of the fourth aspect, further comprising obtaining host cells comprising said target modification (eg, wherein the host cells are comprised by a mixed population comprising said first bacterial species).

[0955] A sixth aspect relates to the method of the first aspect, wherein the cRNA or gRNA is cognate to an endogenous Cas nuclease of the host cells, wherein the nuclease is said first Cas nuclease.

[0956] A seventh aspect relates to the method of the first aspect, wherein the cRNA or gRNA is cognate to an endogenous Cas nuclease of the carrier cells, wherein the nuclease is said first Cas nuclease.

[0957] An eighth aspect relates to the method of the seventh aspect, wherein the nuclease has wild-type nuclease activity.

[0958] A ninth aspect relates to the method of the first, sixth, seventh or eighth aspect, wherein the first Cas nuclease is a Cas9.

[0959] A tenth aspect relates to the method of a ninth aspect, wherein the Cas9 is a *Streptococcus* Cas9.

[0960] An eleventh aspect relates to the method of the first aspect, wherein each engineered nucleotide sequence is comprised by a respective nucleic acid vector, wherein the vectors are capable of horizontal transfer between the carrier and host cells.

[0961] A twelfth aspect relates to the method of the first to eleventh aspect, wherein each engineered sequence is comprised by a respective mobile genetic element, eg, a transposon or plasmid.

[0962] A thirteenth aspect relates to the method of the first aspect, wherein following said transformation of host cells, growth of the host cell sub-population is inhibited.

[0963] A fourteenth aspect relates to the method of the thirteenth aspect, wherein the inhibition of host cell population growth is at least 5-fold compared to the growth of a population of said host cells not transformed with said engineered

nucleotide sequence.

[0964] A fifteenth aspect relates to the method of the thirteenth or fourteenth aspect, wherein host cell population growth on a surface is inhibited.

[0965] In an alternative, any aspect is dependent from any preceding aspect.

[0966] In an example, the method is a method of treating or preventing a disease or condition in a human, animal or plant subject, eg, as described herein, wherein the method effects said treatment or prevention. The disclosure provides a mixed bacterial population obtained or obtainable by the method for such a method of treating or preventing.

[0967] In an example, the method is carried out on a mixed bacterial population of an environment, equipment, apparatus, container, waterway, water, fluid, foodstuff, beverage, microbiota, microbiome or cosmetic, eg, as described herein, wherein the method reduces the proportion of host cells compared to first cells.

[0968] In an example, the product of the method is for administration to the gut of a human or non-human animal for treating or preventing obesity, diabetes or IBD of the human or animal.

[0969] In an example, the first and second species are species of human or non-human animal gut commensal or symbiotic bacteria.

[0970] The product of the method is useful as it can be adminstered (eg, intranasally) to a human or animal so that the bacteria populate one or more microbiomes (eg, gut microbiome) of the human or animal. The first cells act as carriers, especially when those cells are non-pathogenic to the human or animal (eg, non-pathogenic in the gut microbiome). The microbiome can be any other micribiome or microbiota population disclosed herein.

[0971] In an example, the first second bacterial species is capable of populating the gut microbiota of a human or non-human animal, and the first bacteria are commensal or symbiotic with humans or animals. Usefully, the first bacteria can be safely administered to the human or animal and can act as a carrier for transfer of engineered sequences thereafter to host cells of the microbiota.

[0972] In an example, the engineered sequence is comprised by any array or vector disclosed herein. In an example, the method uses any CRISPR/Cas system disclosed herein.

[0973] In an example the first cell is a *Bacteroidetes (eg,* Bacteroidales or *Bacteroides*) cell; *Lactobacillus* (eg, *acidophilus* (eg, La-5, La-14 or NCFM), *brevis, bulgaricus, plantarum, rhammosus, fermentum, caucasicus, helveticus, lactis, reuteri* or *casei* eg, *casei* Shirota); *Bifidobacterium* (eg, *bifidum, breve, longum* or *infantis); Streptococcus thermophiles; Enterococcusfaecium; Alistipes; Alkaliflexus; Parabacteroides; Tannerella; E coli;* or *Xylanibacter* cell.

[0974] In an example, the host cells are of a human microbiota species and the carrier cells are cells of a species that is non-pathogenic in said human microbiota, wherein the target sequence is not comprised by the genome of the carrier cells, the engineered sequence being comprised by a MGE comprising an *ori*T that is operable in the carrier and host cells, wherein the MGE is capable of horizontal transfer from the carrier cell to the host cell. In an example, the engineered sequence, MGE or vector is comprised by a bacteriophage, the bacteriophage being capable of infecting the first cells (carriers) to introduce the MGE into the first (carrier) cells. Thereafter the MGE is capable of horizontal transfer to host cells.

[0975] In an example, the first cells are *Bacteroidetes* or *Prevotella* cells; optionally wherein the MGE is capable of horizontal transfer from the first cell species to *Firmicutes* species (host cells) of said human microbiota. The latter is useful, for example, for treating or preventing obesity in a human when the target sequence is comprised by the *Firmicutes,* but not the first (carrier) cells.

FURTHER EXAMPLES

EXAMPLE 1: Environmental Treatment or Decontamination

Oil, Metal & Mineral Industry

[0976] In an embodiment, the host cell is in an a mineral mine or field; in a metal mine or field; in an oil field or in oil or a petrochemical (eg, for any of these when the host is an anaerobic sulphate-reducing bacterium, eg, a *Desulfovibrio* bacterium). In an example, this composition comprises an oxidising agent (eg, sodium hypochlorite), a quaternary ammonium compound or isothiazolone or is administered simultaneously or sequentially with sodium hypochlorite, a quaternary ammonium compound or isothiazolone. An example of a suitable vector for use in the present invention for modifying a *Desulfovibrio* bacterial host is a bacteriophage. The references below describe suitable methods for isolating phage that infect *Desulfovibrio.* For use as a vector in the present invention, the bacteriophage described by any of the references may be used. Alternatively, the vector is provided by nanoparticles.

[0977] Heidelberg *et al* describe the two copies of the nearly identical mu-like bacteriophage DVUO189-221, DVU2847-79, DVU2688-733 and remnants of bacteriophage are present in the genome of *Desufovibrio vulgaris* Hildenborough. Such a phage can be a basis on which to design a phage vector for use in the present invention.

References:

[0978]

Seyedirashti S et al , J Gen Microbiol. 1991 Jul;137(7):1545-9, "Induction and partial purification of bacteriophages from Desulfovibrio vulgaris (Hildenborough) and Desulfovibrio desulfuricans ATCC 13541;
Seyedirashti S et al , J Gen Microbiol. 1992 Jul;138(7):1393-7, "Molecular characterization of two bacteriophages isolated from Desulfovibrio vulgaris NCIMB 8303 (Hildenborough)";
*Walker CB et al ; Environ Microbiol. 2006 Nov;8(11):1950-9, "Recovery of temperate Desulfovibrio vulgaris bacteriophage using a novel host strain";
Miranda et al, Corrosion Science 48 (2006) 2417-2431, "Biocorrosion of carbon steel alloys by an hydrogenotrophic sulfate-reducing bacterium Desulfovibrio capillatus isolated from a Mexican oil field separator";
Eydal et al, The ISME Journal (2009) 3, 1139-1147; doi:10.1038/ismej.2009.66; published online 11 June 2009, "Bacteriophage lytic to Desulfovibrio aespoeensis isolated from deep groundwater";
Walker CB et al; Environ Microbiol. 2009 Sep;11(9):2244-52. doi: 10.1111/j.1462-2920.2009.01946.x, "Contribution of mobile genetic elements to Desulfovibrio vulgaris genome plasticity".

[0979] **[The sequences described in this article have been deposited in GenBank under Accession No. DQ826728-DQ826732, incorporated herein by reference]*

EXAMPLE 2: Water or Sewage Treatment or Environmental (eg, Soil) Metal Decontamination

[0980] An alternative application of the disclosure provides a HM-CRISPR array, HM-CRISPR/Cas system, HM-crRNA, HM-spacer, HM-DNA, HM-Cas or HM-composition as described herein for water or sewage treatment, eg wherein the host is a sulphate-reducing bacterium, eg, a *Desulfovibrio* bacterium.

[0981] In an example, the target nucleotide sequence in the host is a sequence of a heavy metal resistance gene. Optionally also the host is a *Desulfovibrio* bacterium, eg, *D vulgaris.*

EXAMPLE 3: Medical Use

[0982] An alternative application of the disclosure provides a HM-CRISPR array, HM-CRISPR/Cas system, HM-crRNA, HM-spacer, HM-DNA, HM-Cas or HM-composition as described herein for treating, preventing or reducing (eg, reducing spread of or expansion of) a bacterial infection in a human or animal.

[0983] In a first disclosure, the infection is caused by MRSA host cells in a human. The host cell is a *Staphylococcus aureus* host cell and a HM-array of the invention is contained in a population of Class I, II or III *Staphylococcus* packaged phage *(Caudovirales* or *Myoviridae* phage). The phage population is administered to a MRSA-infected patient with or without methicillin or vancomycin. In one trial, the phage HM-arrays target (i) the region of 20 nucleotides at the 3' of the leader promoter of endogenous *S aureus* CRISPR arrays and (ii) the methicillin resistance genes in the host cells. When vancomycin is administered, a lower dose than usual is administered to the patient. It is expected that host cell infection will be knocked-down and resistance to the phage medicine will not be established or established at a lower rate or severity than usual. In other trials, the design is identical except that the phage in those trials also target the essential *S aureus* gene *ftsZ* (Liang et al, Int J Infect Dis. 2015 Jan;30:1-6. doi: 10.1016/j.ijid.2014.09.015. Epub 2014 Nov 5, "Inhibiting the growth of methicillinresistant Staphylococcus aureus in vitro with antisense peptide nucleic acid conjugates targeting the ftsZ gene").

[0984] A further trial repeated the trials above, but phage K endolysin was administered in addition or instead of methicillin.

References

[0985]

1. Jiang W et al, Nucleic Acids Res. 2013 Nov;41(20):e188. doi: 10.1093/nar/gkt780. Epub 2013 Sep 2, "Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice";
2. Seed KD et al, Nature. 2013 Feb 28;494(7438):489-91. doi: 10.1038/nature11927, "A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity";
3. Semenova E et al, Proc Natl Acad Sci U S A. 2011 Jun 21;108(25):10098-103. doi: 10.1073/pnas.1104144108. Epub 2011 Jun 6, "Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence";

4. Heler R et al, Mol Microbiol. 2014 Jul;93(1):1-9. doi: 10.1111/mmi.12640. Epub 2014 Jun 4, "Adapting to new threats: the generation of memory by CRISPR-Cas immune systems";
5. Gomaa A et al, MBio. 2014 Jan 28;5(1):e00928-13. doi: 10.1128/mBio.00928-13, "Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems";
6. Fineran PC et al, Proc Natl Acad Sci U S A. 2014 Apr 22;111(16):E1629-38. doi: 10.1073/pnas.1400071111. Epub 2014 Apr 7, "Degenerate target sites mediate rapid primed CRISPR adaptation";
7. Wiedenheft et al, Nature. 2011 Sep 21;477(7365):486-9. doi: 10.1038/nature10402, "Structures of the RNA-guided surveillance complex from a bacterial immune system;
8. Bondy-Denomy et al, Nature 493, 429-432 (17 January 2013) doi:10.1038/nature11723, " Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system";
9. Nuñez JK et al, Nature. 2015 Mar 12;519(7542):193-8. doi: 10.1038/nature14237. Epub 2015 Feb 18, "Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity".

EXAMPLE 4: Altering the Ratio of Bacteria in a Mixed Gut Microbiota Population

[0986] Alteration of the ratio of bacteria will be performed according to the present example, which is described by reference to knocking-down *Clostridium difficil*e bacteria in a mixed gut microbiota sample. The sample will contain *Bacteroides* and metronidazole (MTZ)-resistant *C difficile* strain 630 sub-populations. *Ex vivo* the mixed population is combined with a population of carrier bacteria *(Lactobacillus acidophilus* La-14 and/or La-5) that have been engineered according to the invention to contain CRISPR arrays.

[0987] Each CRISPR array is comprised on a plasmid that is compatible with the carrier bacterium and C *difficile* cells. The array is comprised by a *Bacteroides thetaiotamicron* CTnDot transposon that also comprises *ori*T, an *intDOT* sequence*,* a tetQ-rteA-rteB operon, *rteC* and the operon *xis2c-xis2d-orf3-exc.* In one experiment, *mob* and *tra* operons are excluded (instead relying on these supplied by *Bacteroides* cells to which the transposons are transferred in the mixture combined with the carrier bacteria). In another experiment, the *mob* and *tra* operons are included in the transposons.

[0988] Protein translocation across the cytoplasmic membrane is an essential process in all bacteria. The Sec system, comprising at its core an ATPase, SecA, and a membrane channel, SecYEG, is responsible for the majority of this protein transport. A second parallel Sec system has been described in a number of Gram-positive species. This accessory Sec system is characterized by the presence of a second copy of the energizing ATPase, SecA2; where it has been studied, SecA2 is responsible for the translocation of a subset of Sec substrates. In common with many pathogenic Gram-positive species, *Clostridium difficil*e possesses two copies of SecA. Export of the S-layer proteins (SLPs) and an additional cell wall protein (CwpV) is dependent on SecA2. Accumulation of the cytoplasmic precursor of the SLPs SlpA and other cell wall proteins is observed in cells expressing dominant-negative *secA1* or *secA2* alleles, concomitant with a decrease in the levels of mature SLPs in the cell wall. Furthermore, expression of either dominant-negative allele or antisense RNA knockdown of SecA1 or SecA2 dramatically impairs growth, indicating that both Sec systems are essential in C. *difficile.*

[0989] *C. difficile* Strain 630 (epidemic type X) has a single circular chromosome with 4,290,252 bp (G+C content = 29.06%) and a circular plasmid with 7,881 bp (G+C content = 27.9%). The whole genome has been sequenced and found that 11% of the genome consists of mobile genetic elements such as conjugative transposons. These elements provide *C. difficil*e with the genes responsible for its antimicrobial resistance, virulence, host interaction and the production of surface structures. For example, the cdeA gene of C. *difficile* produces a multidrug efflux pump which was shown to be homologous to known efflux transporters in the multidrug and toxic compound extrusion (MATE) family. The protein facilitates energy-dependent and sodium-coupled efflux of drugs from cells. In addition, the cme gene in *C. difficile* has been shown to provide multidrug resistance in other bacteria.

[0990] The array comprises a R1-S1-R1' CRISPR unit for targeting a sequence in an essential gene (SecA2) of *C difficil*e cells. In another experiment, targeting is to the cdeA gene in the presence of MTZ and optionally one or more other anti-C *difficile* antibiotics. Each spacer (S) comprises a 20mer nucleotide sequence of the SecA or cdeA gene, wherein the sequence comprises a PAM of a *C difficil*e strain 630 CRISPR/Cas system that is cognate to the repeat sequences. Each repeat is identical to a *C difficile* strain 630 repeat and has the sequence
5'- ATTTACATACCACTTAGTTAATATAAAAC-3' (SEQ ID NO: 118)

[0991] In an alternative set of experiments, the following sequence is used for the repeats:
5'- GTTTTATATTAACTAAGTGGTATGTAAAT-3' (SEQ ID NO: 119)

[0992] The repeats function with Cas that is endogenous to the *C difficile* cells in the mixed population. The mixed population of bacteria is retreived as an *ex vivo* sample from a stool sample of a human patient suffering from *C difficile* infection. The mixed population is mixed with the carrier bacteria *in vitro* and incubated at 37 degrees centigrade under anaerobic conditions to simulate gut conditions in the presence of tetracycline. It is expected that transposons containing the CRISPR arrays will be transferred to *Bacteroides* and *C difficil*e cells in the mixture. Furthermore, it is expected that

the target sites in the latter cells will be cut by Cas nuclease action, thus reducing the proportion of *C dificile* in the mixed population (and increasing the ratio of *Bacteroides* versus *C dificile).*

[0993] In a follow-on experiment, a drink is produced comprising the carrier bacteria and this is consumed by the human patient once or twice for several consecutive days. The patient is also administered with tetracycline during the treatment period. It is expected that stool analysis will reveal that the proportion of *C dificile* in the stool samples will reduce (and the ratio of *Bacteroides* versus *C dificile* will increase).

EXAMPLE 5: Cholera Treatment or Prevention

[0994] Reference is made to the World Health Organisation (WHO) Cholera Fact sheet N°107 (Updated July 2015). Cholera is an acute diarrhoeal infection caused by ingestion of food or water contaminated with the bacterium *Vibrio cholerae.* Researchers have estimated that every year, there are roughly 1.4 to 4.3 million cases, and 28 000 to 142 000 deaths per year worldwide due to cholera. The short incubation period of 2 hours to 5 days, is a factor that triggers the potentially explosive pattern of outbreaks. Cholera is an extremely virulent disease. It affects both children and adults and can kill within hours. About 80% of people infected with *V. cholerae* do not develop any symptoms, although the bacteria are present in their faeces for 1-10 days after infection and are shed back into the environment, potentially infecting other people. Among people who develop symptoms, 80% have mild or moderate symptoms, while around 20% develop acute watery diarrhoea with severe dehydration. This can lead to death if left untreated.

[0995] Two serogroups of *V. cholerae* - O1 and 0139 - cause outbreaks. *V. cholerae* O1 causes the majority of outbreaks, while 0139 - first identified in Bangladesh in 1992 - is confined to South-East Asia. Non-O1 and non-O139 *V. cholerae* can cause mild diarrhoea but do not generate epidemics. Recently, new variant strains have been detected in several parts of Asia and Africa. Observations suggest that these strains cause more severe cholera with higher case fatality rates. The main reservoirs of *V. cholerae* are people and water-borne sources such as brackish water and estuaries, often associated with algal blooms.

[0996] Reference is made to Nature. 2013 Feb 28;494(7438):489-91. doi: 10.1038/nature11927, "A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity", Seed KD *et al* (incorporated herein by reference), which describes that *Vibrio cholerae* serogroup O1 is the primary causative agent of the severe diarrhoeal disease cholera, and lytic *V. cholerae* phages have been implicated in impacting disease burden particularly in the endemic region surrounding the Bay of Bengal. The authors described the isolation of the ICP1 (for the International Centre for Diarrhoeal Disease Research, Bangladesh cholera phage 1) -related, V. cholerae O1-specific virulent myoviruses that are omnipresent amongst cholera patient rice-water stool samples collected from 2001 to 201114 and in the study described in their publication.

[0997] The authors explain that ICP1 CRISPR/Cas system consists of two CRISPR loci (designated CR1 and CR2) and six cas genes whose organization and protein products are most homologous to Cas proteins of the type 1-F *(Yersinia pestis)* subtype system 17. *V. cholerae* is divided into two biotypes, classical and El Tor, the former of which is associated with earlier pandemics and has since been replaced by the El Tor biotype18. The classical strain, V. cholerae O395, has a CRISPR/Cas system belonging to the type I-E (Escherichia coli) subtype17, and to date there has not been any description of El Tor strains possessing a CRISPR/Cas system. Thus, the origin of the CRISPR/Cas system in ICP1 phage is unknown.

[0998] The RNA sequence of the CR1 and CR2 consensus direct repeat with the partially palindromic sequence forming the predicted stem in the crRNA underlined is as follows:-

GUUAGCAGCCGCAUAGGCUGCUUAAAUA [SEQ ID NO: 75]

[0999] In an example of the invention, the or each repeat of the array comprises or consists of a sequence that is at least 80, 90, 95, 96, 97, 98 or 99% identical to SEQ ID NO: 75 (or is identical to SEQ ID NO: 75).

[1000] The majority of spacers in the ICP1 CRISPR show 100% identity to sequences within an 18 kb island found in a subset of *V. cholerae* strains that include the classical strain O395 isolated in India in 1964, El Tor strain MJ-1236 isolated in Bangladesh in 1994, and several El Tor strains collected at the ICDDR,B between 2001-2011. The 18 kb island resembles the phage inducible chromosomal islands (PICIs) of Gram-positive bacteria, including the prototype *Staphylococcus aureus* pathogenicity islands (SaPIs). SaPIs are induced to excise, circularize and replicate following infection by certain phages. They use varied mechanisms to interfere with the phage reproduction cycle to enable their own promiscuous spread and this can protect the surrounding bacterial population from further phage predation. The organization of the *V. cholerae* 18 kb island targeted by the ICP1 CRISPR/Cas system is similar in length, base composition, and organization to that observed in the SaPIs subset of PICIs, with an integrase homologue at one end and a GC content lower than that of the host species (37% compared to 47.5%). The 18 kb element is therefore referred to as the *V. cholerae* PICI-like element (PLE).

V. cholerae PLE(wt)	AAT	TTA	AAT	AGG	GAA	GAT	AAG	CAA	AGG	GTT	GAC
	N	L	N	R	E	D	K	Q	R	V	D

[the nucleotide sequence= SEQ ID NO: 76 (The 32 bp protospacer sequence (SEQ ID NO: 77) is shaded in grey, the present disclosure includes a sequence that starts at the first T shaded grey and ends at the last C shaded grey; and the amino acid sequence = SEQ ID NO: 78]

[1001] Seed *et al* determined that the CR1 and CR2 arrays operate by recognition of a GA PAM sequence. Seed *et al* also found that the majority of spacers in the studied ICP1-related phage CRISPR arrays showed identity to *V. cholera* PLEs. The spacers are shown in the following Table 3; S1 in an array of the invention is, for example, selected from any one of these sequences. In an embodiment, S1 is selected from any one of the underlined sequences.

[1002] In an example, the array of the invention (or each array) is an engineered array comprising one, more or all of the underlined spacer sequences. The array spacers can comprise a non-naturally occurring arrangement as follows:-

[1003] For example, the array comprises a spacer of type 8a and/or 9a, and 0, 1, 2, 3, 4, 5 or 6 (but not 7) of types 1a-7a. For example, the array comprises a spacer of type 4b, and 0, 1 or 2, 3 (but not 3 or more) of 1b-3b. For example, the array comprises a spacer of type 8a and one, more or all of 9a, 4b, 1c, 3d, 1e and 3e. For example, the array comprises a spacer of type 9a and one, more or all of 8a, 4b, 1c, 3d, 1e and 3e. For example, the array comprises a spacer of type 4b and one, more or all of 8a, 9a, 1c, 3d, 1e and 3e. For example, the array comprises a spacer of type 1c and one, more or all of 8a, 9a, 4b, 3d, 1e and 3e. For example, the array comprises a spacer of type 3d and one, more or all of 8a, 9a, 4b, 1c, 1e and 3e. For example, the array comprises a spacer of type 1e and one, more or all of 8a, 9a, 4b, 1c, 3d and 3e. For example, the array comprises a spacer of type 3e and one, more or all of 8a, 9a, 4b, 1c, 3d and 1e.

[1004] In another non-naturally occurring arrangement, the vector comprises first and second arrays of the invention, wherein the arrays comprise at least two spacers selected from 1a to 1g (eg, at least two spacers selected from 8a, 9a, 4b, 1c, 3d, 1e and 3e) wherein the spacers are not spacers of the same ICP1 phage genome, eg, not all spacers of ICP1_ 2011_A, or ICP1_ 2006_E, or ICP1_ 2005_A or ICP1_ 2004_A (by reference to the spacers in the table above). Thus, in an embodiment:-

[1005] The first array comprises an ICP1_2011_A spacer sequence (eg, 8a and/or 9a), and the second array comprises a spacer sequence of ICP1_ 2006_E, ICP1_ 2005_A or ICP1_ 2004_A (eg, one or more spacers selected from 4b, 1c, 3d, 1e and 3e).

[1006] In an example, the vector comprises 1, 2, 3, 4, 5, 6 or all 7 spacer types selected from 8a, 9a, 4b, 1c, 3d, 1e and 3e. In an example, the vector comprises multiple copies of one or more of said selected types. In an example, the, some or each array in the vector comprises a first spacer (nearest the promoter of the array), wherein the first spacer is selected from 8a, 9a, 4b, 1c, 3d, 1e and 3e. Positioning in this way is advantageous as natural arrays use the first spacer most frequently.

[1007] Reference is made to Nucleic Acids Res. 2013 Oct;41(19):9033-48. doi: 10.1093/nar/gkt654. Epub 2013 Jul 30, "High-resolution definition of the Vibrio cholerae essential gene set with hidden Markov model-based analyses of transposon-insertion sequencing data", Chao MC *et al* (incorporated by reference), which discloses the coupling of high-density transposon mutagenesis to high-throughput DNA sequencing (transposon-insertion sequencing) enables simultaneous and genome-wide assessment of the contributions of individual loci to bacterial growth and survival. HMM results indicate that 128 genes are required for optimal growth of V. cholerae in LB. The target sequence of the invention can be a sequence of any one of these genes (which gene names and sequences are explicitly incorporated herein by reference for use in providing target sequences of the vectors of the present invention and for possible inclusion in the claims herein).

[1008] For example, insertion mutants in *vc0309* and *vc0753,* which had average reads of 5.6 and 4.7, respectively, were severely attenuated in growth. Likewise, *vc0237* and *vc1773* mutants were less fit than wild-type cells in an *in vitro* competition experiment. The list also includes a number of antitoxin genes from putative toxin/antitoxin addiction loci, including *vca0360, vca0477, vca0486* and *vca0488.* Such genes are presumed to be essential when associated with active toxins.

[1009] The authors found the essential *V cholerae* genes in Table 4. The authors identified more than 200 intergenic regions that appear to be essential.

[1010] Thus, in an example of the invention when the host cell is a *Vibrio cholerae* cell, the target sequence is a vc0631, vc2024, vc2626, vc2763-vc2767 or vc2768-vc2770 sequence. In an example of the invention when the host cell is a *Vibrio cholerae* cell, the target sequence is a *vc0309* and *vc0753, vc0237* and *vc1773, vca0360, vca0477, vca0486* or *vca0488* sequence.

[1011] Reference is made to Infect Immun. 2015 Sep;83(9):3381-95. doi: 10.1128/IAI.00411-15. Epub 2015 Jun 8, "A Genome-Wide Screen Reveals that the Vibrio cholerae Phosphoenolpyruvate Phosphotransferase System Modulates

Virulence Gene Expression", Wang Q *et al* (incorporated by reference). The authors used a transposon insertion site (TIS) sequencing-based strategy to identify new factors required for expression of tcpA, which encodes the major subunit of TCP, the organism's chief intestinal colonization factor. Besides identifying most of the genes known to modulate tcpA expression, the screen yielded ptsI and ptsH, which encode the enzyme I (EI) and Hpr components of the *V. cholerae* phosphoenolpyruvate phosphotransferase system (PTS). In addition to reduced expression of TcpA, strains lacking EI, Hpr, or the associated EIIA(Glc) protein produced less cholera toxin (CT) and had a diminished capacity to colonize the infant mouse intestine. The PTS modulates virulence gene expression by regulating expression of tcpPH and aphAB, which themselves control expression of toxT, the central activator of virulence gene expression.

[1012] Thus, in an example of the invention when the host cell is a *Vibrio cholerae* cell, the target sequence is a tcpA sequence or a tcpA modulator sequence (ie, a nucleotide sequence that modulates tcpA itself or *via* its expression product). For example, the sequence is a ptsI or ptsH sequence. In an example, the target sequence is sequence of the phosphoenolpyruvate phosphotransferase system (PTS), or a tcpPH, aphAB or toxT sequence. In an example the target sequence is a gene sequence encoding EIIA(Glc) protein.

[1013] Suitable target sequences for the present invention are also as shown in Table 5 - sequence of any one of the following (Pathogenicity genes are underlined).

[1014] In an embodiment, the cell is a *Vibrio* (eg, *cholera)* cell and the target sequence is a sequence if any of these genes.

[1015] Pathogenicity genes are shown in Table 6.

[1016] In an embodiment, the cell is a *Vibrio* (eg, *cholera)* cell and the target sequence is a sequence if any of these genes.

Genes from TCP and CTX pathogenicity islands

[1017] In an embodiment, the cell is a *Vibrio* (eg, *cholera)* cell and the target sequence is an ace, cep, ctxA, ctxB, orfU, zot, rstA, rstB, rstR, acfA, acfB, acfC, tagE, aldA, int, tagA, tagD, tcpA, tcpB, tcpC, tcpD, tcpE, tcpF, tcpH, tcpI, tcpJ, tcpP, tcpQ, tcpR, tcpS, tcpT or toxT sequence.

COMPARATIVE EXAMPLE 6: Specific Microbiota Bacterial Population Growth Inhibition By Harnessing Wild-Type Endogenous Cas

1. Material and methods

1. 1. Strains

[1018] The following strains were used in the course of this Example and Examples 7 and 8: *E. coli* MG1655, *E.coli* TOP10, *Streptococcus thermophilus* LMD-9 (ATCC BAA-491, Manassas, Virginia), *Streptococcus thermophilus* DSM 20617(T) (DSMZ, Braunschweig, Germany), *Lactococcus lactis* MG1363 and *Streptococcus mutans* Clarke 1924 DSM 20523 (DSMZ, Braunschweig, Germany).

[1019] During the course of media selection and testing of the genetic constructs different *Streptoccoci* strains were used. *Streptococcus thermophilus* LMD-9 (ATCC BAA-491) and *Escherichia coli* TOP10 were considered because of their compatible growth requirements. All strains were cultivated in Todd-Hewitt broth (TH) (T1438 Sigma-Aldrich), in aerobic conditions and at 37°C, unless elsewhere indicated. The strains were stored in 25% glycerol at -80°C.

1. 2. Differential growth media

[1020] All strains were grown on TH media at 37 °C for 20 hours. Selective media for S.thermophilus was TH media supplemented with 3 g l^{-1} of 2-phenylethanol (PEA). PEA was added to the media and autoclaved at 121°C for 15 minutes at 15 psi. Agar plates were prepared by adding 1.5% (wt/vol) agar to the corresponding media. When necessary for selection or plasmid maintenance 30 μg ml^{-1} kanamycin was used for both *S. thermophilus* strains and *E.coli,* and 500 μg ml^{-1} for *S.mutans.*

[1021] In some cases, depending on the strain and plasmid, a longer incubation, up to 48 hours, may be needed to see growth on media supplemented with PEA. In order to control for the viability of the organisms used, a control TH agar must be done in parallel.

1. 3. Cloning

[1022] *E. coli* (One Shot® ThermoFischer TOP10 Chemically Competent cells) was used in all subcloning procedures. PCR was carried out using Phusion polymerase. All PCR products were purified with Nucleospin Gel and PCR Clean-

up by Macherey-Nagel following the manufacturer's protocol. The purified fragments were digested with restriction enzyme DpnI in 1X FD buffer with 1μl enzyme in a total volume of 34 μl. The digested reaction was again purified with Nucleospin Gel and PCR Clean-up by Macherey-Nagel following the manufacturer's protocol. Gibson assembly was performed in 10 μl reactions following the manufacturer's protocol (NewEngland Biolab).

[1023] Plasmid DNA was prepared using Qiagen kits according to the manufacturer's instructions. Modifications for Gram-positive strains included growing bacteria in a medium supplemented with 0.5% glycine and lysozyme to facilitate cell lysis.

1. 4. Transformation

1. 4.1 Electro-competent *E.coli* cells and transformation

[1024] Commercially electrocompetent cells were used for cloning and the experiments (One Shot® ThermoFischer TOP10 Chemically Competent *E. coli).* Electroporation was done using standard settings: 1800 V, 25 μF and 200 Ω using an Electro Cell Manipulator (BTX Harvard Apparatus ECM630). Following the pulse, 1 ml LB-SOC media was added and the cells were incubated at 37°C for 1 hour. The transformed cells were plated in LB-agar containing 50 μg ml^{-1} of kanamycin.

1. 4.2 Preparation of electro-competent *S. thermophilus* cells

[1025] The electroporation protocol was modified from Somkuti and Steinberg, 1988. An overnight culture of *Streptococcus thermophilus* in TH Broth supplemented with 40 mM DL-threonine (T8375 Sigma-Aldrich) was diluted 100-fold in 5 ml of the same media and grown to an OD_{600} between 0.3 - 0.5 (approximately 2.5 hours after inoculation). The cells were collected by centrifugation at 10,000 x g for 10 min at 4°C and washed three times with 5 ml of ice cold wash buffer (0.5 M sucrose + 10% glycerol). After the cells were washed, they were suspended to an OD_{600} of 15-30 in electroporation buffer (0.5 M sucrose, 10% glycerol and 1mM $MgCl_2$). The cells in the electroporation buffer may be kept at 4°C until use (within one hour) or aliquot 50 μl in eppendorf tubes, freezing them in liquid nitrogen and stored at -80°C for later use.

1. 4.3 Electroporation *S. thermophilus* cells

[1026] 1 μl of purified plasmid DNA was added to 50 μl of the cell suspension and electroporation was carried out in 2mm-gap electroporation cuvettes pre-cooled. The electroporation setting were 2500 V, 25 μF and 200 Ω using an Electro Cell Manipulator (BTX Harvard Apparatus ECM630). Immediately after the electric pulse, 1 ml of TH broth was added to the cells and the suspension was kept on ice for 10 minutes, subsequently the cells were incubated for 3 h at 37°C. After allowing time for expression of the resistance gene the cells were plated onto TH-agar plates containing 30μg ml^{-1} of kanamycin. Depending on the construct, colonies were visible between 12 and 48 h of incubation at 37°C.

1. 5. Construction of XylS plasmid

[1027] All the plasmids used in this work were based on pBAV1K-T5, which is a broad-host range expression vector derived from the a cryptic plasmid pWV01 from *Streptococcus cremoris* (Bryksin & Matsumura, 2010), the backbone was amplified using that contain overhangs for assembly with the other fragments using Gibson's method.

[1028] The xylose inducible system was constructed by cloning the promoter gyrA in front of the XylR repressor (Figure 1). The XylR repressor was amplified from *Bacillus Subtilis* strain SCK6 (Zhang *et al.* 2011) with the a reverse primer that includes an overhang for Gibson assembly and a forward primer, that is an ultramer used to introduce the gyrA promoter (Xie *et al.* 2013) and the corresponding overhang for assembly into pBAV1KT5 backbone. The resulting fragment was flanked by an mCherry amplified from pCL002 (unpublished work) with an ultramer that include PIdha+PxylA hybrid promoter (Xie *et al.* 2013). The three resulting PCR products were assembled in a Gibson Master Mix® (NewEngland Biolab) according to manufacturer's instructions. The product was finally transformed in *E. coli* TOP10 electrocompetent cells. See Figure 1.

1. *6. Design and construction of CRISPR array plasmid*

[1029] *Streptococcus thermophilus* has 4 distinct CRISPR systems (Sapranauskas, *et al.* 2011), for this work the type II CRISPR1 (ST1-CRISPR) system was chosen. The design of the target sequence was based on the available genome sequence of LMD-9 (GenBank: CP000419.1). The ST1-CRISPR array was designed to contain only the CRISPR array repeats and spacers under a xylose inducible promoter (Xie *et al.* 2013), followed by the corresponding tracrRNA under

a strong constitutive promoter for *Streptococci* species (Sorg *et al.* 2014) (Figure 2, SEQ ID Nos:).

[1030] The tracrRNA plays a role in the maturation of crRNA and it is processed by *S. thermophilus* endogenous RNase III, forming a complex with crRNA. This complex acts as a guide for the endonuclease ST1-Cas9 (Horvath & Barrangou, 2010). After transcription of the synthetic array from the xylose inducible promoter, the endogenous Cas9 and RNAses will process it into a functional gRNA. The gRNA/Cas9 complex will cause a double stranded break at the target location.

[1031] The design of the array used 2 specific target sequences high on GC content and a reduced portion of the tracrRNA (ie, a less than complete tracrRNA sequence), which has been suggested not to be necessary for proper maturation of crRNA (Horvath & Barrangou, 2010).

[1032] The 2 targets were an essential gene (DNA polymerase III subunit alpha) and an antibiotic resistance gene (tetA-like gene) (SEQ ID NOs:).

[1033] Primers were used to amplify pBAV1KT5-XylR-PIdhA backbone. The CRISPR array gBlock and the backbone with overhangs were assembled in a Gibson Master Mix® according to manufacturer's instructions (NewEngland Biolabs). The product was finally transformed in *E. coli* TOP10 electrocompetent cells.

1. 7. Characterization of Xylose inducible system in Streptoccocus thermophilus LMD-9

[1034] Overnight stationary-phase cultures were diluted 1:100 into TH broth with corresponding antibiotic. Mid-log cells were induced with different concentration of D-(+)-xylose (0, 0.001, 0.01, 0.1, 0.5 and 1 % wt/vol) and the cell cultures were measured either directly in medium to assess the extent of autofluorescence of the media, on the cell suspension or the suspension buffer (PBS buffer). 20μl samples of the cell cultures were diluted 1/10 on PBS buffer, on 96-well plates with flat bottoms. Fluorescence of cell suspensions or media was read on a plate reader. mCherry fluorescence was measured using an excitation wavelength of 558nm and emission at 612nm. Absorbance of the resuspended cells was measured at OD 600 nm. A minimum of three independent biological replicates was done for each experiment.

1.8. Activation of CRISPR array in S. thermophilus

[1035] *S. thermophilus* LMD-9 and *E.coli* TOP10 both with the plasmid containing the CRISPR array targeting the *DNA polymerase III* and *tetA* of *S. thermophilus* were grown overnight in 3 ml cultures supplemented with 30 μg ml^{-1} of kanamycin for plasmid maintenance. The next day 96 well deep well plates were inoculated with 500 μl of 1/100 of overnight culture in fresh TH media, supplemented with 30 μg ml^{-1} kanamycin. Mid-log cell cultures were induced with 1% xylose. The killing effect was tested on *S.thermophilus* and *E.coli* alone. For each strain and condition tested a negative control was kept without xylose. The cells were grown till ~OD 0.5 and next 10-fold serially diluted in TH media and using a 96-well replicator (Mettler Toledo Liquidator™ 96) 5μL volume drops were spotted on TH agar and TH agar supplemented with g l^{-1} PEA plates. The plates were incubated for 24H at 37°C and the colony forming units (CFU) were calculated from triplicate measurements.

2. Results

2.1 Growth condition and selective media

[1036] We first set out to establish the bacterial strains and cultivation protocol that would support growth for all strains we planned to use for the co-cultivation experiments. We used *S. thermophilus* strain LMD-9 which was able to support a similar growth as *E.coli* in TH broth at 37°C (Figure 3).

[1037] Distinguishing the different bacteria from a mixed culture is important in order to determine cell number of the different species. With MacConkey agar is possible to selectively grow *E.coli,* however there is no specific media for selective growth of *S.thermophilus.* PEA agar is a selective medium that is used for the isolation of gram-positive *(S.thermophilus)* from gram-negative *(E.coli).* Additionally, we found that different concentrations of PEA partially inhibit the growth of other gram positives, which allow for selection between the other gram-positive bacteria used in this work (figure 4). 3g l^{-1} of PEA proved to selectively grow *S. thermophilus* LMD-9 while limiting growth of *E. coli.*

2.2 Design and validation of inducible system

[1038] An induction system for Streptococcus species was previously developed based on the *Bacillus megaterium* xylose operon (Figure 5) by creating a heterologous xylose induction cassette (Xyl-S). The *xylR* and *xylA* promoters were replaced with *S. mutans'* constitutively expressed *gyrA* and *Idh* promoters respectively. This expression cassette for Streptococcus species showed differences in sensitivity and expression levels between different species, however

the system was not tested in *S. thermophilus* (Xie *et al.* 2013). Therefore we first set out to validate the xylose induction cassette in *S. thermophilus.*

[1039] An alternative version of the induction cassette was constructed by only replacing the *xylR* promoter with the *S. mutans' gyrA* promoter but left the endogenous *B. megaterium xylA* promoter intact. During the design of the xylose inducible system we considered both versions of the inducible promoter, the natural P_{XylA} promoter found in *Bacillus megaterium* and a hybrid promoter of the highly conserved promoter P_{ldha} fused with the repressor binding sites of P_{XylA} promoter (Figure 5). Only a few Streptococcus species have been reported to metabolize xylose, and thus the presence of a regulatory machinery to recognize the xylA promoter in the other Streptococcus species is not likely. Therefore we constructed both xylose induction systems but only tested the inducibility of mCherry with the $P_{ldha+xylA}$ system.

[1040] In order to determine mCherry inducible expression by xylose, mid-log cultures of cells with the plasmid (pBAV1KT5-XylR-mCherry-$P_{ldha+XylA}$) were induced with different concentrations of xylose. Six hours after the induction we measured mCherry fluorescence in the cultures, where we observed substantially higher overall expression levels in cells carrying the plasmid (figure 6). It is worth noticing that the system showed a substantial level of basal expression even in the cultures where xylose was not added. This means that the system is 'leaky' and in context of the kill-array this can lead to cell death even before the system is induced with xylose. However, in the subsequent course of this study we used both versions of the plasmid (pBAV1KT5-XylR-mCherry-$P_{ldha+XylA}$ and pBAV1KT5-XylR-mCherry-P_{xylA}).

2. 3 Design of CRISPR/CAS9 array

[1041] In order to determine if the genomic targeting spacers in a CRISPR array can cause death in *S.thermophilus* LMD-9, we inserted the CRISPR array we designed into the two xylose inducible systems previously constructed (pBAV1KT5-XylR-mCherry-$P_{ldha+XylA}$ and pBAV1KT5-XylR-mCherry-P_{xylA}). In these plasmids we replaced mCherry with the gBlock containing the CRISPR array (Figure 7). The variant with the $P_{ldha+xylA}$ promoter was expected to be stronger and have a higher basal activity than the P_{xylA} (Xie *et al.* 2013).

2. 4 Inhibition of Bacterial Population Growth Using Endogenous Cas9

[1042] After we constructed the plasmids in *E.coli,* we transformed the plasmids into *S. thermophilus.* This would allow us to determine if we could cause cell death of a specific bacterial species. Interestingly, bacterial host population size (indicated by growing bacteria and counting colony numbers on agar plates) in *S. thermophilus* exposed to the plasmid containing the strong $P_{ldh+xylA}$ hybrid promoter was 10-fold less when compared to *S. thermophilus* exposed to the plasmid containing the weak, normal P_{XylA} promoter (figure 8; 52 colonies with the strong array expression versus 556 colonies with weak array expression, 10.7-fold difference), the 2 strains having been transformed in parallel using the same batch of electrocompetent *S. thermophilus* cells. This suggests to us that the plasmid carrying the CRISPR array targeting *S. thermophilus* genes is able to kill the cells using the endogenous Cas nuclease and RNase III, thereby inhibiting population growth by 10-fold.

[1043] We expect that weak array expression in host cells transformed by the plasmid comprising the P_{XylA} promoter led to a degree of cell killing, albeit much less than with the strong promoter plasmid. We expect that population growth inhibition that is greater than the observed 10-fold inhibition would be determined if a comparison of the activity of strong array expression was made with *S thermophilus* that is not exposed to any array-encoding plasmid (such as bacteria directly isolated from gut microbiota). Thus, we believe that array (or single guide RNA) expression in host cells for harnessing endogenous Cas nuclease will be useful for providing effective growth inhibition of target host cells in environmental, medical and other settings mentioned herein. Co-administration of antibiotic may also be useful to enhance the growth inhibition, particularly when one or more antibiotic resistance genes are targeted.

3. Discussion and outlook

[1044] In this study we set out to design a CRISPR-array to specifically kill *S. thermophilus* using the endogenous Cas9 system. In order to gain control over the killing signal we sought to apply an inducible system that can be applied in *S. thermophilus.* The xylose inducible XylR system from *B. megaterium* was previously applied in *S. mutans* (Xie, 2013) but not in *S. thermophilus.* In this study we demonstrated the functionality of the xylR induction system using the designed XylR-mCherry-PIdha circuit in *S. thermophilus.* We found 0.1 % wt/vol is sufficient to fully induce the XylR system in *S. thermophilus* (Figure 6).

[1045] In order to observe abundance when co-culturing *S. thermophilus* and *E. coli* we established that supplementation of the culture media with 3 g l^{-1} of PEA, allows for the selective growth of *S. thermophilus* while limiting the growth of *E. coli* (Figure 4).

[1046] A ST1-CRISPR array, targeting the DNA polymerase III subunit alpha and a tetA like gene in the *S. thermophilus* LMD-9 genome, was placed under the xylose inducible promoter (Xie et al. 2013). Targeting these regions should lead

to a double strand break and thus limit *S. thermophilus* viability (Figure 9). Since the engineered array was designed to target *S. thermophilus* genome using the endogenous CRISPR/Cas machinery to process the encoded CRISPR array, the array is expected to have no influence on growth of unrelated strains such as *E. coli,* even similar targets could be found on its genome. This was successfully tested in a mixed bacterial population (simulating aspects of a human microbiota) as discussed in Example 8.

[1047] The demonstration of the invention's ability to inhibit host cell growth on a surface is important and desirable in embodiments where the invention is for treating or preventing diseases or conditions mediated or caused by microbiota as disclosed herein in a human or animal subject. Such microbiota are typically in contact with tissue of the subject (eg, gut, oral cavity, lung, armpit, ocular, vaginal, anal, ear, nose or throat tissue) and thus we believe that the demonstration of activity to inhibit growth of a microbiota bacterial species (exemplified by Streptococcus) on a surface supports this utility.

COMPARATIVE EXAMPLE 7: Specific Microbiota Bacterial Population Growth Inhibition In Different Strains

[1048] Example 6 demonstrated specific growth inhibition of *Streptococcus thermophilus* LMD-9. Here we demonstrate growth inhibition can also be obtained in a second strain: *Streptococcus thermophilus* DSM 20617. Methods described in Example 6 were, therefore, applied to the latter strain (except that selective media for S.thermophilus DSM 20617was TH media supplemented with 2.5g l^{-1} of 2-phenylethanol (PEA)).

[1049] *Streptococcus thermophilus* DSM 20617 transformed with the CRISPR array plasmids were incubated for recovery in liquid media for a period of 3 hours at 37°C that would allow for expression of kanamycin resistance. After a recovery period, cells were plated in different selection media in presence of 1% xylose in order to induce cell death, and without xylose as a control (figure 10). It is evident that; (1) by xylose induction the growth of *S. thermophilus* can be inhibited (around 10-fold for the 'strong' promoter plasmid versus control), (2) the 'strong' system (pBAV1KT5-XylR-CRISPR-P_{ldhA}) results in more growth reduction than the 'weak' system (pBAV1KT5- XylR-CRISPR-P_{xylA}).

EXAMPLE 8: Selective Bacterial Population Growth Inhibition In a Mixed Consortium of Different Microbiota Species

[1050] We next demonstrated selective growth inhibition of a specific bacterial species in a mixed population of three species. We selected species found in gut microbiota of humans and animals (*S thermophilus* DSM 20617(T), *Lactobacillus lactis* and *E coli).* We included two gram-positive species (the *S thermophilus* and *L lactis)* to see if this would affect the ability for selective killing of the former species; furthermore to increase difficulty (and to more closely simulate situations in microbiota) *L lactis* was chosen as this is a phylogenetically-related species to *S thermophilus* (as indicated by high 16s ribosomal RNA sequence identity between the two species). The *S thermophilus* and *L lactis* are bothe *Firmicutes.* Furthermore, to simulate microbiota, a human commensal gut species *(E coli)* was included.

1. Materials & Methods

[1051] Methods as set out in Example 6 were used strain (except that selective media was TH media supplemented with 2.5g l^{-1} of 2-phenylethanol (PEA)).

1.1 Preparation of electro-competent L.lactis cells

[1052] Overnight cultures of *L. lactis* in TH media supplemented with 0.5 M sucrose and 1% glycine were diluted 100-fold in 5 ml of the same media and grown at 30°C to an OD_{600} between 0.2 - 0.7 (approximately 2 hours after inoculation). The cells were collected at 7000 x g for 5 min at 4°C and washed three times with 5 ml of ice cold wash buffer (0.5 M sucrose + 10% glycerol). After the cells were washed, they were suspended to an OD_{600} of 15-30 in electroporation buffer (0.5 M sucrose, 10% glycerol and 1mM $MgCl_2$). The cells in the electroporation buffer were kept at 4°C until use (within one hour) or aliquot 50 μl in eppendorf tubes, freezing them in liquid nitrogen and stored at -80°C for later use.

[1053] Electroporation conditions for all species were as described in Example 6.

1.2 Activation of CRISPR array: Consortium Experiments.

[1054] *S. thermophilus* DSM 20617, *L. lactis* MG1363 and *E.coli* TOP10 were genetically transformed with the plasmid containing the CRISPR array targeting the DNA polymerase III and tetA of *S. thermophilus.* After transformation all cells were grown alone and in co-culture for 3 hours at 37°C allowing for recovery to develop the antibiotic resistance encoded in the plasmid. We decided to use transformation efficiency as a read out of CRISPR-encoded growth inhibition. Therefore, after allowing the cells for recovery the cultures were plated in TH media, TH supplemented with PEA and MacConkey agar all supplemented with Kanamycin, and induced by 1% xylose.

2. Results

2.0 Phylogenetic distance between L. lactis, E. coli and S. thermophilus

[1055] The calculated sequence similarity in the 16S rrNA-encoding DNA sequence of the *S. thermophilus* and *L. lactis* was determined as 83.3%. The following 16S sequences were used: *E. coli:* AB030918.1, *S. thermophilus:* AY188354.1, *L. lactis:* AB030918. The sequences were aligned with needle (http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html) with the following parameters: - gapopen 10.0 -gapextend 0.5 -endopen 10.0 -endextend 0.5 -aformat3 pair -snucleotide1 -snucleotide2. Figure 11 shows the maximum-likelihood phylogenetic tree of 16S sequences from *S. thermophilus, L. lactis* and *E. coli.*

2.1 Growth condition and selective media

[1056] *S. thermophilus* and *L. lactis* are commonly used in combination in many fermented foods and yoghurt. We chose these strains since they are commonly known to be gut microbes that form an intimate association with the host and previous characterizations of the 16S ribosomal RNA region of *S. thermophilus* and *L. lactis* have shown that these organisms are phylogenetically closely related (Ludwig *et al.,* 1995). In parallel we also evaluated the growth of *E.coli* for our mixed population co-culture experiments, since this organism is also commonly found in gut microbe communities. We first set out to establish the bacterial strains and cultivation protocol that would support growth for all strains we planned to use for the co-cultivation experiments. We found that all strains were able to support growth in TH broth at 37°C (Figure 3).

[1057] Distinguishing the different bacteria from a mixed culture is important in order to determine cell number of the different species. With MacConkey agar is possible to selectively grow *E.coli,* however there is no specific media for selective growth of *S.thermophilus.* PEA agar is a selective medium that is used for the isolation of gram-positive *(S.thermophilus)* from gram-negative *(E.coli).* Additionally, different concentrations of PEA partially inhibit the growth of the different grams positive species and strains, which allow for selection between the other gram-positive bacteria used in this work. Using 2.5 g l^{-1} of PEA proved to selectively grow *S. thermophilus* while limiting growth of *L. lactis* and *E. coli.*

[1058] All strains were transformed with a plasmid that used the vector backbone of pBAV1KT5 that has a kanamycin selection marker; we found that using media supplemented with 30 ug ml^{-1} of kanamycin was enough to grow the cells while keeping the plasmid.

2. 3 Transformation & Selective Growth Inhibition in a Mixed Population

[1059] We transformed *S. thermophilus, L. lactis* and *E. coli with* plasmid containing the CRISPR array and cultured them in a consortium of all the bacterial species combined in equal parts, which would allow us to determine if we could cause cell death specifically in *S.thermophilus.* We transformed all the species with either the pBAV1KT5-XylR-CRISPR-P_{XylA} or pBAV1KT5-XylR-CRISPR-$P_{ldha+XylA}$ plasmid.

[1060] Figure 12 shows the selective *S thermophilus* growth inhibition in a co-culture of *E. coli, L. lactis and S. thermophiles* harboring either the pBAV1KT5-XylR-CRISPR-P_{xylA} or the pBAV1KT5-XylR-CRISPR-$P_{ldhA+XylA}$ plasmid. No growth difference is observed between *E. coli* harboring the pBAV1KT5-XylR-CRISPR-P_{xylA} or the pBAV1KT5-XylR-CRISPR-$P_{ldhA+XylA}$ plasmid (middle column). However, *S. thermophiles* (selectively grown on TH agar supplemented with 2.5 gl^{-1} PEA, last column) shows a decrease in transformation efficiency between the pBAV1KT5-XylR-CRISPR-P_{xylA} (strong) or the pBAV1KT5-XylR-CRISPR-$P_{ldhA+XylA}$ (weak) plasmid as we expected. We thus demonstrated a selective growth inhibition of the target S thermophilus sub-population in the mixed population of cells.

References

[1061]

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Patrick Boyaval, Moineau, S., ... Horvath, P. (2007). CRISPRProvides Acquired Resistance Against Viruses in Prokaryotes. Science, 315(March), 1709-1712. http://doi.org/10.1126/science.1138140

Bryksin, A. V, & Matsumura, I. (2010). Rational design of a plasmid origin that replicates efficiently in both gram-positive and gram-negative bacteria. PloS One, 5(10), e13244. http://doi.org/10.1371/journal.pone.0013244

Chan CTY, Lee JW, Cameron DE, Bashor CJ, Collins JJ: "Deadman" and "Passcode" microbial kill switches for bacterial containment. Nat Chem Biol 2015, 12(December):1-7.

Horvath, P., Romero, D. A., Coûté-Monvoisin, A.-C., Richards, M., Deveau, H., Moineau, S., ... Barrangou, R. (2008). Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. Journal of Bacteriology, 190(4), 1401-12. http://doi.org/10.1128/JB.01415-07

Ludwig, E. S., Klipper, R., Magrum L., Wose C., & Stackebrandt, E. (1985). The phylogenetic position of Streptococcus and Enterococcus. Journul of Gencwl Microhiologj., 131, 543-55 1.

Mercenier, A. (1990). Molecular genetics of Streptococcus thermophilus. FEMS Microbiology Letters, 87(1-2), 61-77. http://doi.org/10.1016/0378-1097(90)90697-0

Samaržija, D., Antunac, N., & Havranek, J. (2001). Taxonomy, physiology and growth of Lactococcus lactis: a review. Mljekarstvo, 51(1), 35-48. Retrieved from http://hrcak.srce.hr/index.php?show=clanak_download&id_clanakJezik=2828

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., & Siksnys, V. (2011). The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic Acids Research, 39(21), 9275-9282. http://doi.org/10.1093/nar/gkr606

Somkuti, G. A., & Steinberg, D. H. (1988). Genetic transformation of Streptococcus thermophilus by electroporation. Biochimie, 70(4), 579-585. http://doi.org/10.1016/0300-9084(88)90095-8

Sorg, R. A., Kuipers, O. P., & Veening, J.-W. (2014). Gene expression platform for synthetic biology in the human pathogen Streptococcus pneumoniae. ACS Synthetic Biology, 4(3), 228-239. http://doi.org/10.1021/sb500229s

Suvorov, a. (1988). Transformation of group A streptococci by electroporation. FEMS Microbiology Letters, 56(1), 95-99. http://doi.org/10.1016/0378-1097(88)90133-4

Xie, Z., Qi, F., & Merritt, J. (2013). Development of a tunable wide-range gene induction system useful for the study of streptococcal toxin-antitoxin systems. Applied and Environmental Microbiology, 79(20), 6375-84. http://doi.org/10.1128/AEM.02320-13

Zhang, X. Z., & Zhang, Y. H. P. (2011). Simple, fast and high-efficiency transformation system for directed evolution of cellulase in Bacillus subtilis. Microbial Biotechnology, 4(1), 98-105. http://doi.org/10.1111/j.1751-7915.2010.00230.x

[1062] TABLES

Table 1: Repeat Sequences of SRBs For Use in the Invention

Each of R1 and R1' can be selected from these repeat consensus sequences, eg, when the system is an crude oil, natural gas or water recovery, processing or storage equipment.

BACTERIUM	CRISPR_ID	START POSITION	END POSITION	NUMBER OF SPACERS	DR (REPEAT) CONSENSUS	SEQ ID NO:	NOTES
Desulphovibrio desulphuricans ND132	NC_ 016803_1	2325998	2326074	1	CCTGGCCTGCCCCAAGTGCAAGG	50	
Desulphovibrio desulphuricans ND132	NC_ 016803_4	3491653	3498191	98	GTCGCCCCCCACGCGGGGGCGTGGATTGAAAC	51	1
Desulphovibrio vulgaris subsp. vulgaris str. Hildenborough	NC_ 005863_3	175898	177711	28	GTCGCCCCCCACGCGGGGGCGTGGATTGAAAC	51	1
Desulphobulbus propionicus DSM 2032	NC_ 014972_1	1701541	1707620	92	GTCGCCCCCCACGCGGGGGCGTGGATTGAAAC	51	1
Desulphovibrio vulgaris RCH1	NC_ 017311_1	170455	172268	28	GTCGCCCCCCACGCGGGGGCGTGGATTGAAAC	51	1
Desulphovibrio desulphuricans subsp. desulphuricans str. (ATCC 27774)	NC_ 011883_1	676246	676547	3	TGGAGCGGGAAACGGGATTTGAACCCGC	52	
Desulphovibrio desulphuricans subsp. desulphuricans str. (ATCC 27774)	NC_ 011883_2	1083779	1085579	29	GTGTTCCCCACGGGCGTGGGGATGAACCG	53	
Desulphovibrio gigas DSM 1382 (ATCC 19364)	NC_ 022444_2	430661	430743	1	AACCTTTCTGCAAAAAGGTTTCCCC	54	2
Desulphovibrio gigas DSM 1382 (ATCC 19364)	NC_ 022444_3	915564	915638	1	CCGCTGGATCCGGCTGCAGCGCC	55	
Desulphovibrio gigas DSM 1382 (ATCC 19364)	NC_ 022444_4	1994976	1995063	1	GTTCACTGCCGCATAGGCAGCTCAGAAA	56	
Desulphovibrio gigas DSM 1382 (ATCC 19364)	NC_ 022444_5	2555284	2555600	4	CACCCGACTATTGAAGTCGGGCCTCATTGAAG	57	
Desulphurispirillum indicum S5					AACCTTTCTGCAAAAAGGTTTCCCC	54	2
Desulphovibrio hydrothermal is	NC_ 022579_1	10819	11067	3	GTCAAAACCCATACCGATATGGATACCTCTTTTGAG	58	
Desulphovibrio hydrothermal is	NC_ 022579_2	24430	24678	3	GTCAAAACCCATACCGATATGGATACCTCTTTTGAG	59	

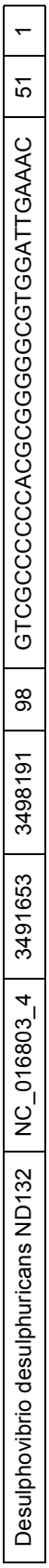
(continued)

Each of R1 and R1' can be selected from these repeat consensus sequences, eg, when the system is an crude oil, natural gas or water recovery, processing or storage equipment.							
BACTERIUM	**CRISPR_ID**	**START POSITION**	**END POSITION**	**NUMBER OF SPACERS**	**DR (REPEAT) CONSENSUS**	**SEQ ID NO:**	**NOTES**
Desulphovibrio hydrothermal is	NC_ 022579_3	36027	36275	3	GTCAAAACCCATACCGATATGGATACCTCTTTTGAG	60	
Desulphovibrio hydrothermal is	NC_ 022579_4	118127	118736	8	GTCAAAACCCATACCGATATGGATACCTCTTTTGAG	61	
Desulphovibrio hydrothermalis	NC_ 022579_5	2366564	2366737	2	CTCAAAAGAGGTATCCATATCGGTATGGGTTTTGAC	62	
Desulphovibrio hydrothermal is	NC_ 022579_6	2574826	2575933	18	GTTCACTGCCGGATAGGCAGCTTAGAAA	63	
Desulphovibrio magneticus RS-1	NC_ 012796_1	1589785	1591828	30	GTCGCCCCCTGCGCGGGGGCGTGGATTGAAAC	64	
Desulphovibrio magneticus RS-1	NC_ 012796_3	4725356	4726585	20	TTTTCTGAGCTGCCTATGCGGCAGTGAAC	65	
Desulphovibrio vulgaris str. 'Miyazaki F'	NC_ 011769_1	241933	242082	1	CATCGACGACGAACCCGGGCACCGCCTGATGGTCCACGCCGTCATG	66	
Desulphovibrio vulgaris str. 'Miyazaki F'	NC_ 011769_3	2444693	2448088	51	GTCGCCCCTCACGCGGGGGCGTGGATAGAAAC	67	
Desulphovibrio vulgaris subsp. vulgaris DP4	NC_ 008741_1	29677	32622	44	GTTTCAATCCACGCCCCCGCACGGGGGGCGAC	68	
Desulphurispirillum indicum S5	NC_ 014836_1	994780	997087	38	TTTCTGAGCTGCCTATGCGGCAGTGAAC	69	3
Desulphurispirillum indicum S5	NC_ 014836_2	1123444	1127359	54	GACCGAAGACCTGTCGGAAACGACGGGGATTGAGAC	70	
Desulphovibrio gigas DSM 1382 (ATCC 19364)					TTTCTGAGCTGCCTATGCGGCAGTGAAC	69	3
Desulphovibrio gigas DSM 1382 (ATCC 19364)	NC_ 022444_4	1994976	1995063	2	GTTCACTGCCGCATAGGCAGCTCAGAAA	70	

(continued)

Each of R1 and R1' can be selected from these repeat consensus sequences, eg, when the system is an crude oil, natural gas or water recovery, processing or storage equipment.							
BACTERIUM	**CRISPR_ID**	**START POSITION**	**END POSITION**	**NUMBER OF SPACERS**	**DR (REPEAT) CONSENSUS**	**SEQ ID NO:**	**NOTES**
Desulphurivibrio alkaliphilus AHT2	NC_ 014216_2	1780099	1782570	40	CGGTTCATCCCCGCGAGTGCGGGGAACAT	71	
Desulphurivibrio alkaliphilus AHT2	NC_ 014216_3	1785014	1791956	115	TTTCTGAGCTGCCTGTGCGGCAGTGAAC	72	
Desulphurobacterium thermolithotrophum DSM 11699	NC_ 015185_1	267992	268349	5	GTTTTATCTGAACGTAGTGGGATATAAAG	73	
Desulphovibrio desulphuricans G20	NC_ 007519_1	885036	886223	19	CGGTTCATCCCCGCGGGTGCGGGGAACAC	74	
Information from CRISPRs Database (www.crispr.u-psud.fr) 1=Repeat sequence (SEQ ID NO: 51) is common across these bacteria; 2=Repeat sequence (SEQ ID NO: 54) is common across these bacteria; 3=Repeat sequence (SEQ ID NO: 69) is common across these bacteria.							

[1063] The entries are read as illustrated by the following example


Desulphovibrio desulphuricans ND132
NC_016803_4
3491653
3498191
98
GTCGCCCCCCACGCGGGGGCGTGGATTGAAAC
51
1

[1064] A CRISPR array is found in *Desulphovibrio desulphuricans* ND132 starting at position 3491653 and ending at position 3498191, wherein the array has 98 spacer sequences, each flanked by repeats, where the repeats each have the sequence of SEQ ID NO: 51. Such a repeat is also found in an array of the other bacteria under note number 1 (last column in the table).

TABLE 2:

BACTEROIDES REPEATS		
SEQ ID NO:	***SPECIES/STRAIN***	***REPEA T SEQUENCE***
107	1. *Bacteroides fragilis NCTC 9343* *2. Bacteroides fragilis 638R*	*GTTGTGATTTGCTTTCAAATTAGTATCTΓΓGAACCATTGGAAACAGC*
108	3. *Bacteroides fragilis NCTC 9343* 4. *Bacteroides fragilis YCH46*	*ATTTCAATTCCATAAGGTACAATTAATAC*
109	5. *Bacteroides helcogenes P 36-108*	*GTTTCAATCCACACACCCGTATAGGGTGTGAC*
110	6. *Bacteroides sp. CF50*	*ACTGTTTCTGATATGTCAAAGATAAAATTTTGAAAGCAAATCACAA C*
111	7. *Bacteroides thetaiotaomicron VPI-5482*	*GAAAAAATACAGTTTCGCTCTCA*
PREVOTELLA REPEATS		
SEQ ID NO:	***SPECIES/STRAIN***	***REPEAT SEQUENCE***
112	8. *Prevotella dentalis DSM 3688*	*GTCGCGTCTCACGTAGGCGCGTGGATTGAAAC*
113	9. *Prevotella denticola F0289*	*ATTGTGCTTGCTACTGCAAAGATACACATTTTGAAGCAATTCACAAC*
114	10. *Prevotella denticola F0289*	*CTCAATGAGTATCTTCCATTAAAACAAGGATTAAGAC*
115	11. *Prevotella intermedia 17*	*GTTGTTTTTAACCTTGCAAACAGCAGGCAGATACAAC*
116	12. *Prevotella intermedia 17*	*GTTGTATTTGCCAATGCAAAGATACTAATTTTAAAGCTAATCACAAC*
117	13. *Prevotella ruminicola 23*	*GTTGTATATCATTCCTTTCCTACATCAAACCACAAC*

Table 3: Underlined = CRISPR spacers that have 100% identity to sequences within the *V. cholerae* PLE

PHAGE SOURCE	ARRAY	SPACER	SPACER SEQUENCE	SEQ ID NO:
ICP1_2011_A	CR1	1a	CATTGCAACTATGCAAAATGATGAAGCTAAAA	79
		2a	TGTTAGAGTCGGTAGTATCTGGATGATCGATA	80
		3a	TTATGTATTGACCCCGACACGCCCCCCGACTG	81
		4a	TTACAGACGACCTAACTCTTCAGTACCATGAT	82
		5a	TACATAAGCTGCAACACGGTGTTCGTTTAAGT	83
		6a	AAAATACGCCTTTTTCCCTTCATCGTTTAAAG	84
		7a	ACCAACAAATCCCATAAACTGATAACCACGTT	85
		8a	GTCAACCCTTTGCTTATCTTCCCTATTTAAAT	86
		9a	TGTTAACCACCGCTTGAAATAATCATGATGCA	87

(continued)

PHAGE SOURCE	ARRAY	SPACER	SPACER SEQUENCE	SEQ ID NO:
ICP1_ 2006_E	CR1	1b	TGTGTCTATACTCAACCAATTTAAGCGCCGCA	88
		2b	CTACTCTCCCCAATATTAGCCATTCCTAATTC	89
		3b	GTCACCTTACCGTAAGACAGGCAGTAAAATTA	90
		4b	AAACTAGTGGACGTAATGCAGTATTCACGGTT	91
	CR2	1c	ATCCACACTACAAATAGAACACTCAACCGTGA	92
ICP1_2005_A	CR1	1d	TGTGTCTATACTCAACCAATTTAAGCGCCGCA	93
		2d	CTACTCTCCCCAATATTAGCCATTCCTAATTC	94
		3d	AAACTAGTGGACGTAATGCAGTATTCACGGTT	95
		4d	ATAATCGTTTTGAGTCTCACCAGCTTTTAGGC	96
	CR2	1e	ATCCACACTACAAATAGAACACTCAACCGTGA	97
		2e	TATTGATTGGTCTCTAACCTTGGGATGATTAA	98
		3e	TTCACGGGTAGCAACAGGGTAATAAACCAATA	99
ICP1_2004_A	CR1	1f	CATTGCAACTATGCAAAATGATGAAGCTAAAA	100
		2f	TGTTAGAGTCGGTAGTATCTGGATGATCGATA	101
		3f	TAGAAGAGTAATAGGAGCTACTGCAAACTTGT	102
		4f	TAACTATGTGTGGTTTATATTTTGTGTGCAAG	103
		5f	TTTTGAAACTATTGACAGAAGGTTGGGAACCT	104
		6f	TTGAGGTTGAACCTCTTCCGGTTCCTCTTCTG	105
	CR2	1g	GTGTATTGCTTGCAGTGGGTTACACACAAGAA	106

Table 4

Essential homolog in *V. cholerae*	**Neutral homolog in *V. cholerae***	**Homolog in *E. coli***	**Essential In *E. coli?***	**Annotated function**
vc0631	*vc0465*	*tyrS*	Yes	Tyrosyl tRNA synthetase
vc2024	*vc0093*	*plsB*	Yes	Glycerol 3 phosphate acyltransferase
vc2626		*dam*	No	Adenine Methyltransferase
vc2763-vc2767		*atpCDGAH*	No	F1 ATP synthase ($\varepsilon,\beta,\gamma,\alpha,\delta$) subunits
vc2768-vc2770		*atpFEB*	No	F0 ATP synthase (B,C,A) subunits

Table 5

Gene	**Feature**
dnaE	DNA polymerase III holoenzyme alpha subunit
recA	recombinase A
ctxB	cholera toxin B
mdh	malate dehydrogenase
gyrB	DNA gyrase subunit B
tcpA	toxin co-regulated pilin A

(continued)

Gene	Feature
ctxA	cholera toxin A subunit
rpoA	RNA polymerase alpha subunit
tcpB	toxin co-regulated pilus biosynthesis protein B
asd	aspartate-semialdehyde dehydrogenase

Table 6

Gene	Feature
ctxB	cholera toxin B
tcpA	toxin co-regulated pilin A
ctxA	cholera toxin A subunit toxin co-regulated pilus biosynthesis
tcpB	protein B
wbet	ogawa specific antigen
hlyA	hemolysin A
hapR	hemagglutinin/protease regulatory protein
rstR	cryptic phage ctxphi transcriptional repressor
mshA	mannose-sensitive hemagglutinin A toxin co-regulated pilus biosynthesis
tcpP	protein P

SEQUENCES:

[1065] In an example, one or more spacers of the disclosure target a respective sequence in this sequence listing. SEQ ID NOs: 1-44 are Type II CRISPR/Cas system sequences, eg, *Streptococcus* sequences.

SEQ ID NO:	SEQUENCES (ALL 5' TO 3')
	PROMOTER
1	TTGAC
2	TATAAT
	TRANSCRIBED LEADER SEQ
3	TATGAAAA
4	A TTTGAG
5	A TTTGAGG
6	GAG
7	GAGG
8	TGAG
9	TGAGG
10	TTGAG
11	TGAGG
12	TTTGAG
13	TTTGAGG
14	A TTTGAG
15	AATTTGAG
16	CATTTGAG
17	GATTTGAG

(continued)

SEQ ID NO:	SEQUENCES (ALL 5' TO 3')
18	TATTTGAG
19	CGATTTGAG
20	ACGATTTGAG
21	TCATTTGAG
22	TTCATTTGAG
23	ATCATTTGAG
24	TTTCATTTGAG
25	AATCATTTGAG
26	AATTCATTTGAG
27	AAATCATTTGAG
28	AAATTCATTTGAG
29	AAAATCATTTGAG
30	AAAATTCATTTGAG
	REPEAT
31	GTT
32	GTTT
33	GTTTT
34	GTTTTT
35	GTTTTTG
36	GTTTTTGT
37	GTTTTTGT A
38	GTTTTTGT AC
39	GTTTTTGT ACT
40	GTTTTTGTACTC
41	GTTTTTGTACTCT
42	GTTTTTGTACTCTC
43	GTTTTTGTACTCTCA
44	GTTTTTGTACTCTCAA
45 The CRISPR leader in the CRISPRI locus of *Streptococcus thermophilus* strain CNRZI 066	CAAGGACAGTTATTGATTTTATAATCACTATGTGGGTATAAAAACGTCAAAATTTCAT TTGA G
46 The CRISPR leader in the CR1SPR1 locus of *E. coli* W3110 CRISPR system	AAACAAAGAATTAGCTGATCTTTAATAATAAGGAAATGTTACATTAAGGTTGGTGGG TTGTTTTTATGGGAAAAAATGCTTTAAGAACAAATGTATACTT AGA

(continued)

SEQ ID NO:	SEQUENCES (ALL 5' TO 3')
47 >tr|J7RUA5 |J7RUA5_ STAAU CRISPR-associated endonuclease Cas9 *OS=Staphylococcus aureus* subsp. aureus GN=cas9 PE=3 SV=1	MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRR HRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNE VEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLK VQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELR SVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEE DIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNS ELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEI PTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNR QTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSV SFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYL LEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRK WKFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIET EQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLY DKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSK KDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNL DVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEV NMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG
48	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFG

(continued)

SEQ ID NO:	SEQUENCES (ALL 5' TO 3')
>sp | Q99ZW2 | CAS9_S TRP1 CRISPR-associated endonuclease Cas9/Csn1 *OS=Streptococcus pyogenes* serotype M1 GN=cas9 PE=1 SV=1	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
49 >ENA | HE980450 | HE9 80450.1 *Staphylococcus aureus* subsp. aureus ORFX gene and pseudo SCCmec-SCC-SCCCRISPR element, strain M06/0171	*[SEQUENCE IS INCORPORATED HEREIN BY REFERENCE FOR USE IN THE PRESENT INVENTION]*
118	GCGGATAACAATTACTAGAGAAAGAGGAGAAATACTATTCTTCTCCTCTTTAAATAAC GAAAACACCCTGCCATAAAATGACAGGGTGTTGATTTCGGCATGAAGCCTTATCTTTG TAGCTTCTGCAAGATTTAAGTAACTGTGTAAGGCGTCCCTTACACTTGCATGTATAGTT

(continued)

SEQ ID NO:	SEQUENCES (ALL 5' TO 3')
pBAV1KTXylR-short1 CRISPR array	ATTATACCAGGGGGACAGTGCAATGTCAAGAATAAACTGTAGAATGACTAGTGACTT AAATCTTGAGAGTACAAAAACCCGTTGGAATCGTGATTAATAGTAACTGTTGTTGTAC AGTTACTTAAATCTTGAGAGTACAAAAACGGCCGAGAAAAGGAGCTGATTCATAGGA CAGTTGTACAGTTACTTAAATCTTGAGAGTACAAAAACTCAAACTTGCCCGTAGTTTA TCTTATAGCCGTTGTACAGTTACTTAAATCTTGAGAGTACAAAAACATTTACCTCCTTT GATTTAAGTGAACAAGTTTATCC
119 Repeat-spacers sequence	TTAAATCTTGAGAGTACAAAAACCCGTTGGAATCGTGATTAATAGTAACTGTTGTTGT ACAGTTACTTAAATCTTGAGAGTACAAAAACGGCCGAGAAAAGGAGCTGATTCATAG GACAGTTGTACAGTTACTTAAATCTTGAGAGTACAAAAACTCAAACTTGCCCGTAGTT TATCTTATAGCCGTTGTACAGTTACTTAAATCTTGAGAGTACAAAAAC
120 tracrRNA-encoding sequence	TTAAATAACGAAAACACCCTGCCATAAAATGACAGGGTGTTGATTTCGGCATGAAGC CTTATCTTTGTAGCTTCTGCAAGATTTAAGTAACTGTGTAAGGCGTCCCTTACAC
121 S1. spacer 1 (DNA Pol III) [PAM= AAAGAAA, in the target is immediately 3' of the 3' terminal GCC]	TGTCCTATGAATCAGCTCCTTTTCTCGGCC
122 S2. spacer 2 (tetA) [PAM= TAAGAAA, in the target is	GGCTATAAGATAAACTACGGGCAAGTTTGA
immediately 3' of the 3' terminal TGA]	
S thermophilus Consensus PAM	NNAGAAW

SEQUENCE LISTING

[1066]

<110> SNIPR TECHNOLOGIES LIMITED

<120> ALTERING MICROBIAL POPULATIONS **&** MODIFYING MICROBIOTA

<130> SNR0001-1 WO

<160> 122

<170> PatentIn version 3.5

<210> 1
<211> 5
<212> DNA

<213> Bacterial

<400> 1
ttgac 5

<210> 2
<211> 6
<212> DNA
<213> Bacterial

<400> 2
tataat 6

<210> 3
<211> 8
<212> DNA
<213> Bacterial

<400> 3
tatgaaaa 8

<210> 4
<211> 7
<212> DNA
<213> Bacterial

<400> 4
atttgag 7

<210> 5
<211> 8
<212> DNA
<213> Bacterial

<400> 5
atttgagg 8

<210> 6
<211> 3
<212> DNA
<213> Bacterial

<400> 6
gag 3

<210> 7
<211> 4
<212> DNA
<213> Bacterial

<400> 7
gagg 4

<210> 8
<211> 4
<212> DNA
<213> Bacterial

<400> 8
tgag 4

<210> 9
<211> 5
<212> DNA
<213> Bacterial

<400> 9
tgagg 5

<210> 10
<211> 5
<212> DNA
<213> Bacterial

<400> 10
ttgag 5

<210> 11
<211> 5
<212> DNA
<213> Bacterial

<400> 11
tgagg 5

<210> 12
<211> 6
<212> DNA
<213> Bacterial

<400> 12
tttgag 6

<210> 13
<211> 7
<212> DNA
<213> Bacterial

<400> 13
tttgagg 7

<210> 14
<211> 7
<212> DNA
<213> Bacterial

<400> 14
atttgag 7

<210> 15
<211> 8
<212> DNA
<213> Bacterial

<400> 15
aatttgag 8

<210> 16
<211> 8
<212> DNA
<213> Bacterial

<400> 16
catttgag 8

<210> 17
<211> 8
<212> DNA
<213> Bacterial

<400> 17
gatttgag 8

<210> 18
<211> 8
<212> DNA
<213> Bacterial

<400> 18
tatttgag 8

<210> 19
<211> 9
<212> DNA
<213> Bacterial

<400> 19
cgatttgag 9

<210> 20
<211> 10
<212> DNA
<213> Bacterial

<400> 20
acgatttgag 10

<210> 21
<211> 9
<212> DNA
<213> Bacterial

<400> 21
tcatttgag 9

<210> 22
<211> 10
<212> DNA
<213> Bacterial

<400> 22
ttcatttgag 10

<210> 23
<211> 10

<212> DNA
<213> Bacterial

<400> 23
atcatttgag 10

<210> 24
<211> 11
<212> DNA
<213> Bacterial

<400> 24
tttcatttga g 11

<210> 25
<211> 11
<212> DNA
<213> Bacterial

<400> 25
aatcatttga g 11

<210> 26
<211> 12
<212> DNA
<213> Bacterial

<400> 26
aattcatttg ag 12

<210> 27
<211> 12
<212> DNA
<213> Bacterial

<400> 27
aaatcatttg ag 12

<210> 28
<211> 13
<212> DNA
<213> Bacterial

<400> 28
aaattcattt gag 13

<210> 29
<211> 13
<212> DNA
<213> Bacterial

<400> 29
aaaatcattt gag 13

<210> 30
<211> 14
<212> DNA
<213> Bacterial

<400> 30
aaaattcatt tgag 14

<210> 31
<211> 3
<212> DNA
<213> Bacterial

<400> 31
gtt 3

<210> 32
<211> 4
<212> DNA
<213> Bacterial

<400> 32
gttt 4

<210> 33
<211> 5
<212> DNA
<213> Bacterial

<400> 33
gtttt 5

<210> 34
<211> 6
<212> DNA
<213> Bacterial

<400> 34
gttttt 6

<210> 35
<211> 7
<212> DNA
<213> Bacterial

<400> 35
gtttttg 7

<210> 36
<211> 8
<212> DNA
<213> Bacterial

<400> 36
gtttttgt 8

<210> 37
<211> 9
<212> DNA
<213> Bacterial

<400> 37
gtttttgta 9

<210> 38
<211> 10
<212> DNA
<213> Bacterial

<400> 38
gtttttgtac 10

<210> 39
<211> 11
<212> DNA
<213> Bacterial

<400> 39
gtttttgtac t 11

<210> 40
<211> 12
<212> DNA
<213> Bacterial

<400> 40
gtttttgtac tc 12

<210> 41
<211> 13
<212> DNA
<213> Bacterial

<400> 41
gtttttgtac tct 13

<210> 42
<211> 14
<212> DNA
<213> Bacterial

<400> 42
gtttttgtac tctc 14

<210> 43
<211> 15
<212> DNA
<213> Bacterial

<400> 43
gtttttgtac tctca 15

<210> 44
<211> 16
<212> DNA
<213> Bacterial

<400> 44
gtttttgtac tctcaa 16

<210> 45
<211> 63

<212> DNA
<213> Bacterial

<400> 45

caaggacagt tattgatttt ataatcacta tgtgggtata aaaacgtcaa aatttcattt 60

gag 63

<210> 46
<211> 103
<212> DNA
<213> Bacterial

<400> 46

aaacaaagaa ttagctgatc tttaataata aggaaatgtt acattaaggt tggtgggttg 60

tttttatggg aaaaaatgct ttaagaacaa atgtatactt aga 103

<210> 47
<211> 1053
<212> **PRT**
<213> Bacterial

<400> **47**

Met Lys Arg Asn Tyr Ile Leu Gly Leu Asp Ile Gly Ile Thr Ser Val
1 5 10 15

```
Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly
            20                  25                  30

Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
        35                  40                  45

Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile
    50                  55                  60

Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His
65                  70                  75                  80

Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu
                85                  90                  95

Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu
            100                 105                 110

Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr
        115                 120                 125

Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala
    130                 135                 140

Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys
145                 150                 155                 160

Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr
                165                 170                 175

Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln
            180                 185                 190

Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg
        195                 200                 205

Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys
    210                 215                 220

Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe
225                 230                 235                 240

Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
                245                 250                 255

Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn
            260                 265                 270
```

```
Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe
        275                 280                 285

Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu
    290                 295                 300

Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys
305                 310                 315                 320

Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr
                325                 330                 335

Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala
            340                 345                 350

Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu
        355                 360                 365

Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
    370                 375                 380

Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile
385                 390                 395                 400

Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala
                405                 410                 415

Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln
            420                 425                 430

Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro
        435                 440                 445

Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile
    450                 455                 460

Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg
465                 470                 475                 480

Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys
                485                 490                 495

Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr
            500                 505                 510

Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp
        515                 520                 525
```

```
Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu
    530                 535                 540

Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro
545                 550                 555                 560

Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
                565                 570                 575

Gln Glu Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu
            580                 585                 590

Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
        595                 600                 605

Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
    610                 615                 620

Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
625                 630                 635                 640

Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu
                645                 650                 655

Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys
            660                 665                 670

Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
        675                 680                 685

Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
    690                 695                 700

Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
705                 710                 715                 720

Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
                725                 730                 735

Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
            740                 745                 750

Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
        755                 760                 765

Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
```

```
    770                 775                 780

Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
785                 790                 795                 800

Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
                805                 810                 815

Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
            820                 825                 830

Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
        835                 840                 845

Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr
    850                 855                 860

Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
865                 870                 875                 880

Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp
                885                 890                 895

Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr
            900                 905                 910

Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val
        915                 920                 925

Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser
    930                 935                 940

Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala
945                 950                 955                 960

Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly
                965                 970                 975

Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile
            980                 985                 990

Glu Val Asn Met Ile Asp Ile Thr  Tyr Arg Glu Tyr Leu  Glu Asn Met
        995                 1000                1005

Asn Asp  Lys Arg Pro Pro Arg  Ile Ile Lys Thr Ile  Ala Ser Lys
    1010                1015                1020
```

Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu
1025 1030 1035

Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly
1040 1045 1050

<210> 48
<211> 1368
<212> PRT
<213> Bacterial

<400> 48

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
180 185 190

```
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
        195                 200                 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
    210                 215                 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225                 230                 235                 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
                245                 250                 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
            260                 265                 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
        275                 280                 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
    290                 295                 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305                 310                 315                 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
                325                 330                 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
            340                 345                 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
        355                 360                 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
    370                 375                 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385                 390                 395                 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
                405                 410                 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
            420                 425                 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
        435                 440                 445
```

```
Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
    450                 455                 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465                 470                 475                 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
                485                 490                 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
            500                 505                 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
        515                 520                 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
    530                 535                 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545                 550                 555                 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
                565                 570                 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
            580                 585                 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
        595                 600                 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
    610                 615                 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625                 630                 635                 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
                645                 650                 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
            660                 665                 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
        675                 680                 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
    690                 695                 700
```

```
Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705                 710                 715                 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
                725                 730                 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
            740                 745                 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
        755                 760                 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
    770                 775                 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785                 790                 795                 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
                805                 810                 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
            820                 825                 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
        835                 840                 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
    850                 855                 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865                 870                 875                 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
                885                 890                 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
            900                 905                 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
        915                 920                 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
    930                 935                 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
```

945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val
1130 1135 1140

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys
1145 1150 1155

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser
1160 1165 1170

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys
1175 1180 1185

```
Glu Val  Lys Lys Asp Leu Ile  Ile Lys Leu Pro Lys  Tyr Ser Leu
    1190                1195                1200

Phe Glu  Leu Glu Asn Gly Arg  Lys Arg Met Leu Ala  Ser Ala Gly
    1205                1210                1215

Glu Leu  Gln Lys Gly Asn Glu  Leu Ala Leu Pro Ser  Lys Tyr Val
    1220                1225                1230

Asn Phe  Leu Tyr Leu Ala Ser  His Tyr Glu Lys Leu  Lys Gly Ser
    1235                1240                1245

Pro Glu  Asp Asn Glu Gln Lys  Gln Leu Phe Val Glu  Gln His Lys
    1250                1255                1260

His Tyr  Leu Asp Glu Ile Ile  Glu Gln Ile Ser Glu  Phe Ser Lys
    1265                1270                1275

Arg Val  Ile Leu Ala Asp Ala  Asn Leu Asp Lys Val  Leu Ser Ala
    1280                1285                1290

Tyr Asn  Lys His Arg Asp Lys  Pro Ile Arg Glu Gln  Ala Glu Asn
    1295                1300                1305

Ile Ile  His Leu Phe Thr Leu  Thr Asn Leu Gly Ala  Pro Ala Ala
    1310                1315                1320

Phe Lys  Tyr Phe Asp Thr Thr  Ile Asp Arg Lys Arg  Tyr Thr Ser
    1325                1330                1335

Thr Lys  Glu Val Leu Asp Ala  Thr Leu Ile His Gln  Ser Ile Thr
    1340                1345                1350

Gly Leu  Tyr Glu Thr Arg Ile  Asp Leu Ser Gln Leu  Gly Gly Asp
    1355                1360                1365
```

<210> 49
<211> 56115
<212> DNA
<213> Bacterial

<400> 49

atgaaaatca ccattttagc tgtagggaaa ctaaaagaga aatactggaa gcaagccata 60

gcagaatatg aaaaacgttt aggcccatac accaagatag acatcataga agttccagac 120

gaaaaagcac cagaaaatat gagtgacaaa gaaattgagc aagtaaaaga aaaagaaggg 180

caacgaatac tagccaaaat taaaccacaa tccacagtca ttacattaga aatacaagga 240

aagatgctat cttccgaagg attggcccaa gaattgaatc aacgcatgac ccaaggccaa 300

agcgactttg tattcgtcat tggcggatca aacggcctgc acaaggacgt cttacaacgc 360

agtaactact cactatcatt cagcaaaatg acattcccac atcaaatgat gcgggttgtg 420

ttaattgagc aagtgtatag agcgtttaag attatgcgag gagaagctta tcataaataa 480

aactaaaaat taagttgtgt ataatttaaa aagttaatga gatgtggagg gattacatat 540

atgaaatatt ggattatacc ttgcaatatc atatgatgtt tatagagtgt ttaataaagg 600

ttctgttgca aagttagaaa aatatagcta attactattt tatcatgtcg ttgttccctt 660

aactggctaa catatgccta atttcgtggc atggcgaaaa tccgtagatc tgaagagacc 720

tgcggttctt tttatataga gcgtaaatac attcaatacc ttttaaagta ttctttgccg 780

tattgatact ttgatatctt gtctttctta ctttaatatg acggtgatct tgctcaatga 840

ggttattcag atatttagat gtacaatgac agtcgggatt cagtttaaac gctttaatta 900

ctttagccat tgctaccttc gttgaaggtg cctgatctgt aatcaccttt tgaggtttac 960

caaattgttt aatgagacgt ttgataaata catatgctga atgattatct cgttgcttac 1020

gcaaccaaat atctaatgta tgtccctctg catcaatggc acgatataaa tagctccatt 1080

ttccttttat tttgatgtac gtctcatcaa tacgccattt gtaataagct tttttatgct 1140

ttttcttcca aatttgatac aaaattgggg catattcttg aacccaacgg tagaccgttg 1200

aatgatgaac gtttacacca cgttccctta atatttcaga tatatcacga taactcaatg 1260

tatatcttag atagtagcca acggctacag tgataacatc cttgttaaat tgtttatatc 1320

tgaaatagtt catacagaag actccttttt gttaaaatta tactataaat tcaactttgc 1380

aacagaaccg tattatggaa tagagatgtt ggtaacattt atacaggatc attatactta 1440

agtttaattt cgttattaca gaaccacaca ttccaaccag aagagaaagt atgtctattt 1500

agttatggtt caggagcagt aggagaaatc tttagtggtt caatcgttaa aggatatgac 1560

aaagcattag ataaagagaa acacttaaat atgctagaat ctagagagca attatcagtc 1620

gaagaatacg aaacattctt taacagattt gataatcaag aatttgattt cgaacgtgaa 1680

ttgacacaag atccatattc aaaagtatac ttatacagta tagaagacca tatcagaaca 1740

tataagatag agaaataaac tagtggccga ttgtgcttga tgagcttggg acataaatcc 1800

taactcgaaa taaataagca tatcactaaa ctgatttttt aaagtttaca gtgatatgct 1860

tatttttta tcttacgatt ttgtacgtgc atgcttgcct aggggtatgg ctcgagccct 1920

ggagccatta gtctctcgca catactattc cctcaggcgt cagcacttac aaaatcggtt 1980

gtaattttca tttttatacg cattcttact gagattatac taataagagg aatagtaaaa 2040

gcaattctaa gtaaaattgc agataagagg tttgttaaaa gcagttctaa gtaaaattgc 2100

agataagagg tacgttaaaa gcaattccat gcaaaattgc tgataagggg taagttaaaa 2160

gcagttctca gtaaaattgc agataagagg tacgttaaaa gcagttctag gcaaaattgc 2220

agataagagg tgcgttaaaa gcagttctca gtaaaattgc tgataagggg taagttaaaa 2280

gcaatcctaa gtaaaattgc agataagagg taagttaaaa gcaatcctaa gtaaaattgc 2340

agataagggg tacagaaaaa ctagacttga ttacaaaatg gagcttggga cataaatgat 2400

tttttaaaaa tgagatgaga cgtagattaa ctccataatc aatacgaatc tatcgacttc 2460

tttatttatg atattcatct ctttttaatg gaaataaaag tgcgattaat gtgataatac 2520

agttacgtta attaaaaaaa taaaaatgca aggagaggta atatgctaac tgtatatgga 2580

catagaggat tacctagtaa agctccggaa aatacaattg catcatttaa agctgcttca 2640

gaagtagaag gtataaactg gttggagtta gatgttgcaa ttacaaaaga tgaacaactg 2700

attatcattc atgatgatta tttagaacgg actacaaata tgtccgggga aataactgaa 2760

ttgaattatg atgaaattaa agatgcttct gcaggatctt ggtttggtga aaaattcaaa 2820

gatgaacatt tgccaacttt cgatgatgta gtaaaaatag caaatgaata taatatgaat 2880

ttaaatgtag aattaaaagg tattactgga ccgaatggac tagcactttc taaaagtatg 2940

gttaagcaag tggaagaaca attaacaaac ttaaatcaga atcaagaagt gctcatttca 3000

agctttaatg ttgtgcttgt taaacttgca gaagaaatca tgccacaata taacagagca 3060

gttatattcc atacaacttc gtttcgtgaa gactggagaa cacttttaga ttactgtaat 3120

gctaaaatag taaacactga agatgccaaa cttactaaag caaaagtaaa aatggtaaaa 3180

gaagcgggtt atgaattgaa cgtatggact gtaaacaaac cagcacgtgc aaaccaactt 3240

gctaattggg gagttgatgg tatctttaca gacaatgcag ataaaatggt gcatttgtct 3300

caatagaaag ttagaggtga gtcttacgtt tcagtgacgg tagacttacc tttaacatgt 3360

tacatactaa aaaattaatt tgaataagaa agagagacat atatgaaata cgatgatttt 3420

atagtaggag aaacattcaa aacaaaaagc cttcatatta cagaagaaga aattatccaa 3480

tttgcaacaa cttttgatcc tcaatatatg catatagata aagaaaaagc agaacaaagt 3540

agatttaaag gtatcattgc atctggcatg catacacttt caatatcatt taaattatgg 3600

gtagaagaag gtaaatacgg agaagaagtt gtagcaggaa cacaaatgaa taacgttaaa 3660

tttattaaac ctgtataccc aggtaataca ttgtacgtta tcgctgaaat tacaaataag 3720

aaatccataa aaaaagaaaa tggactcgtt acagtgtcac tttcaacata caatgaaaat 3780

gaagaaattg tatttaaggg agaagtaaca gcacttatta ataattcata ataaaacagt 3840

gaagcaacca tcgttacgga ttgcttcact gttttgttat tcatctatat cgtatttttt 3900

attaccgttc tcatatagct catcatacac tttacctgag attttggcat tgtagctagc 3960

cattccttta tcttgtacat ctttaacatt aatagccatc atcatgtttg gattatcttt 4020

atcatatgat ataaaccacc caatttgtct gcctgtttct ccttgtttca ttttgagttc 4080

tgcagtaccg gatttgccaa ttaagtttgc ataagatcta taaatatctt ctttatgtgt 4140

tttatttacg acttgttgca taccatcagt taatagattg atattttctt tggaaataat 4200

atttttcttc caaactttgt ttttcgtgtc ttttaataag tgaggtgcgt taatattgcc 4260

attattttct aatgcgctat agattgaaag gatctgtact gggttaatca gtatttcacc 4320

ttgtccgtaa cctgaatcag ctaataatat ttcattatct aaatttttgt ttgaaatttg 4380

agcattataa aatggataat cacttggtat atcttcacca acacctagtt ttttcatgcc 4440

tttttcaaat ttcttactgc ctaattcgag tgctactcta gcaaagaaaa tgttatctga 4500

tgattctatt gcttgtttta agtcgatatt accatttacc acttcatatc ttgtaacgtt 4560

gtaaccaccc caagatttat ctttttgcca acctttacca tcgattttat aacttgtttt 4620

atcgtctaat gttttgttat ttaacccaat cattgctgtt aatatttttt gagttgaacc 4680

tggtgaagtt gtaatctgga acttgttgag cagaggttct tttttatctt cggttaattt 4740

attatattct tcgttactca tgccatacat aaatggatag acgtcatatg aaggtgtgct 4800

tacaagtgct aataattcac ctgtttgagg gtggatagca gtacctgagc cataatcatt 4860

tttcatgttg ttataaatac tcttttgaac tttagcatca atagttagtt gaatatcttt 4920

gccatctttt ttctttttct ctattaatgt atgtgcgatt gtattgctat tatcgtcaac 4980

gattgtgaca cgatagccat cttcatgttg gagctttta tcgtaaagtt tttcgagtcc 5040

ctttttacca ataactgcat catctttata gcctttatat tctttttgtt ttaattcttc 5100

agagttaatg ggaccaacat aacctaatag atgtgaagtc gctttttcta gaggatagtt 5160

acgactttct gtttcattag ttgtaagatg aaatttttt gcgaaatcac ttaaatattc 5220

atccattttt ttaacggttt taagtggaac gaaggtatca tcttgtaccc aattttgatc 5280

catttgttgt ttgatatagt cttcagaaat acttagttct ttagcgattg ctttataatc 5340

ttttttagat acattctttg gaacgatgcc tatctcatat gctgttcctg tattggccaa 5400

ttccacattg tttcggtcta aaattttacc acgttctgat tttaaatttt caatatgtat 5460

gctttggtct ttctgcattc ctggaataat gacgctatga tcccaatcta acttccacat 5520

accatcttct ttaacaaaat taaattgaac gttgcgatca atgttaccgt agtttgtttt 5580

aattttatat tgagcatcta ctcgtttttt atttttagat actttttta ttttacgatc 5640

ctgaatgttt atatctttaa cgcctaaact attatatatt tttatcggac gttcagtcat 5700

ttctacttca ccattatcgc ttttagaaat ataactgcta tctttataaa cttgtttgaa 5760

atttttatct tcaattgcat caatagtatt attaatttct ttatcttttg aagcataaaa 5820

atatatacca aacccgacaa ctacaactat taaaataagt ggaacaattt ttatcttttt 5880

catcaatatc ctccttatat aagactacat ttgtagtata ttacaaatgt agtatttatg 5940

tcaaaataat gttataattt ttgtgatatg gaggtgtaga aggtgttatc atctttttg 6000

gttctgttgc aaagtaaggt tctgttgcaa agtaaaaaat atagctaacc actaatttat 6060

catgtcagtg ttcgcttaac ttgctagcat gatgctaatt tcgtggcatg gcgaaaatcc 6120

gtagatctga agagacctgc ggttcttttt atatagagcg taaatacatt caataccttt 6180

taaagtattc tttgctgtat tgatactttg ataccttgtc tttcttactt taatatgacg 6240

gtgatcttgc tcaatgaggt tattcagata tttcgatgta caatggcagt caggtttaag 6300

tttaaaagct ttaattactt tagccattgc taccttcgtt gaaggtgcct gatctgtaat 6360

taccttttga ggtttaccaa attgtttaat gagacgtttg ataaacgcat atgctgaatg 6420

attatctcgt tgcttacgca accaaatatc taatgtatgt ccctctgcat caatggcacg 6480

atataaatag ctccattttc cttttatttt gatgtacgtc tcatcaatac gccatttgta 6540

ataagctttt ttatgctttt tcttccaaat ttgatataaa attggggcat attcttgaac 6600

ccaacggtag accgttgaat gatgaacgtt tacaccaagt tcccttaata tttcagatat 6660

atcacgataa ctcaatgcat atcttagata gtagccaacg gctacagtga taacatcctt 6720

gttaaattgt ttatatctga aatagttcat acagaagact ccttttggtt aagattatac 6780

tataaattca actttgcaac agaacctctt aatcttattg gatacaataa ctaacttaaa 6840

cacattatct acaagtaaaa acttgtaaaa agaacatgtg ttcttaatta gtgtataatt 6900

agtttaaatt gattatgggg tgatgtaatg aatgagagta atcaaattat attcgatcat 6960

ataacaaagg ctattgaaaa tatgaaaaat aaaggctggg aaacaaggta tgggcaagaa 7020

actttaatgt atgatgtatt tgatgcttat gattcaagag aaaatttaat agtagaagcg 7080

caagtaggta taggaaaatc atttggatat ttaattcctg gtatattgat ttcaaaaagt 7140

actaaaaaac ctttaattgt aacgacttcc tcaattcaat taactgagca actagtcaat 7200

gatatacagc aagttgaaga aattttacat atatctgttg attgtatagt aggtaaaggt 7260

gtaactaact atccttgttt taaagaaatt catcgtaaaa atttaagtaa attaaattta 7320

atggaggctt tagatatagc aaataatggt ttaacaaaac aaactgtacg tgagactaat 7380

ttaaaatgga atcaaatatc taccaataat tgtattatga gtaaatgtca ctataaaaat 7440

gaatgcgctt attttaaaat gagaaataag ttaaaagagg gaaatagata tattagactc 7500

aatgaatata agccgaaagt cataatcgtt aatcaagata tgttaatgat gaattttaaa 7560

aagttaacct ttggaaaaga gagtttaatc tatgatgatc cttgtatgct aatcgtagat 7620

gaagttcaca acttagaaga gaaacaacgt gctaacatga ctaaaactat taactctaaa 7680

actatgataa ataaaataaa agaaggcgca aaccgagttg gtagtagatc tcggtattta 7740

aaaaatataa aaatgattga agattggttt aatcttcaaa aggatagtgc caaaagtgag 7800

atatacaaaa atggtaattg tttatctact ggacgagttg atataaaacc tgttagtcat 7860

catcaaatgg ccaaattaat tagtatgaca aaggaaatag ttgaagaatt cgatgttaat 7920

aatattgttg tatttaaaca atcttcgtta gttagtaatg aacaattaga aatagctaat 7980

aatctgttaa cattatttaa aaatcttcaa aataacagcg acaaatatat attttggact 8040

gaaataacta agcaagacca aattgatatt tcattttgtc ccaataatat tgctgaaaca 8100

cttagaaaaa cagtttttag tacaagttac ccagttgtat gtctttctgc aacaattact 8160

aataaaacta ataatgaaaa tagctatgag tacattaaag aaataattgg ttttaaaggc 8220

tatgaggaag atattaagta taatgatttt ccgtacaagg aaagcaggct ctacatacct 8280

ccaaacttac caaaatttga taaacgtgat gttaaatatt atgaagaaat aggtaaacat 8340

atttttgaat tggcaagcca aaataaggga ggcactttga tattgtttac tgctaaagat 8400

gacattgacg gtgtatataa tgatttatca aaaaggaaat ttaataaaac tatctatgta 8460

gatgatggta gtaaaagtca aaatgaaatt attgaatcat ttaaaaaaac taaaggtgta 8520

atattgggta cgggagtatt ttgggaagga atagacttaa aaaatgaatt attaacatta 8580

ttagttattg tcagattacc ttttcctacg attgatccaa ttacaaaata taagattact 8640

aaattaaatg atagtaatga agcagttata gttcctgaaa tgataattaa actgaaacaa 8700

ggtgtaggaa gattagtaag aacaaagcag gataaaggtt tattagtgtt attagattca 8760

agaatgaata aacccattta taagcataaa gaagcagtat tagatgcatt accaattaaa 8820

aatatgattc ataaagatga agtcaaagat ttccttaata acatttaaag taagtagtaa 8880

tagttatcca atattttcta gttgattctc ttataatatg tccaattaag tgtagacgat 8940

tcaagatgat attgagggaa ggatgagaat gtgtctttta ggtagctcat ttgatgataa 9000

ccttaaacat aatatttact ttcatcctgc tgaaaaaaga ggttacagtt tcaagccgat 9060

agaatacatt gggctatata aagacaaagc aataaaagct atcggaaaag tagacaaagt 9120

aatagtgacg gaaataaatg agagctcatt atccttgaag acagtttatc ctgtaggtac 9180

tgagttatcg attgatgagt atgaaatagt taaaaataaa ataatggtgc ttgggaaaga 9240

aaagtggact aatctattga acgaacccca ttattattac ttaattgaag attttataga 9300

gacagactac aagaaaacat ctaaaggtgg atcgatggga gttaagtatt ttaatgtgaa 9360

tgaaattttg aatagggatt gcttaaccac cgaacaaatc gctaaagaat tatgtaataa 9420

agattgggaa tgataaatag tgaatgaagc tataatgcaa taatgtatgt ataacataat 9480

caagtaatga aattaatgat aagggagtat ggagtaaaga attgaattag tttgcatggt 9540

ctatgtatct taaaggtgtt gaaggagata caccgaagta tgcagtacca attcttgaaa 9600

cagactttag tgatttacca tctacacata cgacagtcgg atcacttgat cggtttcgag 9660

ataagacaat gagcattaac agaagcggga gtaactgttg aattccattt atatccagtt 9720

agatactaga gtttgaagtg ataaatcttg attcggatgc tgaaataagt aaaagagcag 9780

ttactgaata taatataagt ataagtatgc cactagatat tttaaggggt atttattcac 9840

ttttataaaa tcatctgcta agatagtgtg gaagattttt ataggagctg attataatgc 9900

gacatataaa tatgtcatac atcatacgct atagtactct aactcaaaga taagaagtac 9960

ttagttatga atagcattca tttcacaaag tacttcaaca aaaatggagg aatatgaaat 10020

gaataaaata gaagtgtata agtttgttaa agtaaagcag ttagtatatc aattgattaa 10080

gttatatcgt acaaacgata tgaattccca taaaacacaa aaagattttt tactaaatga 10140

aattaatgat atctttaaag aaaaagatat tgatatctcg gactttatta catcgattga 10200

cgatgtaaaa ttaactaaga aaaaagcaga acatctttta aatgaattaa aagtgtacat 10260

ccaagatttt gaaatacctt catcaagtca actggagaaa atttttcgta aagtaaaaaa 10320

attaaagaga ccagatataa atttaattga tacaaaagaa atttcatatt taggatggaa 10380

tgataattct tctaaccgaa aatatatcgt ttataaaaat ttagatgata aattcgaagg 10440

tatatatggc gaaatttcac caaataaagt aaaaggattc tgtaaaattt gtaatcagga 10500

atctgataca tcactctttc tcaataaaac taaacataat aagagtagtg gaacatatac 10560

taaaaaagga gattacattt gttatgacag ttttaaatgt aatcagaacc tagatgatat 10620

aaataatctt tacgaattta ttgttaaaat aaaatagatc caaacagccc tgatctttag 10680

aactaatgat cagggttgtt atattttgta cctcactctt aatcgatacg ctgagcctta 10740

taataaataa agctaagatt ggttagtgtg taactttatg caatttgaaa agatataatg 10800

aaataagaga gttttggtga tgatattaat atttaaatag atctatagac aattaatcgc 10860

tgttgatttt taacaaatca attcaattta gttatgtttt gtgatggcat actaattaaa 10920

aaaagaagaa ctaggttttg tgccctagtt cttcttttta tttaattgtt attcctattt 10980

cttaataagc gtaatgaatt aagtatgaca agtatcgcag cgcctgtatc acttaatact 11040

gccaaccaaa gtgttaatag cccaggaaat acaagaatga atgcaatcaa ttttatgatg 11100

atagaaaaat atatgttttg cttaatgata tttttagctt ttttacttat agaaatcgta 11160

cgtgttaatt gatttatatc atctgacata agtacgacat cagctgtttc cattgctgta 11220

tctgaaccaa tacctcccat tgcaatacct acgtcacttt gtgctaaagc aggggcatca 11280

tttatgccgt ctccaatcat agcgacacga tatcctttat tttgcaaatc ttttattgca 11340

gctaatttgt caccaggcat taattcagca taaacttctt taatacctga aagttgtgcg 11400

atttttgggg cagtaccttt attatctcct gttagcatta tagtgttttt aatgtgagta 11460

ccatttaatt gttgtattat ttgcttgata tttgaacgta atgggtctgc aatagtaatc 11520

aagccatgaa tcatgctaga agatgcaata atgatgacgg taaaaccttc ttgctcatat 11580

ttgtttattt cttctttata gttgttaatt tttttgttaa tagactcaat taattttacg 11640

ttaccagcat aaacgtaact ttcatttatt ttacctttga tgccttgacc aacgatattt 11700

tcaaaatcag ttacataagc ttttttaaca tttaattgca ttgcgtaatc tacgatagcg 11760

ttgctaatag ggtgggttga gtaggattcc aaactcatag caatatttag caatgttgtt 11820

tcattagcat ctatagtttt aatcgtatcc acctttggac gaccttcagt caaagtacct 11880

gttttatcaa aagccaacgc agataaagta cctagacctt ctaaatggtt accacccttg 11940

ataagtacgc catttttgc tgcacttcca atcgcagtaa cgatagctac cggtgttgaa 12000

atgacaagtg cacatggaca tgcgataact agtaattcaa gccctttata taaccattcg 12060

ccccacgtac ctaaagagaa tagaggtggt atgaccatga ctaataatgc taatacaaag 12120

acgataggtg tataaatttc tgaaaatcga tctataaaag cttgcgtcgg tgctttattt 12180

tcttgtgctt cttcaaccat atgaatgatt ttagaaagcg ttgtatcttc taccaattta 12240

gacacccgga tgtataaagt accattttca ttgatagagc ctgcatatac ttcatcgtta 12300

atcgttttat caacaggtat cgattcaccg gtaataggtg cttggtttaa actggaactt 12360

ccatcggtaa ttgttccatc taaagggacg cgatcacctg gtttgataag taagatttca 12420

ccaactcgta tatctgtcaa atctttagta gttgttccgt tttctgttat gacatttgct 12480

gttgtggaag taatatccat taatgattgt atagaattac gtgttttatc tattgaaatt 12540

gtttgcagca aggttccaat tgtaaataac aatacaacaa tcgcaccttc aaaatattca 12600

ccaatgaata tagcgccaat aacagctacg gacatcaaca cattcatatc taagctttta 12660

gatttaatgg cataatatgc gcttttcaat ggttttatgc cactaacaat aatggctatg 12720

atatacatta aatttgcaat taatactggc aatgttgtca gagttacaac taaaccaagg 12780

ataagaaata atgttgaaat gataggtttt ctgaatactc tccatttgga attactactt 12840

ttttttgttg gacttaatgt tgcagaataa ccgattttag aaacttcttt ttcaatattt 12900

ttagttttat taccttcaaa aggttctgtt gcaaagttga atttatagta taattttaac 12960

aaaaaggagt cttctgtatg aactatttca gatataaaca atttaacaag gatgttatca 13020

ctgtagccgt tggctactat ctaagatata cattgagtta tcgtgatata tctgaaatat 13080

taagggaacg tggtgtaaac gttcatcatt caacggtcta ccgttgggtt caagaatatg 13140

ccccaatttt gtatcaaatt tggaagaaaa agcataaaaa agcttattac aaatggcgta 13200

ttgatgagac gtacatcaaa ataaaaggaa aatggagcta tttatatcgt gccattgatg 13260

cagagggaca tacattagat atttggttgc gtaagcaacg agataatcat tcagcatatg 13320

cgtttatcaa acgtctcatt aaacaatttg gtaaacctca aaaggtaatt acagatcagg 13380

caccttcaac gaaggtagca atggctaaag taattaaagc ttttaaactt aaacctgact 13440

gtcattgtac atcgaaatat ctgaataacc tcattgagca agatcaccgt catattaaag 13500

taagaaagac aaggtatcaa agtatcaata cagcaaagaa tactttaaaa ggtattgaat 13560

gtatttacgc tctatataaa aagaaccgca ggtctcttca gatctacgga ttttcgccat 13620

gccacgaaat tagcatcatg ctagcaagtt aagcgaacac tgacatgata aattagtggt 13680

tagctatatt tttttacttt gcaacagaac ccttttttgca ataatacatt taattaaaac 13740

cttgttttaa acatcgatga caagatctaa tgtaggacgt ggagacatca ttttcggaat 13800

gatataaata aatccataaa tggagataaa tcgaaaacat ttatacccat agggggtatg 13860

tgttatagta taagtatagc taatataata ttttcataaa taggaggggt taatttgaat 13920

aataatggtg aagagcataa tcctcaaaat cacatgaatc attccaatca catgcatcat 13980

gataaccatg cctcacatca tcatagtggc catgcacatc atcatggaaa ttttaaagtt 14040

aagttttttg tttcattaat tttggcaata cctatcattc ttttatcgcc aatgatgggt 14100

gttaacttac cttttcaatt cacatttcca ggttctgaat gggtagtgtt aatattaagt 14160

acaattttat tcttttatgg tggtaaaccg ttcttgtctg gtggtaaaga tgaaattgct 14220

acaaaaaaac caggcatgat gaccttagtt gccctaggta tttcagtagc ttatatttat 14280

agcttgtatg ctttttatat gaataacttt agtagtgcaa ctggtcatac aatggacttt 14340

ttttgggaat tagcaacctt aattttaatt atgctattag gacattggat agaaatgaat 14400

gctgtcggaa atgctggaga tgctttaaag aaaatggcag aactgttacc taatagtgct 14460

attaaagtta tggataatgg ccaacgcgaa gaagttaaaa tatcagacat catgactgat 14520

gatatcgtcg aagtaaaagc cggagaaagc attccgacag atggtattat cgttcaagga 14580

caaacatcta tagatgaatc cctagtcact ggagaatcta aaaaagtaca aaaaaatcaa 14640

aatgacaacg tcatcggggg ttctattaat gggtctggaa caatacaagt caaggttaca 14700

gctgtgggag aagatggata tctttctcaa gttatgggac ttgttaatca agcacaaaat 14760

gataaatcta gtgctgaatt gttatctgat aaagtagcgg gttatttatt ctactttgct 14820

gtaattgttg gcgtgatttc ttttattgtc tggatgctca ttcaaaatga tgttgatttt 14880

gcattagaac gtcttgtaac tgtgttagtc attgcttgtc cacatgcttt aggcttggca 14940

ataccttttag tcactgcacg ttctacttca attggtgcac ataatggttt aattattaaa 15000

aatagagagt ctgtagaaat agctcaacat atcgattatg taatgatgga taaaactggt 15060

actttaactg agggtaactt ttctgtgaat cattatgaga gctttaaaaa tgatttgagt 15120

aatgatacaa tattaagcct tttcgcctca ttagaaagtc aatctaatca cccattagct 15180

ataagtattg ttgattttgc gaaaagtaaa aatgtttcat ttactaatcc acaagacgtt 15240

aataatattc caggtgtcgg attagaaggt ctaattgata ataaaacata taaaataaca 15300

aatgtctctt atcttgataa acataaactt aattatgacg atgacttatt tactaaatta 15360

gctcaacaag gtaattcaat cagctattta attgaggatc aacaagtcat tggcatgatt 15420

gctcaaggag atcaaattaa agaaagctca aaacaaatgg tagctgattt actatcaaga 15480

aatattacac cagtcatgct tacaggtgac aataatgaag tggcacacgc tgtcgcaaaa 15540

gaattaggta ttagtgatgt ccacgcacaa ctcatgccag aagataagga aagcattata 15600

aaagattatc aaagtgacgg taataaagtc atgatggtcg gagacggtat caacgatgcg 15660

ccgagtctta taagagccga tattggtata gcaattggtg caggcacaga tgttgcagta 15720

gattcaggtg atatcatact tgttaaaagt aatccatcag atattattca tttcttgacc 15780

ctttcaaata atactatgag aaaaatggtg caaaacttat ggtggggtgc aggttataat 15840

attgttgctg tacctttagc agctggtatt ttagcgttta tcggcttgat tttatcacct 15900

gcaataggtg ctattttaat gtcgctaagt acggttattg ttgcaattaa tgcctttaca 15960

ttaaaattaa aataaaagat aggagtttta ttatgattaa aaaattattt tttatgatat 16020

taggatcatt actaatatta tcagcttgct ccaataatga tgaaaaagat aaagacacta 16080

atgaccaaaa aagtgagagc catatgaagc ataatgatga aagtaaagtt cctgaaggta 16140

tgaaatcgac taatgagggt gaatttaaag taggagataa agtaacgatt acagcagggc 16200

atatgccagg tatgaaaggt gcagaagcta ctgtaaaagg tgcgtataaa acatatgctt 16260

atgttgtaag ttataaaccc acaaatggaa atgaaaaagt aaacaatcat aaatgggtcg 16320

taaacgaaga gatcaaagat gcacctaaag atggatttag taagggcgat actgttaaat 16380

tagaagcaag tcatatgtct ggtatgaaag gtgctacagt caatatagat aacgtgaaaa 16440

agactactgt ttacgtagtt gattacaaat ccaaagataa tggtaaaatc attaaaaatc 16500

ataaatggat gacaggaaat gaactgaaag cacgataaaa atctagttct agattgagaa 16560

ataaatagat ataaagatat cctccttaat caataattta aataacttat tattgttaag 16620

gaggatattt tttagtgtgt aaattaaaaa gaattttaga agaataacat ttatcaaaaa 16680

actgttcatt accttattaa atgaaattat ataattaaaa accgcatcat taaccgatac 16740

gcagaagcgt accacaaata aaactaaaaa atagattgtg aaaaatttat ctttgaaagg 16800

ggtcgaatga gagacataga agtagaaaaa gcctataaga tgatgaaaga acaaggtaaa 16860

ataaatttta gtgaatacca taagacacta ttacatatac atactctagg atcacatgat 16920

ttaaaattat atgatgaatg gaacaaacaa tatttctatc tattagccaa gatgaaatat 16980

ttagaattgc taaaaaagaa aatattcttt acagtgagta ttctaatata aatgatttaa 17040

aggaattaga tatttttaag aactataaaa atctaaagga aatgctttct tattttattt 17100

tggtatacag catatatact aatggaatat caatagttgt tgttgcagat cataatactg 17160

tgtgaagaat ggaataaaga attctgtacg gcttctgttc tactcttttt tattaaaaat 17220

atataataaa gaatttttta aaatagtaaa atttatactt tgattaactt ttgacataat 17280

tatgaacatc taaataagta aaaatcaaat atttgaaatt gtgttaatat gaatgtagaa 17340

actaaggtaa taatttcatt taaaatagct tttaaaattt atatacatat tagcgttgtg 17400

atacacgtag aaaggggggaa tcaaaagtgg aaacgctaat ggttaaagca atggggacgg 17460

tgatacgttt atcgattgaa catcaacatt cgagtacatt acttcaagaa gctgaattaa 17520

aaattcatga ttgggaatct caatttagtg ctaatgatcc aaaatcagat ttgatgaatg 17580

tgaatcaaca tgcaggtatc gcaccagtca aggttaatcc tgagatgttt aacatgatac 17640

gttatggtta tgaaactaca ctatcttcta attttaagat gaacatattg atagggccat 17700

tagttaaatt atggaaaatt ggtttcaaag atgcattgaa acctaaagaa gaagatatac 17760

aacgcgcttt aatgtgtatg aatcctgaaa atcttgttct aaattcaaag acacatgaag 17820

tatttcttac acgaccagga atggagattg atttaggagc cattgtaaaa ggttattttg 17880

ctgatcaatt acagcaatac tttttatctc atggtgtagc ttctgccatt atcgatttag 17940

gtggtaatgt tttaacaatt ggtagacaac cggaaactct agaaaaatgg catgtaggta 18000

tacgcaatcc tttttataaa gatgcattat cactcgttac attaaatgta gagcatcaat 18060

cagttgttac ttcaggcatc tatgaacgct atttcataca ggaacatcaa ttatttcatc 18120

atattttgga ttccaataca gggtatcctg tagataatga tatcgctagc gtgacaatca 18180

tatctgatca tgggattgat ggcgaggtat ggagtacaat ttgtagtttt ggtcaatcac 18240

aacaaaatat tgaattatta aacctcatag atggtgttga aggcatcatt gtgacaagag 18300

atggaaacgt tttaatgact tcaaaaatgc aaaggtattt ataaaatttt aaaaaatact 18360

cataaagagt attttatttt taagtataaa tgtgaaaaaa atcacaatgt attaggagga 18420

attgctatga gtaaagaatt atttgatact tttaaattta aatgtggtgc cgaattaaaa 18480

aatagagtat taatggcgcc catgactatc caagctggat attttgatgg aagtgttaca 18540

tcagaaatga ttgattatta tcaatttagg gctggtgatg cttcagcaat tattgttgaa 18600

agttgttttg ttgaaaacca tgggagaggc ttcccaggag ctataggtat tgataacgat 18660

gacaaaatac ctggacttaa acgtctagca gaagcgattc aatctaaggg gtcaaaagcg 18720

attttacaac tttatcatgc tggaagaatg gcgaatccta agttcaatga aggtgaacaa 18780

cccatatctg cgagtcctat tgcagcatta agacctgatg cggtaccacc aagagaaatg 18840

acacatgctc aaatcaatca gatgatagaa gatttcggag aagcgacacg tcgtgctata 18900

gaagcaggat ttgatggtgt tgaaattcat ggtgccaaca catacttatt acaacaattt 18960

ttctctccgc attctaatcg aagagaagat acatggggag gcagtcgtga aaaacgtacg 19020

cgattcccaa tcgaagtact aacaaaagtt caacatgtcg ttgatgaaaa tgaagcttct 19080

caatttataa taggatatcg attctcacct gaagaaattg aagagccagg catacgtttt 19140

gaagatacaa tgtttttact aaatacttta gcaaaatatg aacccgatta cttccatata 19200

tcagcaaaca gttatcaacg tacgtccata gtgaatcaag aagatacaga acctttaatt 19260

aataaatatc ataaaatgca aagtgcacaa ttggcacaaa ttccattaat cggtgttggt 19320

agtattgcgc aacgaaaaga tgcagaacat gcccttgaac tcggttataa tcttttaagt 19380

gttgggaaag cctatttagt agaacctcaa tggacggata aaatttcaca aaacgaagaa 19440

gttgaacaat ttgtcgatat acatgaccaa aaaatacttc acataccatc tcctctatgg 19500

aaagtgatgg acttcatgat tttagataaa gaagaagaac atcgtaaata tgaaagatta 19560

aaagcactac aaaataaaaa agtgaaattt aacaaaggca cgtatcatgt ttatgcaaaa 19620

ggtcataatg gcaatctacc aatgaaagtt caattgtcgg aagataaaat tgtaagtatt 19680

gaagtagatg atagtggaga gtctgaaggt atagcgaacc cagtttttga acgtttacct 19740

caagatatca tcaatggaca aacgttaaat gtcgatgtca tatcaggtgc gacagtaaca 19800

agtgaaggca tcgtacaggg cattgcagat gcaattgaac aagcaggcga agatccagat 19860

attttacgag cacgtcctaa accagtcgtt cagtggtctg atgaggtcgt tgaagagacg 19920

actgacgtag tcgtgattgg tacaggtggt gctggtttaa gtgcagctgc tacggtttta 19980

gacgaaggaa aagaagtcat catgcttgag aaatttgcgg ctataggtgg caataccatc 20040

cgaacaggtg gtcaagtcaa tgctgctgag cctaaatggc aaagtgcatt ccctgcactc 20100

gctggtgaaa aagacacatt gatacagtta ttaaatcatg atgaaaagga tatagatgag 20160

gcttatcttg aagatttcaa tactttaaag cgtcaaatta aagattacct tgaaaaaagt 20220

agcaataaag atgattatct tttcgattct gttgaattac atcgcattca aacgtatcta 20280

ggtggaagac gtaaagatcg taataatgtc gaaatttcag gtgattatga tttagttaaa 20340

acactcacag ataatgtttt ggaatcagta tactggttga aagataaagg ggtacatttt 20400

gatcgttcat ttgtagatat gcctgttggt gcactatggc gtcgtggtca taaaccaatg 20460

aaagcacaag gtttagagta cattgaaaat ctaggagact acgttaaaca gcatcatggt 20520

cgtattttta cagaaactac tgcagaaaaa ttaattaaag aaggtaatca agttgttggt 20580

attgaagcac gtaaagcaaa tggtgctaaa gttaagattc acactcgtca cggcgtagtg 20640

ttggccacag gtggttttgg agcgaatact aaaatgatac aacaatataa tacgtattgg 20700

gataacattc cagatgatat taaaacaaca aattctcctg caattacagg tgatggtata 20760

cgtttaggtg tgcaagcagg cgctgacatc gtaggtatgg gattctctca aatgatgccg 20820

atttctgatc ctaagacagg tgcattattc acaggattaa tcgtgacacc ttcaaacttc 20880

gtcttcatta ataaggaagg ccatcgtttt gttaacgaat ttgaaagtag agatgtatta 20940

tctaaggcag cattagaaca aaaagacggc atcttctata ttattgcaga cgcaaatatt 21000

aaagcgctag ctatgaatac gactgaagac aaaataaatc aagaattaga agatggcact 21060

ttaataaaag ctgatacctt agaagtgtta gctcaaaaat taaacataga tgcaactact 21120

tttgtgaata cgattgaaag atataatact ttcgtagaac taggacatga tgaagatttc 21180

aataaaaatg catttgattt aaaaatcgaa aaagcaccat tctatgcgac accacgaaaa 21240

cctgcaatac atcatactat gggtggtttg aaaatcaaca tacatgcaca agtcgttgat 21300

gtagaaggtc atatcattaa aggattgtat gcagctggtg aagttgctgg tggtatacat 21360

gcgggtaacc gtctaggtgg taatgcactg gcagatattt tcacttttgg tcgcatcgct 21420

ggtcaaagtg ctgtaatgaa ataaataaag aagaacagga ggacatttaa tgatgagtga 21480

acagtttaat aaaggaaaac aactcattgt agcgattatt ttagggattc tcacatattg 21540

gttgtttgca cagtctttct tgaatatcgc accacatgtt caacgtttct atgatgtaga 21600

tatgagtatt gttaacattg cggtcagtct cacatcactc ttaactggtg tatttatagt 21660

tgtagcagga ggtttatctg ataagttagg tcgagttaaa attacttacg caggtttgat 21720

acttagtata cttgggtcta ttgcactgat tatatctcat gcgccacttt tattattggt 21780

aggtcgagtg ttacaaggtt tatctgcagc atgcctactt ccagcaacta ttgcgttaat 21840

taattcattc tttcatggtg aagaaagaca aaaggcacta agttattggt catttggttc 21900

atatggtgga acaggattag catcactatt tgctggaatg atagccactt ttattgagtg 21960

gcgttggatt ttcgtgttgt caatcatttt ctcaattctt gccttaatat tgcttaaagg 22020

cataccagaa tctaaagatg aatctgcaca taataaaaaa tttgatatta ttggtattat 22080

tatttttgtc attatgatgt taagtattaa catagtaata acccaaggtg atagaattgg 22140

atggttaaat cctattattg taatattaat tgccatattc attgtgacat taatagcttt 22200

ctatatattt gaaaaacgtc aaaatgaacc tttcatagat tttagtttat tttcaaataa 22260

tatttatatt ggcacaacat tagctaattt aatggtgaat atggatatag gctcacttgc 22320

attatttaat atctatgttc aagatgataa acatttaaca ggagcacaag caggtttaat 22380

tacaattcca tatatgttgt gtagtttatt aatgatacgt gtaggtgaac gttttatgca 22440

aaatcgagga ccgcaattgc cacttatgct aggtcctata tcaattacag tgggtattat 22500

acttttagcc ttcacttcgt tgcctaacat gatttattat attgtagcgt gcatcggctt 22560

tatctttata ggtttaggat taggtttctt tgcgacacca gcgctatcta cagctgtatc 22620

taatgttccg tcagaaaaag caggtactgc atcagggatt atcaaaatga cttctacact 22680

cggtgcagca tttggtatcg ctgttgtaac tacaatctat acggcattgt ccgtcaatca 22740

tccagcacaa tttgcagcta caatcgcatt tatcgtgggg gcaggtttag tgtttatcgc 22800

atttattgcc gcgtattgtt taattcctag aaagaatgta gatgcttaaa tcgagagttg 22860

gcgcagtaaa tatcattgct tgaattttga tcgtatgttg aatgatgaat aaatagttaa 22920

acttgatgat agtcttgaaa taggagatgc gtttgaagcc ttttcttcag aattacgctt 22980

ggaagatatc aatcaccatc cgctgttgca aacttagtat cattcttagc ttctaatgat 23040

tctgactata ttactggtca attcatttta actgatggtg gattagtgta tagataaaaa 23100

gggatacttc gtcatgataa ttaatggata aaaaaagcgt ttaaatattt atatataatg 23160

attgtgtgct tatacaaaat gtcaactcaa tgcattcatg tgagagtggg atgatgagtc 23220

tgagacataa agtctctgga taagtcaata aagacaattt ctattgaaat aatatagaaa 23280

ttgtcttttt tataattttt tgattatttt cagattattg agctgctact tttcttatat 23340

tgagtgccat taatacaaat ccaagttctc ttttaatttt atttattcct ctaatggata 23400

ttcgaatgaa acccgttcta gtaatttagt tttgaaccca cctaaatgaa tatatgagtt 23460

atttttttata ctataaaata tattcagatt tcaatagtga cataaaatag gcatctttat 23520

atttaccttt agtgtagaat tgctctttga gtaatccttc agttttaaat ccttgtgatt 23580

cgtatatatg cacagctttt ttgttatctg tatcaacata tagataaatt ttgtgcatat 23640

ttaatatatc taatgcataa tttatcgctt tttcgaatgc gaattttgca taacctttac 23700

cactgaactc aggtttaata attatttgta tttcacaatt acgatggatg taattaattt 23760

ctactaattc aacaattcct atgacttgat tttcatcttc aacaataaaa cgtctctctg 23820

attcatctaa taaatgctta tcaaataaat attgaagttc cgttaaggat tcataaggct 23880

cttcaaacca ataagacata atagaatatt cattatttaa ttcatgaaca aaaagtaaat 23940

cactatactc taatgctctt agtttcataa ttccactcct aaaattttcc tatatatttg 24000

cattataaat ataaataacg aataagttat cattcactgt gaatactcta ttttaacaat 24060

tcactacata ctaattctca ttttcttatt attctctaat atctctgatt tattactctt 24120

actatgaaac atataaaact gtcacatttg tttatattaa taataaatac atcaatgata 24180

acaaaaagaa aataaaaaaa taaaaataga gcatccctca ccgtaaaagt gaaggatgct 24240

ttatttttat tagaatatag ttttcatttt gatagataat ggttaataat tgtctataaa 24300

tcattattaa ggaccgttgt aatacaatcg tcattcataa aattttcata gattttatcc 24360

aatatttttt catcttcaat tgctgtgaaa tgatatacta gccctttac tattgataca 24420

tactttcttc ttatgaggtt ttagcatttc catcacttgt tcgacacgtt ctacaacaac 24480

tggtcgcttc gcataagcac cataagctag aatcactgtg tcactttcgc taatcgcttt 24540

catcaagtga atatcagtgt gctcatcgta tggatttttg atgtgtttga gattttctgg 24600

acttttaata ttagagaata gatttacaag atatacagca ccgtatcttt ctgaattagc 24660

taattggttg aggatgagaa cagttgtgag atcgagtgat aatacaccat ctaaatgagg 24720

atacatcgtt atcactgtac aaacaggttt cttttcatcc cacgtctttt tgagtaagta 24780

tcggtgtttt tcatcgtcgc taaatatcgc ttctgtgtat atcgtacttt tgattgtatt 24840

catcatcttc actcctttta gtattcttct ggtaaaagca tcacataata aaaagtgtct 24900

acatcatctt cacgaatgac gtagactttc ttaggtactg cattttgatt ttttacatag 24960

tttgtatagt gatattccaa tttgtatgcg ggctgttctt gttcatgagt aatagagagt 25020

atattattat cttcttgtag tttaaaaatg tgtaggtaat ctgtatgacc ttgttgatct 25080

ctttctttta ccatattcca aagtaagatt tgaaggtcta gagatagttg ttcactaatg 25140

cctcttgtga tgtatcgatt gattttcata ctattttcct ccattttgct tttctttcat 25200

gatgtcaatc acttcgttaa tgactgtaac agatatttat gccactttga tccaattatt 25260

catagttatt ccctccttct cttagtaaat gacgttcatc gataatcgta tttttagtat 25320

ctgtgaggta tagaaagtcc atatcaaaat gatccagata accaatgctg atgagttggt 25380

tattggcgta cattagaaat ggatagatac ttagctcatg tagctcatta ttatagtagg 25440

tataagtttc aagtatgaga tgtgccagtg gagaattatt aataaaacgt tccggtagaa 25500

tattttctgc tgcttcttct aacgcttcac attcccatgt ttcgttgtta gatagttgga 25560

atagatgagt tatatattgt ttgagttctt gagtgattgt ttccatatca ttgcctccta 25620

gatagtgtga tagtgatgta gttcatatac atcaatggga taatatatat ttgatttgtc 25680

atttattacg aatcccggta agaataagag aaaattccga tcaaaaaccc gtgataaacg 25740

ttggtgtaac aaggaaatcc cggaatccca ctcattttga cgaacaatca actcattatt 25800

tataagtatt gatgataggg ttgggtttct gcttccttat atatattatt tatttataaa 25860

gaataacggg attttgggat tgtgcttgca caatccttct gcttcttcga gtaagcaaat 25920

cccattcatt tcccgttaaa aaatcatcgt gggatgttct ttagcaattt caatataagc 25980

ttcgtgtagt agtagataga taatgcaatt tgctaccata ttagataagt gttattgaaa 26040

ttgataaaga gaattcttat aaatagttat aaaaagtaat cagaagaaga aaataaagaa 26100

ctttctataa taaaaagtct tcactattaa attttagtga agacttttta cttaaaattt 26160

tattatttat aaaactaagc caataatact tgctgaaata attgatgcta atgttgcacc 26220

taataataat cttagggcaa aagaagctac tttatttccc tgtttactat caataccttt 26280

aatagctcct accataattc caacagtacc aaagttagca aagcttacaa gataaacaga 26340

tattatacct tgagtctttg gagatagaga agaggcaatt ttttgaaaat ctaacattgc 26400

tacgaactca tttgtaatta actttgttgc cataattgat ccagcacgta tagcttcact 26460

ccatggaata cccataatga aagcgatagg agcaaatatg taaccaatca atttttaaa 26520

atctagaccg actaaatcaa atacaatagt aattcctttc attaaagata taaaagctaa 26580

caacattatt gctacagtaa tagcaattct gaagccatca atagcacttt caccaatcat 26640

ttgaaagaaa gttgtttttt taatattttt ttctgattca gcttctaatt ctttattatc 26700

atttttgtaa ggattaatga cgctagcaat tatcaaagca ctgaaaatat ttaaaacaac 26760

tgcagtaact acatatttag gatcgatcat ttgcatatat gagcctagca tagccatact 26820

gacagcactc atacctgaag tagcgatagt atataacttt tcttttgaca agcttggaat 26880

aatattttta actgttaaat aaacctctgg ttgacctaac atagaagtag atattgcgat 26940

atagctttca agttggccca tatgagttat tttattaatt atccaaccta taattttaat 27000

tagaaaaggt aaaattttaa aatagtttaa cataccaatt aaaactgaaa taaaaactaa 27060

aggtagtaat acatttagaa agaaattata gccgttttta ttttgtaaat ctccaaaaac 27120

gaattcaata ccttctttac taatattcat taaaccttca aaaaatgtac tcattgcggt 27180

taataaatta agacctattg atgtattcat cataaataaa acaagaacaa tttgtattac 27240

tatcataatc attggttttt ttatatctat agcttttcta ttaaaactaa aaatatatgc 27300

aactaataaa gcaaaaacta ttcctaaaaa tgtgaagatt attgacatag tcaagcacct 27360

acattttgct gaaagctaca gctattttag ctcctaaaac agcgttattt tctaccaact 27420

taatatttgt ttctaagctt ttgccatctg ttttttcaac gattgtttta agaaggaaag 27480

gtgtagagtc ttttccacta attccattgg caatagaatc tttaacagct tcttcaatta 27540

tagtgtcaat ataatcttta gaaagttcat gttctttagg aacaggattt gcaatgacca 27600

ttccaccatg taaatttaac tccattttgg tcttgaaaat ttcagctatt tctgtaggtt 27660

cctcaactga actatttagt tttaaaccgc tttctctagt gaaaaaagca ggtagttcag 27720

ttgtttgata tccaattacc ggaacacctt tagtttctaa gtattcaagc gttttaggca 27780

aatctaaaat agattttgct cctgcgcata taactgtaac atttgtatgc gctagttctt 27840

ctaaatctgc tgaaatatcc atactagttt ctacgccctt atgaacacca ccaattccac 27900

cagtgacaaa aaatggtatc ccagctaatt ctgaacaaat catagtagaa gcaacagttg 27960

tagccccaat ttttttagta gcaactattt ctgctaagtc gcgacgagat actttagcga 28020

cattaggact atttgctaga gtctccaagt cattttcgtt taaaccaatt tttattttcc 28080

catccattat ggcaatagta gcaggtgtgg cgccattgtc tcgtattatt tgttcaactt 28140

tttttgccat ttcaatgttt tgtgggtagg gcataccatg agagattatt gtagattcta 28200

aagctactat gggactattg ttttttttgg cttcttcaac ttctgtagaa aattcaagaa 28260

atttttttcat ttttaaattc ctccaattca atttctaatt gtttttgatc taaattttgt 28320

ctaacagtgt aattagtttc aacagttttc ttcgagttta acaaagctaa gttaataatg 28380

tcttcattag tataattttt tagccaacta tatattactg cgctagtaaa tgaatcacca 28440

gcaccagtta catcaactac atttttagat gaaattactg agtgaaatga ttctgaacta 28500

ttatctttga aaatgagtgg ttttagtcca tttgtaataa tgacgttaga cacaccaaga 28560

ttaattaatt tctcagcagc aatactcatg tcatcattac tatcaatttt catttttaaa 28620

aaagtttcag tttcatcgcg attaataatt aaccaatcaa ttgaatgtag tgaattaggc 28680

atatggatca ttttaggtga agatactggt acaattacta acttaatatt gtttttttct 28740

gcgaatgaag tgaggtactc aatagatttt ttgggtaaat ttaaatctat tactaaacat 28800

ttagctttta ataagataga tgactgcttt gacagaaatt caggtgtaat ttcctcgtaa 28860

atattcatat ctgcaaaacc atactgcata tcacctttt tatctataat agctgtatat 28920

gatcctgtac ttgcattttc tataggttga atataatttg tattcataaa aggctcagat 28980

agtttttaa tcaaactcca ttcgctatca tttccagctg ctgttaaaaa tgtaacattt 29040

tcaccaagtc tccctaagtt ttcagcaata tttcttgcta cacctccaat cgattttgtt 29100

gattttactg gattggaagt catagattgt agatcagtat cacaataaaa ctttctatca 29160

acattagcag ctccaataca gattatggaa tcttcctcat taacaatgta tgcttttcct 29220

aatatatatt tctttttttat taaaccagat attatattag ctgtggctgg ccttgataaa 29280

ccaatacgtt ctgctaactc ttgttgagta ataaaaggat ttttttaat ttcgtttaaa 29340

ataaattgct ctttatttttt caaaatagca cctcctttta aagaaacaat agttcacaaa 29400

caaaacattt gtttgtttca ttcaaagaga taataacata tatacatttt tgtgtaaacg 29460

ttttcttgaa gggttttatt ttttattcta ttaaaagata attatactca ctttatttac 29520

taaataataa gatacagtgt ttgggacaaa aataagaggc tgggacaaaa taaaatgtct 29580

cagcctattt tatcattttg tcagtagctg actgatttga aaatgcgctt gaatcaagct 29640

tttttcaaac ctagtcatcc ttgctggggt aggacaacga aataaatttt gagaaaatat 29700

tatttctgtc ccactcccaa gtggatggtt attaaatagt tgtttgataa tctcattatt 29760

cagtttgagc gtaacgtaaa attgcttctt atgatgctca tcttttctta tgtcaatacg 29820

atcgatgatt gttagatata gtgatttcaa ctgagatttc tctaatttgt ctatatcttt 29880

gaatattgct tgtaatacat tagcaatcat atctgcatca taatgtggag attccgcttg 29940

tttatcttgt tctagttgat ggatttgatt atttatttga ttcaattcat cttggtagct 30000

aagtattgtt ggtttcaata cgctatctaa gtcaggggag tcttcgatag ttttcgttag 30060

tgtatgcatt ttggctttta tttcttcaca ttgtgactgt ttataggcaa tatcatgatt 30120

caaagaagat acgtctattt gacttctttc atttaccttt tcaaccaatt gcttcaatac 30180

ctttttgctt ttgataattt ccaatatctg atccataaca tatttttcta gtacatctgc 30240

tctaacacta ttggctgaac aaacttttga acctttgttt ctaaaattac tacatgagta 30300

atatctgatt cttttcttag tgccatcttt taatgtattg gtcgtattac ttgcggccat 30360

tgctgcacca catttcggac attttacaat gccagtcagt aggttagtcc ctttaccgtg 30420

aacttgtggt ttcttgcgac tctcttggcg tttaaactgt actttatccc gtaaagcttt 30480

atcaataata ggcgcatgct taccatctgc aatgattggt tcttcattga gtccttttcg 30540

ttttttatcg ctccagtgtc tatacttcgc aaactgtatc tttccaatgt aaaaagggtt 30600

tgagatgatg taagtaatgg acgaaatgct aaaaggtttc gctttcttag tcacataacc 30660

tttatgattc aatgcgttcg caatcttacg ataaccatga cctttagcgt atgactcaaa 30720

aatatattta acaatattcg cttcatgttg attgatcatg agctcttttt tactgtcagg 30780

tactttatca tagcctaaag gtaaattacc ttggtaataa ccttcaatcg ctcgttgtct 30840

ttgaccgtta taaacgtttt ctataatcgt atttctttca aattctgcga agctggctaa 30900

aatttggagc atcaatttac ctgttgagct ggcaatttct atttttcag ttagactaaa 30960

aaattcgaca ttgattttat acaatttttc cacaatattt aataagtctg aggtatttct 31020

agctaaacga ttggttttgt agaccataat acagtctaac ttaccgtcat ttgcatcttt 31080

taacatacgt tgtaattctg gacgttccat tgttttaccc gatataccac gatcggtata 31140

ttcatcaacg acctcataac cttgaaattg acaatactct gtaagttggt ttaattgacc 31200

ttgaatactg taaccatttt tttgagtttc agtcgataca cgggcatata atccaatatg 31260

tttcttttta agttgtttca tattacttcc atcctttcag atattattaa atgtgattgt 31320

tcaacgatat tgagtggact atttttaaag tagatacctt gtaattgttt agtttgagtt 31380

attgtaatgt cattaatcaa aggtgctatg gtttctaatg tcaatttatc tttgattata 31440

tgacggattg tttcggtaat ttgtttatgt gaataattag agtaacgttc ttctctttca 31500

aagccagcag ataaacgttt aaacgtttcg atatcaattt gattttgtgc cagtttttcg 31560

atgagttgtt cttgtgttaa atgtgattgt ctgtgtttca tttgttggtg ttttagagct 31620

tttagtattg tcttatttaa tttgtcatga aatgacggat tatcacaata tgctttacaa 31680

gtgttgagga cctctgtttc aatcatttgc gcattgatac ctttaaatgg acacgtatga 31740

taggcttcat tcatattctt aggacaaaca tagtaacgta gggagtgttt ctctttttt 31800

atcgttaagt ttgttaatgt tgattgacag taaggacatt tgattcgtcg ctttagctta 31860

tttctagaat tggatcgatt gtgttgctta tgaatacgac gctcttgtgc ttcttcaaac 31920

atatcaacat caataatagg cggaacgata tcattaaatg tgccatattt attaatgaca 31980

cgaccgcaat agttaggatt caagagaata tttctaactt gatagggctt acgaggaata 32040

agcttaggat ggttatctaa atgttgggaa atctttttgt agcctagacc gtgtaagtac 32100

cagcgataaa ctgctttgac tgtataggct tcttcttcat gtacaacaaa atgaccttgt 32160

ctataacgat acccaaaagg tgcatgagtt gtgattagct taccttgttt ggctttttct 32220

cttgtcccat ttcgtgtctg ttcactgata ttattagact ccatttcagc tagactcatg 32280

agtatgttta agcgaaagca atcaaacttt taagacaaat caaaatatcc atcgttaaca 32340

ctgatgattg tgacgtgatg ctttttacaa atttcaaaga attgtatggc atttttcaaa 32400

ttacgatgga gccgatttaa gcgatagcaa cataatactt tacattttcc agatgtaatc 32460

atttctacca ttttttgata acccgaatgt tttgtatgtc gacctgtttt cttatcatca 32520

taaaatgtta cattggacca tccatattgc ttagcagtgt ctataatgag ggatttctga 32580

gtcgctaagc tttgttgttt gagtgtactt tgacgtacat aagcaataac ttcttccatg 32640

ttatacacct ccaaaaagat aatatatatt tgtgagtaaa ttcaaataaa ggtccaacgt 32700

gcttctaaca cgttggaccg caatgattaa taatcatgat tttatttttc aataaagtaa 32760

ttagttagtg atgtaataag tttatatttc aaatatgaat tatttaataa atattgtaaa 32820

atttcattaa gttatttatg ataaatataa attaaaacgt tatttttatt ttatagtgaa 32880

aggattaata ataatgaaac acatgcattt atctgtttta gcttatagta caggtctaca 32940

tcctgcctct tggagactac cacattcata tgttgaagaa gttggagata ttgattttca 33000

aattaagtta gcgaaattag ccgaaaaagg taaactggat gcattcttct taggtgatgg 33060

tcaatatatt tctggagaag aaacaggaca tatttcttat tattttgaac cactaacggc 33120

gcttgctgcc atttcacgtg aaacacattc aattggtttg ataggtacca tgtcttcttc 33180

tttttatgaa ccttatcttg cagcacgtat gctttcaagt ttacatcaaa tatctcatgg 33240

gcgtattggg gcaaatattg taacatctca atttgattta gaagcacaaa actactcgat 33300

gcaagcttta ccacatcttg aaaaacgtta tgaacgagca gatgagttta ttaatgtaat 33360

gaaaaaatta tgggaatcat ttactgtaga ggctatcgtt aatcacaaaa actcaggaat 33420

tggcttaaat catcagtata ttcatccact tcattatcaa ggaaagtatt ttcaagttgc 33480

cggagctatt aacattccaa cacccaaata tggtcgacca cgtttatttc aagcaggaac 33540

ttccattcca ggtcgtgatt tagcagtacg acacgtggat gctatttttt caattgcatg 33600

gaatttacaa gatggccaac agttcagaca agacattcat caacgtgcca tggaagcaaa 33660

tcgtcaaccc cctcttgtat taccaggtct cactgtttat gttgatgaga atgaagagga 33720

tgcttataaa cttaagcaac aattagatga tatgattcca cttgaaaaac gcaagaaaca 33780

attatctaaa gctattggtc tagatataag tgaatggaaa ctagatgaag ttataccgtc 33840

cttaccagat tatgaaacat tatcgaaaaa ggtcgtaaaa tcagtatatg aagccataca 33900

acgtgccatt aagactgagc atcttacatt aagagattta ttagatcgtt ttggtacttg 33960

ggtagggcac aagacaatag taggaactcc cgaacaagtg gcggatgaaa tgatcgaatg 34020

gtttaaaaaa gaagcgtgcg atggttttat gttgatggcc ccgacgtatc cagaatcatt 34080

tgaaaaattt attcatttag tgataccaat tcttcaagaa cgtggagtgt ttagaagaga 34140

ttatgagagt catctgttaa aaaatcattt gggtataaaa aactaaaaat aaaccgcatc 34200

attaaccgat acgcagaagc ttatcacaaa taaaactaaa aaattggtaa gtagatgaag 34260

agtgtaaatc agagtaatta ataataatgt atcaaattta aacaaagggg tctttaagta 34320

tgaatttaag aggtcatgaa aataggctta aatttcttga gaaatatgat gtgacaccta 34380

tatcacattt aaaattatta gaagatcaaa agaaagacgg tgaaggtggc atactgacag 34440

atagctatta ctgtttttca tacagcttaa aaggtaattc taaaaaagtt ttaggtacgt 34500

ttaattgtgg ttatcatatt gctgaagact tactaaaatt atcaaatcaa gacaaattac 34560

ctttgtttaa cccgtttaaa gtaattaatg aaggtaatca attgcaggaa gtagcgaata 34620

aaggtaattt agatattaat aggcaaagaa aacagtataa tgaagtagct ttacagcttt 34680

caaatgctat aaatttaatc atattttgtt atgagggtaa tattaaagaa ccactttcaa 34740

caataaaaca cgaaaccgat aaatattatt atagtaaacc atatgatcgt aaaattaaag 34800

cagtaaatac tattatcggt aatttgtttg ataagaaatt agttgagaaa atatctgaaa 34860

ttaatggaaa tcaggaaatt agaaaatttg atttttcatt actcactgat gtaattaaaa 34920

aacacgagat acaaaataat tttgaataat aagtggatgt aagaagatat taaatattga 34980

agctagtatt aaggatattc aaaatattat aaattatttt actgaatgtg gcgtgaataa 35040

taagtattat gagagtgaga aaagaagtag tctaggaaag ttggtaagtg ttattggaca 35100

cgaagaaaat gttagaaatc caaaataacc tgaattagat tttgaagatt tcatatctga 35160

atgtataagt gaaaaacaaa aaagaattag acataaacta ggaggcagag tatgatgtca 35220

tcgaatgaga tagtaaaaaa attatattca aaggaagata ttcgcataga attagagctc 35280

actccacata agtttaaaaa aagaatggaa accattacta agcttttaa aattgatatg 35340

aaaatttttc acaattacaa aggccaagat aaaaataatc aatacacatt taatggcgtc 35400

gcaaaagaat taataaaggt gctactaaaa agtattgact attacccagt ggatataaat 35460

tccaaagctt ttaaacaaaa cggaaaatct aaaaaagaga tgattgaaaa tataaaagca 35520

tctagttaca tggaatatac ttatcagtta atgaaatcta tcaatgaaat taaatacaaa 35580

agattaattg cagacataca tacgaaagac gtatatcaaa atactaaagc atggttaagt 35640

attggtgaat caattaataa aaaggaacaa gaattatacc aatatatgac acctttgccc 35700

atacataaaa gaattgagct acaaaatgaa gtattaaaat ctatagatga aacgattttt 35760

caatttatgg caaaagaacg tagaaacacc caaataaaag aaaataatga attagaagaa 35820

tacacaaatg caataaaaga aggtaagtat cccaaaaagg attacgaatt aaaccaccta 35880

ttatacgaga aaaatgaaat gccattagat agcttaatta aagatgtatg ggactatgat 35940

gaaacagagt atatggagct tgattggtta attgcagata tgctaaaacg ttcacaaaga 36000

gctgatagta attttatcga aagtttagat aagaaaaaca aattgagaaa aaatattcat 36060

aaagaatcta ttaaaagtat atctaagttg attgatcatg aattggttaa aaataataga 36120

aagtgggacg tagcagaatt ttataagaat caacaatttt ggaacaatag tacattaaga 36180

cgtagctctc aaaattttgg gattttatat tacaaaaaaa ataggttaca tacaatgaac 36240

caaagtaata ttattaaaaa aataaaaatt atcgaggcta atatagagag tgcaaataaa 36300

cagccgagct atttagattt ctttttacaa gctaagtttg acttagagag gttagaaagt 36360

gaattaatga aatatagcat taatccgaaa tattttaata agataaatag tgattacgta 36420

agatatgcta ttgaagatgt acataatgca attattaaaa ttgacagata tataagtaac 36480

gataaagtga agttaaacta ttcaaatgcc caaatggtag tgagaaacga aataacagaa 36540

caaggaagct actttgtaaa ccaggcattg aatacactgt caaaggtaga agagtctggc 36600

tacagctttg ataatttcat ggtaggacct ttcgtagaag actttaaaaa agcaattaaa 36660

aatcgtagat ttggttctaa aacagattaa gattgttttt tagagttaaa cataagaata 36720

aatttaaata aaaaaacaaa cccaacaaac cttgttatat caacgtttgt tgggtttatt 36780

tttttggtga ttttttaata actcaattac accttagtat aatattgcta aattatacat 36840

tgtcgaggcc aagacaaata aaaaattaaa gcaaaagtta gcgcgcaact ttctgccata 36900

aaggagcaga tttatatgga gacaagagac aaactgatgt cattgactca aagtgacaaa 36960

acacaacaat ggttaatgaa caaatcatct aaccaaactg atgttcaaca attgcagcaa 37020

caattcagtc agcagcttga ccaacaatat aatgcacttt tagctgatga acaagctaaa 37080

ttagatcaat atgtggaagt ccatcatgaa ttggaagcat taaaggaaga aattgaagca 37140

gaacctatta cacttaatat tgataaatta cccgatatca aagcgataat gcttgaaaag 37200

gctaagaaca atgaacattc tgataaaatc gaacagttat ttgatagatt agatcagtca 37260

ttaaatggta cgaatcgatt atatacacaa ttatcgttga ttggtacacg aacacatcga 37320

atcacaacta aaaattttaa tattcagggc ttacctaaag cagtccaatg cacgatttta 37380

ccttcaaaat ataagaaggt atatactgtc gattttaaat catttgaacc ttcggttgtt 37440

gcgtacatga ctcaagattc aaaactgatt gacttttttaa atcagaaaga cggattgtat 37500

gacgcattgc taagtgaatt aggcttatca tatgagctac gagtatttgt taaacgtgca 37560

tttattggtt cgtttctatt tggtggcaac ttcgatagcg ctaaattcaa gttgaaacaa 37620

tatgtaagtg aagtgcaatg gttggatgcg gtcagccaat ttacaaaagt cattgaactt 37680

aagaaacaaa tcgaagagca taaaaccatt cctatgcctt atggcattga acatgatatg 37740

agcgcatttc aaggtagcag tattatagca ttttatgtac aaactgtagc gagttatgtt 37800

ttcaagcata ttctgttgga agtgtacaaa gcacaatgcg aacaaaaaac gtttaaaatt 37860

attgtaccta tacatgatgc gattatgatt gaatgtgatg atgaagaaac ggcgcaaaat 37920

gtatcgcagc taatgaaaag taaggcaaat gaattgttta atggtgagtt tgcacacgtg 37980

acagtggaag cattaggggg tgtaaacaat gaataatgat agaggaaaaa gtttacaaat 38040

tccaaaaagc atagccctaa ctgaaggata tatatatgtt gcttcattat attctgtaac 38100

ccaaacagat ttcaatgggg atgttaaaca tcagtttacg tatgaaatag aggtagatca 38160

gaaagttatt tatgtaaatc gtagtatttc gaaagatgct accgataagc agttgtcaat 38220

caatgattgg cttaaacatc acagtaatta tagctccact cataagaatt atgcccccta 38280

cattgataga aagcatttaa tccatgtagg ttgctttaac ggtaattatt atgtacaaga 38340

tgtagcatca ttaaatgaag atggaggatt attataatga aacatatatt cgatatgtta 38400

acaaaaataa taagtgtggc taaggaggca gtcgaccgtc aaggtctgat tgctatccta 38460

accatgagta ttggcaatga tgatgaaata gaagaaacgg tacaaggtga aacagtgtat 38520

aacgaactta tcgatcagtt acgacttaat atcccaaaag atacggatta tcgacctaac 38580

atctatagct atttcggtat taagaagaaa ccaagtgaca caatattaat agacatgatg 38640

ataaaagttt ttcatatcaa acgcttcaat tcggaactgt atatattcaa aattaacggt 38700

tggcaaaagc taagtgaaga cgagttaaaa ggatttgtat caaaaatgat acaagtactg 38760

ttaatcggtt atacgcctac gcaaagtgca ctgaataatg tggttgaagg tctgcaaaaa 38820

tcaacagatg tagaagaact tgtcgagaac gagcactata ttggttgtgg tgaaaatatg 38880

ttcgatctta atacattcaa agtcgttaaa aattcaatcg acatttttcc aaaaacgcga 38940

ttaaatttag cgttagataa aaatgatgag attactgacc agatacccaa acatttcaat 39000

caatatatgt tagaacttgc aaattatgac gatgatttac aatactttct ttttcaacat 39060

acagcagtgt tattgacggc tgatactaaa taccgtcgag gtctcatatt atatggtgga 39120

gctaagaatg ggaaatctgt atatattgaa ctagttaaat catttttcta tagcaaagat 39180

attgtgtcta agccacttaa tgagcttgaa ggtcgctttg acaaagaaag tttaattgac 39240

aaaagtctaa tggcaagtca tgaaataggg caatctagga ttcaagaaca gattgtaaat 39300

gacttcaaaa agttattatc tgtagaatca atgcatgttg atcgtaaagg aaaaactcaa 39360

gtggaagtca ctttggattt gaaacttatt tttagtacaa atgcggtact taattttcct 39420

cctgaacatg cgaaagcttt ggagcgtcga attaatatta ttccatgtga gtattatgtt 39480

gaaaaagcgg acacttcatt gattgataag ctccagagtg agaagaaaga aatctttctt 39540

tacttgatgt atgtgtatca acagattgta aaagcagata tcgagtacct tgaaaatagt 39600

cgtgtcactg aaattactca cgattggtta aattttggat atgaatttgt ttctagcagg 39660

tccgcaagta ttgcaaatca gaaagagtgt attaatttac tcagaaaact tatagaaatc 39720

aaaccaggat cacgaatcaa agtatccgag ttaaataaag ttattaatga agaaataaag 39780

gtgagttctc aggttattaa acagttaatt caagaaaact ttgatactca aaccaaacta 39840

tacaatggct acgattattg gattgattta ggttggaaag aagccgataa aaaagagatt 39900

catgatattt cgaaaaaaga taatattatt tcattagaaa aaaatgaaaa tataacagac 39960

gatgaggcat tagatgaaga gaatttggac tttgattggg aggactttga cgatgaataa 40020

tgaacaaatt gaagcatttg tagaagtact tgtgcctatc atagaagaac gtatcaataa 40080

aggtaagtaa tctaattacg tactacaggc agttgcctgt agtactcata tgattaagtg 40140

ggaaaagtga taaaaatgaa acgaaattat aaatatatat tatccatatg atgttacaag 40200

accgatggtc tgtagcaata atctaataaa aggagcggta tgatatgaag ggtaaaattg 40260

cactttattc acgcgttagt acgtcagagc agtcggagca tgggtactca atccatgagc 40320

aagaacaagt actcatcaaa gaggttgtga aaaactatcc aggttatgac tatgaaactt 40380

atatcgattc tggtatatca ggtaaaaata ttgaaggtcg accagcaatg aaacgtctat 40440

tacaagatgt taaggataat aaaatcgaaa tggtgttaag ttggaaattg aatcgtatct 40500

cacgatcaat gagagacgtg tttaatatta ttcatgaatt caaagaacat ggcgtagggt 40560

ataaatcgat ttctgagaat attgatacat ccaatgcttc tggtgaagtt ttagttacaa 40620

tgtttgggtt aataggatct atagaacgcc agactttgat ttccaatgtg aaactttcta 40680

tgaatgctaa ggcacggagc ggagaggcaa tcaccggtcg tgttttaggc tacaaattat 40740

cacttaatcc attgacacag aaaaatgatt tagttattga tgaaaatgaa gctaatattg 40800

tacgtgaaat tttcgattta tatttgaatc ataataaagg acttaaagca atcacgacaa 40860

ttctaaatca aaaaggatat cgcaccatta atcaaaaacc attttcagtg ttcggcgtga 40920

aatatatttt gaataatcca gtctataaag gttttgttag atttaataac catcaaaact 40980

gggcagttca gcgaagaggt ggtaaaagtg atgaaaatga tgtgatattg gtcaaaggta 41040

agcatgaagc cattataagt gaagatgtat ttgatcaagt tcatgagaaa ctagcttcta 41100

aaagttttaa accgggtcga cctattggtg gagatttcta cttacgtggt cttattaaat 41160

gcccagaatg cggaaataat atggtatgtc gacggactta ttataaaacg aaaaagtcaa 41220

aagaacggac aatcaaacgc tattacattt gttcattatt taaccgttct gggagttctg 41280

cttgtcatag taattccatc aatgctgaag tcgtcgagcg tgtaattaat gttcatttga 41340

atcgtattct gtctcaacca gatattatca agcagattgc gtcaaatgtg atcgaagaac 41400

tgaaacaaaa gcatagtaac caaacagaaa ttaaatatga cattgatagt ttagaaaaac 41460

aaaaagctaa gcttaaaaca caacaagaac gattattaga tttatatttg gatgaggaaa 41520

tagatagcga aatattgaaa gataaacaaa gtcaaatgaa tcaacagtta gaagtattag 41580

atcaacaaat taaagaagcg caacaagcaa atcaatcaca ggatgaaata cccaattttg 41640

ataaattaaa aggacgactc attttgatga taacacgatt cagcgtgtac ttaagagagg 41700

ctacacccga atctaaaaat caacttatga aaatgttaat tgattcaatt gaaattacga 41760

cagataaaca agtaaaactt gtaaggtata aaattgacga aagtcttatc ccgcaatcca 41820

ttaaaaagga ttgtgggaat ttttttatgc ctaaatttaa ctttgtgata aatgtcacaa 41880

ataaaaatag gattgaaaat ttatcacttt taccactttt ttagagtgac aaaagtggag 41940

gagttttgaa atattttata aatatatatt ttatttatgg agtacacatt attaattaag 42000

gaggtcatta tagtgacgct aagcaaacaa cttaaaacgt atattactga acgatttaaa 42060

ttaaattatc aagaaacttg gagttgtgaa accgtagacg cggtggctga agacgtatta 42120

cctgaaaaat atattaaaaa tagtccactc gaacataaaa tattgaacac atttacttat 42180

tataacgatg aattacatga aatcagtatt tatcctttct tatactattt agataaggaa 42240

ttaatagcaa taggttattt agataatttt gatttagact ttatattttt aaatgacact 42300

catcaaatta ttattgatga acgctacttg ttacaaaaag ggggcgagta attatgaact 42360

ggattaaggt tgctcaacta tctgttacag tcattaacga agtgattgag atcatgaaag 42420

aaaagcagaa tggaggaaaa tagtatgaac atcaatcgat acatcacaag aggcattagt 42480

gaacaactat ctctagacct tcaaatctta ctttggcaca tggtagaaga aaaagataac 42540

cagcctcata ccgattacct acacattttc aaactacaag atgatgataa tatgttgtca 42600

attacacatg aacaagaaca gcccgtatac aagttagaat atcactatac aaactatgaa 42660

aaaaatcaaa atgcattacc taagaaagtc tacgtcattc gtgaagatga tgtagacgtt 42720

ttttattatg tgatgctttt accagaagaa tactaaaagg agtgacgttg atgaatacaa 42780

tcaaaagtac gatacacaca gaagctatat ttagcgatga tgaacaacac cgctacttac 42840

tcaaaaagac gtgggatgaa aagaaacctg tttgtacagt gataacgatg tatcctcatt 42900

tagatggcgt attatcactc gatcttacta ctgttcttat cctcaaccaa ttagcgaatt 42960

ctgaacaata tggcgctgta tatctagtga atctattctc taatattaaa acaccagaga 43020

atcttaaaca tattaaagaa ccttatgata aacacacaga catccatttg atgaaagcaa 43080

ttagtgagag tgatacagtg attctagctt atggtgctta tgcgaagcga cccgttgtcg 43140

tcgaacgtgt tgagcaagta atggaaatgt taaaatctca taaaaagaaa gtcaaaaaac 43200

tcataaaccc agcaacgaat gacattatgc atccgcttaa tcctaaagca cgtcaaaaat 43260

ggacattgaa ataaaggagg attatatatg aaccataaca ctacacaatc agattggcga 43320

atagttgcta gttgcttagc atcacaaaat tatatatcta tcgtaaaagg attagtacat 43380

cattttacag cgattgcaga tgaggaaaca cttgataaaa tctatgaaga ttttatgaat 43440

gatgactcta taaccacggt acttaataat gatttacaga caattattaa ccattatctg 43500

tcaaaatgaa aattatattc taataaaact agagcatcct tcacttagtg gtgagggatg 43560

ctcttatctt tatttttttg attaatattt tactcaggaa attttacatt aaatttaaaa 43620

gtattataca attaaattat tagtttctag tgaagaaata cgtattaaat ttattattaa 43680

ctgtagcaca ttatcaagca attgtactgt taaaagggat tggtcattca atgagacaat 43740

atgtttatca aaatgatatt aatttaataa actctttata tgagagtgat ttttggaaga 43800

taataaaaga agacgcagca tattatcata aaaataataa attcaagaaa gataatgcta 43860

tacgaatact tgaatctcta attaaatcaa tatatgtaga tcctgatggt tttgataaag 43920

ctcttgcagc tgaaatgcaa gatttctata acaaaatgca aaaatcacaa tacatcaaag 43980

agtcttatta tcttagtatt aatcaccaga aatgtagtct agatgctctc atcggatgga 44040

aaccactttt ccgatttaga aatggtgata aaaaatggtt agatgactta gaacttatta 44100

gagggaatag aatgggccat ttagcttttc cagtacaaaa aaattctctt aatcaactaa 44160

gaggaatttt attaaaagat agaattgatt atactctatt tgatattaaa ttattttatg 44220

aaaatgcagc tcatttaaaa ttacagaaag cttatgaaca agaaccaaca cgaaaatggc 44280

taaaatcatt tggtaccttt aatcaattta tagagcgtat gcaattgaac tacttcgtct 44340

ataaagatcc aatcacatta aaatatgatg ttattgactt atcattgcca tataataatg 44400

ataagtcaca ttgcttaaaa gaaattccaa aaaaaattaa ggtcgaagaa acttatatta 44460

caaatatact taattatatt aaaaaatatg gcgaaaaatt atctactatc catgtggact 44520

taatgattga ttatcatgta taaataaaag aacaaccctt caccgtaaag tgaagggtgc 44580

tctttagctt ctttcattat ttacaaattt catctctttt tttgatacaa ttctaactta 44640

attaaatata atgctactct ataaatctcc tacatattta tagctcatat cctctaaaat 44700

aaccagtgag tatctataca tatattttat ggtaacgtag gctttttcta gtttacattt 44760

tattgattac tcgctataag ttctgtttgc tcgacaactt gttcgacagt catcttttgt 44820

aaatcaggag gataaccata ttttttaagc aatcgtctaa aagttaagcg catttttact 44880

tttgcactat cgcgtttaga taaattgacg cccatatttt ttcactgttt tagttagctc 44940

atgagcaatt gcacgtagtt ctttatcacc attgcctctt tagttatttc atgtgatgaa 45000

gctaatgcat aataaaatac aaggtcctta attactattg agatgttaga aaatattaaa 45060

gaaagcaatt tttactaaat tttaaggaca ttacgttttt ataattaaat taaagttagc 45120

ttttataaga aagatattca tacttaataa aaaaaaacta ttaaattttg tttgactaaa 45180

gttggatagt acattatagt ataattacaa tactaatgtg tattgtactt atacctaaaa 45240

ttacagaatc tactaaaaca aggcaaaatg ccgtgtttat ctcgtcaact tgttggcgag 45300

atttttttat atcaataata taaatttcat gtatctggga gtggtcatat gaaaaggaat 45360

tatatcttag gattagatat cggaattaca tcagtgggtt atggaattat tgattatgaa 45420

actagagatg tcatagatgc gggcgtacgt ttatttaaag aggctaatgt tgaaaataat 45480

gaaggacgac gatcaaaaag aggtgccaga aggcttaaga ggcgtcgtag acatagaata 45540

caaagagtaa agaaactttt atttgattac aatttgttga cagatcatag tgagctaagt 45600

ggaatcaatc cttacgaggc gcgcgtaaag ggattaagtc aaaaattaag tgaagaggaa 45660

ttttctgcgg cattgctaca tttagcaaag cgtagaggtg tacataatgt taatgaagtg 45720

gaagaagata caggtaatga attatccact aaagaacaaa tttcaagaaa tagtaaagcg 45780

ttagaagaga agtatgttgc agaattacag ttggaacgtt tgaaaaaaga cggtgaagtg 45840

agaggttcga ttaaccgttt caaaacatct gactatgtaa aagaagcaaa gcagttatta 45900

aaagtacaaa aagcatatca tcaacttgat caatcattta tagacactta tattgattta 45960

ttggaaacaa gaagaacata ttatgaggga ccaggtgaag gtagcccatt tggatggaaa 46020

gatattaaag aatggtatga aatgttaatg ggacattgta cgtatttccc agaagaatta 46080

cgtagtgtga aatatgccta taatgctgat ttatataatg cgctgaatga tttgaacaac 46140

ttggttatta cacgagatga gaatgagaag ctagagtatt atgaaaaatt ccaaattatc 46200

gagaatgtct ttaaacaaaa gaaaaagccg acgcttaaac aaattgcgaa ggaaatcttg 46260

gtgaatgaag aagacatcaa aggctatcgt gtcacaagta caggtaaacc agaatttaca 46320

aacttgaaag tttatcacga tatcaaagat attacagcaa gaaaagaaat tatcgagaat 46380

gcagagctac tcgatcaaat agctaaaata ttaactattt accagtcatc agaagatata 46440

caagaagaat taacaaacct aaattcagaa ttgacacaag aagagattga acaaatttca 46500

aacttgaaag gttatacagg aactcataac ctttcactaa aggcaataaa tttaatatta 46560

gacgaattgt ggcatacgaa cgataatcaa atagctattt tcaatcgttt gaaacttgta 46620

cctaaaaagg tagatttaag ccaacaaaaa gaaattccta ctactttagt tgatgatttt 46680

atactgtctc cagtagtgaa acgttcattt atacaatcta ttaaagttat taacgctatt 46740

attaaaaaat acggtttgcc aaatgatatt attattgaac ttgcgagaga aaagaattct 46800

aaagatgcac aaaaaatgat taatgaaatg cagaagagaa atcgtcaaac gaatgaacgt 46860

attgaggaaa ttataagaac gacaggtaaa gaaaatgcta aatatttaat tgaaaaaatt 46920

aagctgcacg atatgcaaga agggaaatgt ttatactcgt tagaagcaat ccctctagaa 46980

gatttactta ataatccatt caattacgaa gtagaccata tcattccacg ttctgtttct 47040

ttcgataact ctttcaataa taaagtgttg gtgaaacaag aagaaaatag taaaaaaggt 47100

aatagaacgc catttcaata tttaagttct tcagattcta aaataagtta tgagacattc 47160

aaaaagcata ttttaaatct tgctaaaggc aaaggtagaa tctctaagac gaaaaaagaa 47220

tatttgttag aagaacgaga tatcaatcgc ttcagtgtcc aaaaagattt tattaaccgt 47280

aacttagtag atacacgcta tgcgacaaga gggttaatga acttattaag atcttatttt 47340

agagtgaata acttagatgt caaagtgaaa tcgattaatg gcggattcac aagtttctta 47400

agaaggaaat ggaagttcaa aaaagaaaga aataagggct acaaacacca tgctgaagat 47460

gcactgatta ttgcgaacgc tgattttatt ttcaaagaat ggaaaaaact agataaagct 47520

aaaaaagtga tggaaaatca aatgtttgaa gaaaagcaag ctgaaagtat gcctgaaatt 47580

gagactgagc aagagtataa agaaattttt ataacgcctc atcaaattaa acatattaag 47640

gattttaaag attataaata ttcacataga gttgataaaa agccgaatag agagttaata 47700

aatgatacat tatattctac gagaaaagat gacaagggta atacattaat cgttaataac 47760

ttaaatggtt tatacgataa agataatgat aaattgaaaa aattaattaa taaatcacct 47820

gaaaaattat tgatgtatca tcatgatcca caaacatatc aaaaattaaa attgatcatg 47880

gaacaatatg gcgatgagaa aaatccgctt tataaatatt atgaagaaac aggcaattac 47940

ttaacaaaat atagtaaaaa agataacgga ccagtcatca aaaaaattaa atattatggt 48000

aacaagctaa atgcgcattt agatattacg gatgattatc caaatagcag aaataaagta 48060

gtaaaacttt cattaaaacc atatcgcttt gatgtttatt tagataatgg ggtatataaa 48120

tttgtgacag ttaaaaattt agatgttatc aaaaaagaaa actactatga agttaattca 48180

aagtgttatg aagaagcaaa aaaactgaag aaaattagta atcaagcaga atttatcgca 48240

agtttttaca ataatgactt gattaagatt aacggagaat tatatagagt cataggtgta 48300

aataatgatc tacttaacag aattgaagta aatatgatag acatcacata tagagaatat 48360

ttagagaaca tgaatgataa aagaccacct agaataatta aaacaatagc aagcaaaaca 48420

caatctatta aaaagtattc tacagatatt ctaggcaatc tttatgaagt gaagagtaaa 48480

aagcatcctc aaatcataaa gaaaggatga ttggatattg gcatggagaa ttgtatatct 48540

ttcagaagta gataagcttg cattaaattt aaataactta aaagtcgcaa aaggggattt 48600

ggaggtaaaa atccccctta gcgatatttt cgcaattgtg attgaagatt taacaactac 48660

tataacctct aggcttatga tagagctatc taaatacaat attttagtga ttttttgtaa 48720

tcaaaagcat ttaccagagt gcatgattca accgattaat ggtcatttta atcaatatgg 48780

ccaaatgaag gaacaactac aatggtcgga agaaagaaaa gctgaattat ggaaacatat 48840

tctgtttcaa aaaatacaaa atcaaattca gtgtatgaaa cattttggag tggatgaaga 48900

aagaataatt aaaatgaaaa aattacaatc aatggttaca tctaaagatg aaatgaatat 48960

agagggccaa gcggctaagt attattttaa tacatttttc aaggacttca gacgcgatga 49020

tgatttttta gtagaaaata tggtactgga ttttgggtat actattttaa attcagcaat 49080

cgcaaggacc attgtagcta aagggttaat tccagcattg ggaatacatc atattggcgc 49140

aagaaaccac tttaatttag catcagatat cattgaagta tttaggccat tggtagattt 49200

atatcttttg aaacatcctc cgaaagtgga atatatgact agagagtata gattagagct 49260

gattaacttg ctgcatgcta gagtacaaat tgatggaaag cagcatacag ttattagagc 49320

gattgaaatt atggttcaat caataattga gtattttaga acggggagat ataaaacgct 49380

caaattacca aatattaagc aatattcatt ttataaatca tgagttataa gtttatgaga 49440

gtaatgatta tgtttgatct cccaactgac acaaaagaag aacgacggtt atatagtaag 49500

tttagaaagg gactcatttc ggaaggcttt ttgatgatgc agtattctat ttacattaaa 49560

tgttgtatta ataaagatgc ggcaaatagt agtgtaaatt atgtgaaacg tattttacca 49620

ccaaacggac atgttagagc tttaacaatt accgaaaagc aatataaaaa aatgaaaata 49680

ttattaggag atgaagataa aaatatcgaa atattaggag ataatcgtac aattttattt 49740

taggagaaaa cctatgtatg aatcaatttt agaaatcaat attgatatgg tacccaaagt 49800

tattttaaat actagtgtat atacatctat atataatgaa gattcaacac atgaagaata 49860

tattttagat atgcttaaga attattttca atctagaaat tcaaataagg atattgttac 49920

gattacagat atttcagatc atgaaagagt atctcataat cggtttagag gatttctctt 49980

agatcattct actattgaaa atgaacacca attatcatca tcgagtattt tgactaagaa 50040

gattgttaga gaaatgcaaa gtgatttgga aatagatggt tatattaatt ctatcaatgt 50100

actgttggat gacttacgag ataatttaaa gggcacactg ccactttcta ttaagtcttt 50160

cgatttaaaa caattcataa agctactctc atttaaatat gattattcgg aagattatgc 50220

cagactaatt actcggttag aacaagtcgt cccattagta gtggatgaaa tgaattatca 50280

agtaaaagaa aaagcatttg taatctattt atatccagaa gctggcctaa gcattaaaga 50340

acaaataaga tttagaaaaa tactaaataa cttacctgtt ccggtcattg ttctaaccga 50400

atcaccacgt tttttatctg aaagtattaa aggactaaat tattttatta ataatgttca 50460

aatgatcaca gaatcttttt ttgaagatag ctattggaat gcaccaatta atgtagaaaa 50520

agaacagatt gaagagaggt ttgaatattt atttgtaaaa tattgttcag ttttagaaat 50580

aaaaccgatt atttctaact atttacttgc ggatatcatt gtttttaatg atattgatct 50640

ttatatcctt gtttcatttt tgaaacactg taatttagag tataaattag atatcgataa 50700

aacaaaaatt tcattagcat tatcgtcata tctattttct taattatttt gaggatttat 50760

aatgatgcta gaggttttag tactctgtaa ttttaggtat gaggtagacc aggtttgtac 50820

ccactgacac caacaagcgg ttttagtact ctgtaatttt aggtatgagg tagaccgtaa 50880

agactttaac gaatttatgg ataaagtttt agtactctgt aattttaggt atgaggtaga 50940

caggaacaac ctgatttaaa tatgtggttg ggttttagta ctctgtaatt ttaggtatga 51000

ggtagactgt agcttcttca agttctttaa atagttcgtt ttagtactct gtaattttag 51060

gtatgaggta gacaaagtgg cgatgttatt aatatcaaag gttgttttag tactctgtaa 51120

ttttaggtat gaggtagact aatgtttgcg gtagaaggtt tacctgtacg ttttagtact 51180

ctgtaatttt aggtatgagg tagacaagtt ggtgaaattt tatataacgc aaacagtttt 51240

agtactctgt aattttaggt atgaggtaga ctaacgcatg tgtaatggtt ctaccctgcc 51300

agttttagta ctctgtaatt ttaggtatga ggtagaccca ctgcaacaag tcgtcgtgag 51360

cgtgtgcgtt ttagtactct gtaattttag gtatgaggta gacaaatata acgtggttta 51420

ttacgtcaga cgggttttag tactctgtaa ttttaggtat gaggtagaca aataccaaat 51480

tttgaattga acccaaatcg ttttagtact ctgtaatttt aggtatgagg tatatctttt 51540

ttaggtcggc tttttatttt gaatagtttt agtactctgt aattttagat atgaggtata 51600

tagcaaaatc actgataggg agaggcgtat cataagtgat gagtttgttt attataagat 51660

atacgaaaaa tattaattta aagatatcag agaaacatag agtatttata taactgttta 51720

gaagtataaa ataatcattt tttgtatcat acagttggtt ctataatgaa tcactaagaa 51780

tgattgtatt caataaatta ataaaaagca tatacatggt attagcaaat agaaattaac 51840

taggttttat tggtttatag cctttaaata aatatattta atttaaattc gcgattaggt 51900

agtagtaatt aatgtttgag attgaaaata gatttaaata tagcaattat atattaacct 51960

attaaagtgt taggctataa tgattagaat attctaggag atataaataa aatgaggaaa 52020

ttatacaatt taaaagaatt tttaaatatt tattcaactg aagatatagg tttaattaat 52080

acatatatcg aatctgataa tagaccactt aaagaaaatg aagtagagga tttaaaatca 52140

ttaataaatc tcttgaagct aaattataaa agttgtgaag gtttttttct tagttatgaa 52200

gttcctcata tcggtcacga aatagactta ttgaaagtta cggataatgc tattttaaat 52260

atcgaattaa aaagtagatc tgagcttgca aagataaata aacaacaaaa gtataattat 52320

ttttatctta atacacttaa caaaaatgta gacattataa cctataatag acatgaagaa 52380

aagttttata aatattgtca agaaagtgag aaatcacttg aaataaatat agataaagta 52440

agggataaaa tttttgaata tggaaataaa gagtttattt cagatattga ttatttattt 52500

aaaccttcta atttccttat ttcacctttt aatgatactg acaaatttct tgaagacaaa 52560

tattatttaa cggcccagca agaaaaaata aaaaaagaca ttataaatca taatgggact 52620

attgctatta ttgaaggaag accaggcaca ggtaaatcac ttttattata tgatattgct 52680

aaaaaactaa aaaaagagtt taatatatta ataatacatt gtgcattgtt aaatgaagga 52740

cacgaattaa tgattaataa aaattggaat ataaaaagtg ctcataaaaa ttgggtatac 52800

gatgaagaga attatgaagg taaaaattat atttttattg atgaatcaca taggttttac 52860

cctaaacaat ttgagtatgt aatagaaact gcacttaaaa aagatattaa gataataaca 52920

actatagacc cattacagta tctaagagat aatgaaaatg attttgaaaa cttacagaga 52980

ttaaaatctc attctaacat ctattataaa aaaattgata ataaaattag atctaatcca 53040

gaaatatcat catttgttaa aagtattatt aatttaagaa atatacctga atataaattt 53100

aagaatattt caatagaata tttgtcaaaa gaagatgatt ggaaatcata tttgctaaat 53160

ttacgtgata gaggatggac atatttacct tatacgaaaa gtagagataa tattagttat 53220

cataaatatt gttcggctaa tactcattta aatactcata gaattattgg gcaagaattt 53280

aataaagtta ttgttatttt agatgaagga tttaaatatg ttgatggata tttaggatat 53340

tacggagatt attattataa tcctagacaa atgttgttcc aaaatttaac tagagctaga 53400

gaaaaaataa aaattataat cattaataat acagatttat ataaaaagat taatcaagtg 53460

ttaactaagt cttctaataa aaatgagtta tcacttaaat tttttattcc tgaacttaat 53520

aatgataaat taaaatcatt ttatagtgat attttaggtt ttaagttaat tagtaatcaa 53580

aaacaaccgc ctggaaaaag aaagataagt tttcaaatag gtgaaacaaa tttagggttt 53640

tatgaaaaag gagatatttc taaaaacaat aaatgcgaaa taaatatagt tgtaagtaat 53700

ttgaaagatt taaagaaaag aatatcaaat gaaaatatgg tattgattag tgaagaagaa 53760

ggtaaaaaaa gtaaatttgt tactttaaat gatccattaa acaactattt aaagttaata 53820

gagtatatat aactagtatt ttgaattatt taaagattta ttgaagtttt gatattaatt 53880

tgtcgaaaaa cgaaccaaat tattaaatgc tacacaaaag aatatcaagt ataagttatt 53940

aatcataaat cagattgggt actcacctat tggaagtgaa gatgccaaac tattattcca 54000

attaatagat ttaaggtatg agaagaaaag tacaatagtt accacaaata tcaaattcaa 54060

cttatgtaac gaaatatttg atgattccaa aatcgcaaat gcgattttgg ataggatatt 54120

gcatcattca agcgtagtta aaatcacagg taaatcatat agattaaaag atcattcacc 54180

taagaaggaa attacttagc aaacaatttt gtacattctt atacaatcat ttatgtacat 54240

ttttatattg acatttacaa gtaagcaagt gaacttatct gctccagctg taagagaaag 54300

ggtaaataaa cttgaagagc aaggtattat acaaggttat acaattaaca tagattataa 54360

acaattaggt tacgatgtag aaaaattaat tgagttaact ataaagaata atagatacag 54420

ccatatgggc agacaaaatg tcagtttatc ttttcatcag ttgtctaaac tacatggaaa 54480

aagtttaata agcataggaa agatttttaa cagagagaag tgacttattt tggaatataa 54540

atcttacaac aaaatgagta ataatatttt atctatcttg ttaagtgaaa aaaatttcga 54600

taataaaatg gtagtctcaa caggtagtta tattttgtct gagttgccat gtattgttgc 54660

ctataaaaat aatgatattg tggggctact aacgtataaa gtttatgatg agtacattga 54720

aatcatttct ttagatagtt ttgtagagaa caagggtatt ggtagtcatt tacttaatta 54780

tgctgaaata attgcctcag atatgagtaa aagaagtatc cgtgttatta ctacaaatga 54840

aaatattaag gcactctatt tctatcaaaa aaataaatat agaatcacag acgttatttt 54900

tgatgctgta acagaagcta ggaaaataaa gccgtcgatt cctacaattg ataaaaatgg 54960

cattgaaatt agagatgaaa ttgttttgag gaaatgtttg taatgtataa atacactcag 55020

ctttaaatca ccctagttca tacctaaaag actttaacac aaacaaaaaa ggaggaacat 55080

taaattcctc ctcaaccttt attactcata ctacaattct attttaacgt cttcgtccat 55140

ttgggcttca aattcatcta gaagtgctct tgcttctgca attgattgtg tgttcatcaa 55200

ttgatggcga agttcgctag cgcctcttat gccacgcaca tagattttaa agaatctacg 55260

caagctcttg aattgtcgta tttcatcttt ctcatatttg ttaaacaatg atagatgcaa 55320

tctcaacaga tctaatagtt ctttgcttgt gtgttcacgt ggttcttttt caaaagcgaa 55380

tggattgtgg aaaatgcctc taccaatcat gacgccatca atgccatatt tttctgccag 55440

ttcaagtcct gtttttctat cgggaatatc accgttaatt gttaacaatg tatttggtgc 55500

aatttcgtca cgtaaatttt taatagcttc gattaattcc caatgtgcat ctactttact 55560

catttcttta cgtgtacgaa gatgaataga taaattggca atgtcttgtt cgaagacgtg 55620

cttcaaccaa tctttccatt catcgatttc atagtagcca aggcgtgttt taacacttac 55680

cggaagccca cctgctttag tcgcttgaat aatttcggca gcaacgtcag gtcttaagat 55740

taagccggaa cccttaccct ttttagcaac atttgctaca ggacatccca tatttaagtc 55800

tatgccttta aagcccattt tagctaattg aatactcgtt tcacggaact gttctggctt 55860

atctccccat atatgagcga ccatcggctg ttcatcttca ctaaaagtta agcgtccgcg 55920

cacactatgt atgccttcag ggtggcaaaa gctttcagta tttgtaaatt cagtgaaaaa 55980

cacatccggt ctagctgctt cacttacaac gtgtcgaaag acgatatctg taacgtcttc 56040

cattggcgcc aaaataaaaa atggacgtgg taattcactc caaaaatttt ctttcataat 56100

atatttatac cctct 56115

<210> 50
<211> 23
<212> DNA
<213> Bacterial

<400> 50
cctggcctgc cccaagtgca agg 23

<210> 51
<211> 32
<212> DNA
<213> Bacterial

<400> 51
gtcgcccccc acgcgggggc gtggattgaa ac 32

<210> 52
<211> 28
<212> DNA
<213> Bacterial

<400> 52
tggagcggga aacgggattt gaacccgc 28

<210> 53
<211> 29
<212> DNA
<213> Bacterial

<400> 53
gtgttcccca cgggcgtggg gatgaaccg 29

<210> 54
<211> 25
<212> DNA
<213> Bacterial

<400> 54
aacctttctg caaaaaggtt tcccc 25

<210> 55
<211> 23
<212> DNA
<213> Bacterial

<400> 55
ccgctggatc cggctgcagc gcc 23

<210> 56
<211> 28
<212> DNA
<213> Bacterial

<400> 56
gttcactgcc gcataggcag ctcagaaa 28

<210> 57
<211> 32
<212> DNA
<213> Bacterial

<400> 57
cacccgacta ttgaagtcgg gcctcattga ag 32

<210> 58
<211> 36
<212> DNA
<213> Bacterial

<400> 58
gtcaaaaccc ataccgatat ggatacctct tttgag 36

<210> 59
<211> 36
<212> DNA
<213> Bacterial

<400> 59
gtcaaaaccc ataccgatat ggatacctct tttgag 36

<210> 60
<211> 36

<212> DNA
<213> Bacterial

<400> 60
gtcaaaaccc ataccgatat ggatacctct tttgag 36

<210> 61
<211> 36
<212> DNA
<213> Bacterial

<400> 61
gtcaaaaccc ataccgatat ggatacctct tttgag 36

<210> 62
<211> 36
<212> DNA
<213> Bacterial

<400> 62
ctcaaaagag gtatccatat cggtatgggt tttgac 36

<210> 63
<211> 28
<212> DNA
<213> Bacterial

<400> 63
gttcactgcc ggataggcag cttagaaa 28

<210> 64
<211> 32
<212> DNA
<213> Bacterial

<400> 64
gtcgcccccc gcgcgggggc gtggattgaa ac 32

<210> 65
<211> 29
<212> DNA
<213> Bacterial

<400> 65
ttttctgagc tgcctatgcg gcagtgaac 29

<210> 66
<211> 46
<212> DNA
<213> Bacterial

<400> 66
catcgacgac gaacccgggc accgcctgat ggtccacgcc gtcatg 46

<210> 67
<211> 32
<212> DNA
<213> Bacterial

<400> 67
gtcgcccctc acgcgggggc gtggatagaa ac 32

<210> 68
<211> 32
<212> DNA
<213> Bacterial

<400> 68
gtttcaatcc acgcccccgc acgggggggcg ac 32

<210> 69
<211> 28
<212> DNA
<213> Bacterial

<400> 69
tttctgagct gcctatgcgg cagtgaac 28

<210> 70
<211> 36
<212> DNA
<213> Bacterial

<400> 70
gaccgaagac ctgtcggaaa cgacggggat tgagac 36

<210> 71
<211> 29
<212> DNA
<213> Bacterial

<400> 71
cggttcatcc ccgcgagtgc ggggaacat 29

<210> 72
<211> 28
<212> DNA
<213> Bacterial

<400> 72
tttctgagct gcctgtgcgg cagtgaac 28

<210> 73
<211> 29
<212> DNA
<213> Bacterial

<400> 73
gttttatctg aacgtagtgg gatataaag 29

<210> 74
<211> 29
<212> DNA
<213> Bacterial

<400> 74
cggttcatcc ccgcgggtgc ggggaacac 29

<210> 75
<211> 28
<212> RNA
<213> Bacterial

<400> 75
guuagcagcc gcauaggcug cuuaaaua 28

<210> 76
<211> 33
<212> DNA
<213> Bacterial

<400> 76
aatttaaata gggaagataa gcaaagggtt gac 33

<210> 77
<211> 31
<212> DNA
<213> Bacterial

<400> 77
tttaaatagg gaagataagc aaagggttga c 31

<210> 78
<211> 11
<212> PRT
<213> Bacterial

<400> 78

Asn	Leu	Asn	Arg	Glu	Asp	Lys	Gln	Arg	Val	Asp
1				5					10	

<210> 79
<211> 32
<212> DNA
<213> Bacteriophage

<400> 79
cattgcaact atgcaaaatg atgaagctaa aa 32

<210> 80
<211> 32
<212> DNA
<213> Bacteriophage

<400> 80
tgttagagtc ggtagtatct ggatgatcga ta 32

<210> 81
<211> 32
<212> DNA
<213> Bacteriophage

<400> 81
ttatgtattg accccgacac gccccccgac tg 32

<210> 82
<211> 32
<212> DNA
<213> Bacteriophage

<400> 82
ttacagacga cctaactctt cagtaccatg at 32

<210> 83
<211> 32
<212> DNA
<213> Bacteriophage

<400> 83
tacataagct gcaacacggt gttcgtttaa gt 32

<210> 84
<211> 32
<212> DNA
<213> Bacteriophage

<400> 84
aaaatacgcc tttttccctt catcgtttaa ag 32

<210> 85
<211> 32
<212> DNA
<213> Bacteriophage

<400> 85
accaacaaat cccataaact gataaccacg tt 32

<210> 86
<211> 32
<212> DNA
<213> Bacteriophage

<400> 86
gtcaaccctt tgcttatctt ccctatttaa at 32

<210> 87
<211> 32
<212> DNA
<213> Bacteriophage

<400> 87
tgttaaccac cgcttgaaat aatcatgatg ca 32

<210> 88
<211> 32
<212> DNA
<213> Bacteriophage

<400> 88
tgtgtctata ctcaaccaat ttaagcgccg ca 32

<210> 89
<211> 32

<212> DNA
<213> Bacteriophage

<400> 89
ctactctccc caatattagc cattcctaat tc 32

<210> 90
<211> 32
<212> DNA
<213> Bacteriophage

<400> 90
gtcaccttac cgtaagacag gcagtaaaat ta 32

<210> 91
<211> 32
<212> DNA
<213> Bacteriophage

<400> 91
aaactagtgg acgtaatgca gtattcacgg tt 32

<210> 92
<211> 32
<212> DNA
<213> Bacteriophage

<400> 92
atccacacta caaatagaac actcaaccgt ga 32

<210> 93
<211> 32
<212> DNA
<213> Bacteriophage

<400> 93
tgtgtctata ctcaaccaat ttaagcgccg ca 32

<210> 94
<211> 32
<212> DNA
<213> Bacteriophage

<400> 94
ctactctccc caatattagc cattcctaat tc 32

<210> 95
<211> 32
<212> DNA
<213> Bacteriophage

<400> 95
aaactagtgg acgtaatgca gtattcacgg tt 32

<210> 96
<211> 32
<212> DNA
<213> Bacteriophage

<400> 96
ataatcgttt tgagtctcac cagcttttag gc 32

<210> 97
<211> 32
<212> DNA
<213> Bacteriophage

<400> 97
atccacacta caaatagaac actcaaccgt ga 32

<210> 98
<211> 32
<212> DNA
<213> Bacteriophage

<400> 98
tattgattgg tctctaacct tgggatgatt aa 32

<210> 99
<211> 32
<212> DNA
<213> Bacteriophage

<400> 99
ttcacgggta gcaacagggt aataaaccaa ta 32

<210> 100
<211> 32
<212> DNA
<213> Bacteriophage

<400> 100
cattgcaact atgcaaaatg atgaagctaa aa 32

<210> 101
<211> 32
<212> DNA
<213> Bacteriophage

<400> 101
tgttagagtc ggtagtatct ggatgatcga ta 32

<210> 102
<211> 32
<212> DNA
<213> Bacteriophage

<400> 102
tagaagagta ataggagcta ctgcaaactt gt 32

<210> 103
<211> 32
<212> DNA
<213> Bacteriophage

<400> 103
taactatgtg tggtttatat tttgtgtgca ag 32

<210> 104
<211> 32
<212> DNA
<213> Bacteriophage

<400> 104
ttttgaaact attgacagaa ggttgggaac ct 32

<210> 105
<211> 32
<212> DNA
<213> Bacteriophage

<400> 105
ttgaggttga acctcttccg gttcctcttc tg 32

<210> 106
<211> 32
<212> DNA
<213> Bacteriophage

<400> 106
gtgtattgct tgcagtgggt tacacacaag aa 32

<210> 107
<211> 47
<212> DNA
<213> Bacterial

<400> 107
gttgtgattt gctttcaaat tagtatcttt gaaccattgg aaacagc 47

<210> 108
<211> 29
<212> DNA
<213> Bacterial

<400> 108
atttcaattc cataaggtac aattaatac 29

<210> 109
<211> 32
<212> DNA
<213> Bacterial

<400> 109
gtttcaatcc acacacccgt atagggtgtg ac 32

<210> 110
<211> 47
<212> DNA
<213> Bacterial

<400> 110
actgtttctg atatgtcaaa gataaaattt tgaaagcaaa tcacaac 47

<210> 111
<211> 23

<212> DNA
<213> Bacterial

<400> 111
gaaaaaatac agtttcgctc tca 23

<210> 112
<211> 32
<212> DNA
<213> Bacterial

<400> 112
gtcgcgtctc acgtaggcgc gtggattgaa ac 32

<210> 113
<211> 47
<212> DNA
<213> Bacterial

<400> 113
attgtgcttg ctactgcaaa gatacacatt ttgaagcaat tcacaac 47

<210> 114
<211> 37
<212> DNA
<213> Bacterial

<400> 114
ctcaatgagt atcttccatt aaaacaagga ttaagac 37

<210> 115
<211> 36
<212> DNA
<213> Bacterial

<400> 115
gttgttttta ccttgcaaac agcaggcaga tacaac 36

<210> 116
<211> 47
<212> DNA
<213> Bacterial

<400> 116
gttgtatttg ccaatgcaaa gatactaatt ttaaagctaa tcacaac 47

<210> 117
<211> 36
<212> DNA
<213> Bacterial

<400> 117
gttgtatatc attcctttcc tacatcaaac cacaac 36

<210> 118
<211> 487
<212> DNA
<213> Synthetiic

<400> 118

gcggataaca attactagag aaagaggaga aatactattc ttctcctctt taaataacga 60

aaacaccctg ccataaaatg acagggtgtt gatttcggca tgaagcctta tctttgtagc 120

ttctgcaaga tttaagtaac tgtgtaaggc gtcccttaca cttgcatgta tagttattat 180

accaggggga cagtgcaatg tcaagaataa actgtagaat gactagtgac ttaaatcttg 240

agagtacaaa aacccgttgg aatcgtgatt aatagtaact gttgttgtac agttacttaa 300

atcttgagag tacaaaaacg gccgagaaaa ggagctgatt cataggacag ttgtacagtt 360

acttaaatct tgagagtaca aaaactcaaa cttgcccgta gtttatctta tagccgttgt 420

acagttactt aaatcttgag agtacaaaaa catttacctc ctttgattta agtgaacaag 480

tttatcc 487

<210> 119
<211> 221
<212> DNA
<213> Streptococcus thermophilus

<400> 119

ttaaatcttg agagtacaaa aacccgttgg aatcgtgatt aatagtaact gttgttgtac 60

agttacttaa atcttgagag tacaaaaacg gccgagaaaa ggagctgatt cataggacag 120

ttgtacagtt acttaaatct tgagagtaca aaaactcaaa cttgcccgta gtttatctta 180

tagccgttgt acagttactt aaatcttgag agtacaaaaa c 221

<210> 120
<211> 112
<212> DNA
<213> Streptococcus thermophilus

<400> 120

ttaaataacg aaaacaccct gccataaaat gacagggtgt tgatttcggc atgaagcctt 60

atctttgtag cttctgcaag atttaagtaa ctgtgtaagg cgtcccttac ac 112

<210> 121
<211> 30
<212> DNA
<213> Streptococcus thermophilus

<400> 121
tgtcctatga atcagctcct tttctcggcc 30

<210> 122
<211> 30
<212> DNA
<213> Streptococcus thermophilus

<400> 122
ggctataaga taaactacgg gcaagtttga 30

Claims

1. A nucleic acid vector comprising an engineered host modifying (HM) CRISPR array for use in treating or preventing, in a human or animal subject, a condition or disease mediated by bacterial host cells by altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria, the second bacteria comprising host cells, wherein the mixed population is comprised by the subject,

   wherein the second sub-population consists of a different bacterial species to that of the first sub-population,
   wherein each host cell has a host target sequence and at least one nucleic acid sequence encoding a Cas nuclease endogenous to the host cell,
   wherein the engineered host modifying (HM) CRISPR array comprises a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to the host target sequence;
   wherein the vector can transform the host cell,
   wherein the HM-crRNA guides the Cas nuclease to the target to modify the target sequence in the host cell, whereby the host cell is killed or host cell growth is reduced,
   wherein the nucleic acid vector is a plasmid vector comprised by a carrier bacterium, or the vector is a phage vector.

2. A nucleic acid vector comprising an engineered host modifying (HM) CRISPR array and a nucleic acid sequence encoding a Cas nuclease for use in treating or preventing in a human or animal subject a condition or disease mediated by bacterial host cells by altering the relative ratio of sub-populations of first and second bacteria in a mixed population of bacteria, the second bacteria comprising host cells, wherein the mixed population is comprised by the subject,

   wherein the second sub-population consists of a different bacterial species to that of the first sub-population,
   wherein each host cell has a host target sequence,
   wherein the vector can transform the host cell,
   wherein the engineered host modifying (HM) CRISPR array comprises a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to the host target sequence,
   wherein the HM-crRNA guides the Cas nuclease to the target to modify the target sequence in the host cell, whereby the host cell is killed or host cell growth is reduced,
   wherein the nucleic acid vector is a plasmid vector comprised by a carrier bacterium, or the vector is a phage vector.

3. Use of a nucleic acid vector comprising an engineered host modifying (HM) CRISPR/Cas array for altering the relative ratio of sub-populations of first and second bacteria in an *ex vivo* mixed population of bacteria, the second bacteria comprising host cells,

   wherein the second sub-population consists of a different bacterial species to that of the first sub-population,
   wherein each host cell has a host target sequence and at least one nucleic acid sequence encoding a Cas nuclease endogenous to the host cell,
   wherein the engineered host modifying (HM) CRISPR array comprises a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to the host target sequence;
   wherein the vector can transform the host cell,
   wherein the HM-crRNA guides the Cas nuclease to the target to modify the target sequence in the host cell, whereby the host cell is killed or host cell growth is reduced,
   wherein the nucleic acid vector is a plasmid vector comprised by a carrier bacterium, or the vector is a phage vector.

4. Use of a nucleic acid vector comprising an engineered host modifying (HM) CRISPR/Cas array and a nucleic acid sequence encoding a Cas nuclease for altering the relative ratio of sub-populations of first and second bacteria in

an *ex vivo* mixed population of bacteria, the second bacteria comprising host cells,

wherein the second sub-population consists of a different bacterial species to that of the first sub-population,
wherein each host cell has a host target sequence,
wherein the vector can transform the host cell,
wherein the engineered host modifying (HM) CRISPR array comprises a spacer sequence (HM-spacer) and repeats encoding a HM-crRNA, the HM-crRNA comprising a sequence that is capable of hybridising to the host target sequence,
wherein the HM-crRNA guides the Cas nuclease to the target to modify the target sequence in the host cell, whereby the host cell is killed or host cell growth is reduced,
wherein the nucleic acid vector is a plasmid vector comprised by a carrier bacterium, or the vector is a phage vector.

5. The use of claims 3 or 4 for treating an industrial or an *ex vivo* medical fluid, surface, apparatus or container; or for treating a waterway, water, a beverage, a foodstuff or a cosmetic, wherein the host cell(s) are comprised by or on the fluid, surface, apparatus, container, waterway, water, beverage, foodstuff or cosmetic.

6. The use or vector for use according to any preceding claim, wherein alternatively HM-crRNA and tracrRNA are comprised by a single guide RNA (gRNA).

7. The use or vector for use according to any preceding claim, wherein host cell population growth is reduced by at least 5-fold compared to the growth of a population of said host cells not transformed with said HM-array or a nucleotide sequence encoding said gRNA.

8. The use according to claim 3, claim 4 or any one of claims 5-7 when dependent on claim 3 or 4, wherein host cell population growth on a surface is inhibited.

9. The use or vector for use according to any one of claim 1 or 3, or any one of claims 5-8 when dependent on any one of claims 1 or 3 , wherein each vector lacks a Cas nuclease encoding sequence.

10. The use or vector for use according to any preceding claim, wherein the vector comprises a DNA sequence for expressing a tracrRNA sequence.

11. The use or vector for use according to any preceding claim wherein the mixed population comprises *S. thermophilus, L. lactis* and *E.coli.*

12. The use or vector for use according to any preceding claim wherein the Cas nuclease is a type I Cas nuclease, a type II Cas nuclease or a type III Cas nuclease.

13. A vector for use, or use according to any preceding claim, the vector comprising nucleic acid sequences for expressing a plurality of different crRNAs (eg, comprised by gRNAs) wherein a first of said crRNAs is capable of hybridising to a first nucleic acid sequence in said host cell; and a second of said crRNAs is capable of hybridising to a second nucleic acid sequence in said host cell, wherein said second sequence is different from said first sequence; and

(a) the first sequence is comprised by an antibiotic resistance gene (or RNA thereof) and the second sequence is comprised by an antibiotic resistance gene (or RNA thereof); optionally wherein the genes are different;
(b) the first sequence is comprised by an antibiotic resistance gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof);
(c) the first sequence is comprised by an essential gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof); or
(d) the first sequence is comprised by a virulence gene (or RNA thereof) and the second sequence is comprised by an essential or virulence gene (or RNA thereof).

14. A vector for use, or use according to any one of claims 1, 3, and 6-12 when dependent on claim 1 or 3, the vector comprising more than 1.4kb of exogenous DNA sequence, wherein the exogenous DNA encodes one or more components of a CRISPR/Cas system and comprises an engineered array for expressing HM-crRNAs or gRNAs in host cells; wherein at least 2 different cRNAs or gRNAs are encoded by the exogenous DNA (eg, by at least 2 HM-CRISPR arrays).

15. The vector for use, or use according to any one of claims 1, 3 and 6-14 when dependent on claim 1 or 3 wherein the vector is devoid of a nucleotide sequence encoding a Cas nuclease that is cognate to the cRNA(s) or gRNA(s).

Patentansprüche

1. Nukleinsäurevektor, umfassend eine manipulierte, wirtsmodifizierende (HM) CRISPR-Anordnung zur Verwendung in Behandlung oder Prävention, in einem menschlichen oder tierischen Subjekt, eines Zustands oder einer Erkrankung, der bzw. die durch bakterielle Wirtszellen vermittelt wird, indem das relative Verhältnis von Subpopulationen erster und zweiter Bakterien in einer gemischten Population von Bakterien verändert wird, wobei die zweiten Bakterien Wirtszellen umfassen, wobei die gemischte Population das Subjekt umfasst,

wobei die zweite Subpopulation aus einer anderen Bakterienart als jener der ersten Subpopulation besteht,
wobei jede Wirtszelle eine Wirtszielsequenz und mindestens eine Nukleinsäuresequenz aufweist, die für eine Cas-Nuklease codiert, die für die Wirtszelle endogen ist,
wobei die manipulierte, wirtsmodifizierende (HM) CRISPR-Anordnung eine Trennsequenz (HM-Spacer) und Wiederholungseinheiten umfasst, die für eine HM-crRNA codieren, wobei die HM-crRNA eine Sequenz umfasst, die imstande ist, an die Wirtszielsequenz zu hybridisieren;
wobei der Vektor die Wirtszelle transformieren kann,
wobei die HM-crRNA die Cas-Nuklease zu dem Ziel führt, um die Zielsequenz in der Wirtszelle zu modifizieren, wodurch die Wirtszelle getötet oder Wirtszellenwachstum verringert wird,
wobei der Nukleinsäurevektor ein Plasmidvektor ist, der ein Trägerbakterium umfasst, oder der Vektor ein Phage-Vektor ist.

2. Nukleinsäurevektor, umfassend eine manipulierte, wirtsmodifizierende (HM) CRISPR-Anordnung und eine Nukleinsäuresequenz, die für eine Cas-Nuklease codiert, zur Verwendung in Behandlung oder Prävention, in einem menschlichen oder tierischen Subjekt, eines Zustands oder einer Erkrankung, der bzw. die durch bakterielle Wirtszellen vermittelt wird, indem das relative Verhältnis von Subpopulationen erster und zweiter Bakterien in einer gemischten Population von Bakterien verändert wird, wobei die zweiten Bakterien Wirtszellen umfassen, wobei die gemischte Population das Subjekt umfasst,

wobei die zweite Subpopulation aus einer anderen Bakterienart als jener der ersten Subpopulation besteht,
wobei jede Wirtszelle eine Wirtszielsequenz aufweist,
wobei der Vektor die Wirtszelle transformieren kann,
wobei die manipulierte, wirtsmodifizierende (HM) CRISPR-Anordnung eine Trennsequenz (HM-Spacer) und Wiederholungseinheiten umfasst, die für eine HM-crRNA codieren, wobei die HM-crRNA eine Sequenz umfasst, die imstande ist, an die Wirtszielsequenz zu hybridisieren,
wobei die HM-crRNA die Cas-Nuklease zu dem Ziel führt, um die Zielsequenz in der Wirtszelle zu modifizieren, wodurch die Wirtszelle getötet oder Wirtszellenwachstum verringert wird,
wobei der Nukleinsäurevektor ein Plasmidvektor ist, der ein Trägerbakterium umfasst, oder der Vektor ein Phage-Vektor ist.

3. Verwendung eines Nukleinsäurevektors, umfassend eine manipulierte, wirtsmodifizierende (HM) CRISPR/Cas-Anordnung zum Ändern des relativen Verhältnisses von Subpopulationen erster und zweiter Bakterien in einer *ex vivo* gemischten Population von Bakterien, wobei die zweiten Bakterien Wirtszellen umfassen,

wobei die zweite Subpopulation aus einer anderen Bakterienart als jener der ersten Subpopulation besteht,
wobei jede Wirtszelle eine Wirtszielsequenz und mindestens eine Nukleinsäuresequenz aufweist, die für eine Cas-Nuklease codiert, die für die Wirtszelle endogen ist,
wobei die manipulierte, wirtsmodifizierende (HM) CRISPR-Anordnung eine Trennsequenz (HM-Spacer) und Wiederholungseinheiten umfasst, die für eine HM-crRNA codieren, wobei die HM-crRNA eine Sequenz umfasst, die imstande ist, an die Wirtszielsequenz zu hybridisieren;
wobei der Vektor die Wirtszelle transformieren kann,
wobei die HM-crRNA die Cas-Nuklease zu dem Ziel führt, um die Zielsequenz in der Wirtszelle zu modifizieren, wodurch die Wirtszelle getötet oder Wirtszellenwachstum verringert wird,
wobei der Nukleinsäurevektor ein Plasmidvektor ist, der ein Trägerbakterium umfasst, oder der Vektor ein Phage-Vektor ist.

4. Verwendung eines Nukleinsäurevektors, umfassend eine manipulierte, wirtsmodifizierende (HM) CRISPR/Cas-Anordnung und eine Nukleinsäuresequenz, die für eine Cas-Nuklease codiert, zum Ändern des relativen Verhältnisses von Subpopulationen erster und zweiter Bakterien in einer *ex vivo* gemischten Population von Bakterien, wobei die zweiten Bakterien Wirtszellen umfassen,

wobei die zweite Subpopulation aus einer anderen Bakterienart als jener der ersten Subpopulation besteht,
wobei jede Wirtszelle eine Wirtszielsequenz aufweist,
wobei der Vektor die Wirtszelle transformieren kann,
wobei die manipulierte, wirtsmodifizierende (HM) CRISPR-Anordnung eine Trennsequenz (HM-Spacer) und Wiederholungseinheiten umfasst, die für eine HM-crRNA codieren, wobei die HM-crRNA eine Sequenz umfasst, die imstande ist, an die Wirtszielsequenz zu hybridisieren,
wobei die HM-crRNA die Cas-Nuklease zu dem Ziel führt, um die Zielsequenz in der Wirtszelle zu modifizieren, wodurch die Wirtszelle getötet oder Wirtszellenwachstum verringert wird,
wobei der Nukleinsäurevektor ein Plasmidvektor ist, der ein Trägerbakterium umfasst, oder der Vektor ein Phage-Vektor ist.

5. Verwendung nach Anspruch 3 oder 4 zur Behandlung eines industriellen oder eines *ex vivo* medizinischen Fluids, einer Oberfläche, einer Einrichtung oder eines Behälters; oder zum Behandeln einer Wasserstraße, von Wasser, eines Getränks, eines Nahrungsmittels oder eines Kosmetikprodukts, wobei die Wirtszelle(n) in oder auf dem Fluid, der Oberfläche, der Einrichtung, dem Behälter, der Wasserstraße, dem Wasser, dem Getränk, dem Nahrungsmittel oder Kosmetikprodukt umfasst ist (sind).

6. Verwendung oder Vektor zur Verwendung nach einem vorstehenden Anspruch, wobei alternativ HM-cRNA und tracrRNA eine einzelne Führungs-RNA (gRNA) umfassen.

7. Verwendung oder Vektor zur Verwendung nach einem vorstehenden Anspruch, wobei Wirtszellenpopulationswachstum mindestens um ein Fünffaches im Vergleich zu dem Wachstum einer Population der Wirtszellen, die nicht mit der HM-Anordnung oder einer Nukleotidsequenz, die für die gRNA codiert, transformiert sind, verringert ist.

8. Verwendung nach Anspruch 3, Anspruch 4 oder einem der Ansprüche 5-7, wenn von Anspruch 3 oder 4 abhängig, wobei Wirtszellenpopulationswachstum auf einer Oberfläche gehemmt ist.

9. Verwendung oder Vektor zur Verwendung nach einem von Anspruch 1 oder 3, oder nach einem von Ansprüchen 5-8, wenn von einem der Ansprüche 1 oder 3 abhängig, wobei jedem Vektor eine Cas-Nuklease codierende Sequenz fehlt.

10. Verwendung oder Vektor zur Verwendung nach einem vorstehenden Anspruch, wobei der Vektor eine DNA-Sequenz zum Exprimieren einer tracrRNA-Sequenz umfasst.

11. Verwendung oder Vektor zur Verwendung nach einem vorstehenden Anspruch, wobei die gemischte Populationi *S.- thermophilus, L. lactis* und *E. coli* umfasst.

12. Verwendung oder Vektor zur Verwendung nach einem vorstehenden Anspruch, wobei die Cas-Nuklease eine Typ I Cas-Nuklease, eine Typ II Cas-Nuklease oder eine Typ III Cas-Nuklease ist.

13. Vektor zur Verwendung oder Verwendung nach einem vorstehenden Anspruch, wobei der Vektor Nukleinsäuresequenzen zum Exprimieren einer Vielzahl von unterschiedlichen crRNAs (z.B. gRNAs umfassend) umfasst, wobei eine erste der crRNAs imstande ist, an eine erste Nukleinsäuresequenz in der Wirtszelle zu hybridisieren; und eine zweite der crRNAs imstande ist, an eine zweite Nukleinsäuresequenz in der Wirtszelle zu hybridisieren, wobei sich die zweite Sequenz von der ersten Sequenz unterscheidet; und

a) die erste Sequenz ein Antibiotikum-Resistenzgen (oder dessen RNA) umfasst und die zweite Sequenz ein Antibiotikum-Resistenzgen (oder dessen RNA) umfasst; wobei die Gene optional unterschiedlich sind;
b) die erste Sequenz ein Antibiotikum-Resistenzgen (oder dessen RNA) umfasst und die zweite Sequenz ein essenzielles oder Virulenzgen (oder dessen RNA) umfasst;
c) die erste Sequenz ein essenzielles oder Virulenzgen (oder dessen RNA) umfasst und die zweite Sequenz ein essenzielles oder Virulenzgen (oder dessen RNA) umfasst; oder
d) die erste Sequenz ein Virulenzgen (oder dessen RNA) umfasst und die zweite Sequenz ein essenzielles

oder Virulenzgen (oder dessen RNA) umfasst.

14. Vektor zur Verwendung oder Verwendung nach einem der Ansprüche 1, 3 und 6-12, wenn von Anspruch 1 oder 3 abhängig, wobei der Vektor mehr als 1,4 kb exogene DNA-Sequenz umfasst, wobei die exogene DNA für eine oder mehrere Komponenten eines CRSPR/Cas-Systems codiert und eine manipulierte Anordnung zur Expression von HM-cRNAs oder gRNAs in Wirtszellen umfasst; wobei mindestens 2 verschiedene cRNAs oder gRNAs durch die exogene DNA codiert werden (z.B. durch mindestens 2 HM-CRISPR-Anordnungen).

15. Vektor zur Verwendung oder Verwendung nach einem der Ansprüche 1, 3 und 6-14, wenn von Anspruch 1 oder 3 abhängig, wobei es dem Vektor an einer Nukleotidsequenz fehlt, die für eine Cas-Nuklease codiert, die mit den cRNA(s) oder gRNA(s) artverwandt ist.

Revendications

1. Vecteur d'acide nucléique comprenant un réseau CRISPR modifié de modification d'hôte (HM) destiné à être utilisé dans le traitement ou la prévention, chez un sujet humain ou animal, d'une affection ou d'une maladie médiée par des cellules hôtes bactériennes en altérant le rapport relatif de sous-populations de première et seconde bactéries dans une population mixte de bactéries, la seconde bactérie comprenant des cellules hôtes, dans lequel la population mixte est constituée par le sujet,

dans lequel la seconde sous-population se compose d'une espèce bactérienne différente de celle de la première sous-population,
dans lequel chaque cellule hôte présente une séquence cible d'hôte et au moins une séquence d'acides nucléiques codant une nucléase Cas endogène à la cellule hôte,
dans lequel le réseau CRISPR modifié de modification d'hôte (HM) comprend une séquence d'espaceurs (espaceur HM) et répète le codage d'un ARNcr HM, l'ARNcr HM comprenant une séquence qui est capable de s'hybrider à la séquence cible d'hôte ;
dans lequel le vecteur peut transformer la cellule hôte,
dans lequel l'ARNcr HM guide la nucléase Cas vers la cible pour modifier la séquence cible dans la cellule hôte, moyennant quoi la cellule hôte est tuée ou une croissance de cellules hôtes est réduite,
dans lequel le vecteur d'acide nucléique est un vecteur plasmidique constitué par une bactérie porteuse ou le vecteur est un vecteur phagique.

2. Vecteur d'acide nucléique comprenant un réseau CRISPR modifié de modification d'hôte (HM) et une séquence d'acides nucléiques codant une nucléase Cas destiné à être utilisé dans le traitement ou la prévention, chez un sujet humain ou animal, d'une affection ou d'une maladie médiée par des cellules hôtes bactériennes en altérant le rapport relatif de sous-populations de première et seconde bactéries dans une population mixte de bactéries, la seconde bactérie comprenant des cellules hôtes, dans lequel la population mixte est constituée par le sujet,

dans lequel la seconde sous-population se compose d'une espèce bactérienne différente de celle de la première sous-population,
dans lequel chaque cellule hôte présente une séquence cible d'hôte,
dans lequel le vecteur peut transformer la cellule hôte,
dans lequel le réseau CRISPR modifié de modification d'hôte (HM) comprend une séquence d'espaceurs (espaceur HM) et répète le codage d'un ARNcr HM, l'ARNcr HM comprenant une séquence qui est capable de s'hybrider à la séquence cible d'hôte ;
dans lequel l'ARNcr HM guide la nucléase Cas vers la cible pour modifier la séquence cible dans la cellule hôte, moyennant quoi la cellule hôte est tuée ou une croissance de cellules hôtes est réduite,
dans lequel le vecteur d'acide nucléique est un vecteur plasmidique constitué par une bactérie porteuse ou le vecteur est un vecteur phagique.

3. Utilisation d'un vecteur d'acide nucléique comprenant un réseau CRISPR/Cas modifié de modification d'hôte (HM) pour modifier le rapport relatif de sous-populations de première et seconde bactéries dans une population mixte *ex vivo* de bactéries, la seconde bactérie comprenant des cellules hôtes,

dans lequel la seconde sous-population se compose d'une espèce bactérienne différente de celle de la première sous-population,

dans lequel chaque cellule hôte présente une séquence cible d'hôte et au moins une séquence d'acides nucléiques codant une nucléase Cas endogène à la cellule hôte,
dans lequel le réseau CRISPR modifié de modification d'hôte (HM) comprend une séquence d'espaceurs (espaceur HM) et répète le codage d'un ARNcr HM, l'ARNcr HM comprenant une séquence qui est capable de s'hybrider à la séquence cible d'hôte ;
dans lequel le vecteur peut transformer la cellule hôte,
dans lequel l'ARNcr HM guide la nucléase Cas vers la cible pour modifier la séquence cible dans la cellule hôte, moyennant quoi la cellule hôte est tuée ou une croissance de cellules hôtes est réduite,
dans lequel le vecteur d'acide nucléique est un vecteur plasmidique constitué par une bactérie porteuse ou le vecteur est un vecteur phagique.

4. Utilisation d'un vecteur d'acide nucléique comprenant un réseau CRISPR/Cas modifié de modification d'hôte (HM) et une séquence d'acides nucléiques codant une nucléase Cas pour modifier le rapport relatif de sous-populations de première et seconde bactéries dans une population mixte *ex vivo* de bactéries, la seconde bactérie comprenant des cellules hôtes,

dans lequel la seconde sous-population se compose d'une espèce bactérienne différente de celle de la première sous-population,
dans lequel chaque cellule hôte présente une séquence cible d'hôte,
dans lequel le vecteur peut transformer la cellule hôte,
dans lequel le réseau CRISPR modifié de modification d'hôte (HM) comprend une séquence d'espaceurs (espaceur HM) et répète le codage d'un ARNcr HM, l'ARNcr HM comprenant une séquence qui est capable de s'hybrider à la séquence cible d'hôte ;
dans lequel l'ARNcr HM guide la nucléase Cas vers la cible pour modifier la séquence cible dans la cellule hôte, moyennant quoi la cellule hôte est tuée ou une croissance de cellules hôtes est réduite,
dans lequel le vecteur d'acide nucléique est un vecteur plasmidique constitué par une bactérie porteuse ou le vecteur est un vecteur phagique.

5. Utilisation selon les revendications 3 ou 4 pour traiter un fluide médical *ex vivo,* une surface, un appareil ou un récipient ; ou pour traiter une voie d'eau, de l'eau, une boisson, des denrées alimentaires ou un produit cosmétique, dans laquelle la ou les cellules hôtes sont constituées par, ou sur, le fluide, la surface, l'appareil, le récipient, la voie d'eau, l'eau, la boisson, les denrées alimentaires ou le produit cosmétique.

6. Utilisation destinée à être utilisée, ou vecteur destiné à être utilisé, selon l'une quelconque des revendications précédentes, dans laquelle, ou dans lequel, alternativement un ARNcr HM et un ARNtracr sont constitués par un seul ARN guide (ARNg).

7. Utilisation destinée à être utilisée, ou vecteur destiné à être utilisé, selon l'une quelconque des revendications précédentes, dans laquelle, ou dans lequel, une croissance de population de cellules hôtes est réduite par d'au moins 5 fois par comparaison avec la croissance d'une population desdites cellules hôtes non transformées avec ledit réseau HM ou une séquence nucléotidique codant ledit ARNg.

8. Utilisation selon la revendication 3, la revendication 4 ou l'une quelconque des revendications 5-7, lorsqu'elle dépend de la revendication 3 ou 4, dans laquelle une croissance de population de cellules hôtes sur une surface est empêchée.

9. Utilisation destinée à être utilisée, ou vecteur destiné à être utilisé, selon la revendication 1 ou 3, ou l'une quelconque des revendications 5-8, lorsqu'elle dépend de l'une quelconque des revendications 1 ou 3, dans laquelle, ou dans lequel, chaque vecteur manque d'une séquence codant une nucléase Cas.

10. Utilisation destinée à être utilisée, ou vecteur destiné à être utilisé, selon l'une quelconque des revendications précédentes, dans laquelle, ou dans lequel, le vecteur comprend une séquence d'ADN pour exprimer une séquence d'ARNtracr.

11. Utilisation destinée à être utilisée, ou vecteur destiné à être utilisé, selon l'une quelconque des revendications précédentes, dans laquelle, ou dans lequel, la population mixte comprend *S. thermophilus, L. Lactis* et *E. coli.*

12. Utilisation destinée à être utilisée, ou vecteur destiné à être utilisé, selon l'une quelconque des revendications

précédentes, dans laquelle, ou dans lequel, la nucléase Cas est une nucléase Cas de type I, une nucléase Cas de type II ou une nucléase Cas de type III.

13. Vecteur destiné à être utilisé, ou utilisation, selon l'une quelconque des revendications précédentes, le vecteur comprenant des séquences d'acides nucléiques pour exprimer une pluralité de différents ARNcr (par exemple, constitués par des ARNg), dans lequel, ou dans laquelle, un premier ARNcr desdits ARNcr est capable de s'hybrider à une première séquence d'acides nucléiques dans ladite cellule hôte; et un second ARNcr desdits ARNcr est capable de s'hybrider à une seconde séquence d'acides nucléiques dans ladite cellule hôte, dans lequel, ou dans laquelle, ladite seconde séquence est différente de ladite première séquence ; et

(a) la première séquence est constituée par un gène de résistance aux antibiotiques (ou d'un ARN de celui-ci) et la seconde séquence est constituée par un gène de résistance aux antibiotiques (ou d'un ARN de celui-ci) ; facultativement dans lequel, ou dans laquelle, les gènes sont différents ;
(b) la première séquence est constituée par un gène de résistance aux antibiotiques (ou d'un ARN de celui-ci) et la seconde séquence est constituée par un gène essentiel ou de virulence (ou d'un ARN de celui-ci) ;
(c) la première séquence est constituée par un gène essentiel (ou d'un ARN de celui-ci) et la seconde séquence est constituée par un gène essentiel ou de virulence (ou d'un ARN de celui-ci) ; ou
(d) la première séquence est constituée par un gène de virulence (ou d'un ARN de celui-ci) et la seconde séquence est constituée par un gène essentiel ou de virulence (ou d'un ARN de celui-ci).

14. Vecteur destiné à être utilisé, ou utilisation, selon l'une quelconque des revendications 1, 3 et 6-12 lorsqu'elle dépend de la revendication 1 ou 3, le vecteur comprenant plus de 1,4 kb de séquence d'ADN exogène, dans lequel, ou dans laquelle, l'ADN exogène code un ou plusieurs composants d'un système CRISPR/Cas et comprend un réseau modifié pour exprimer des ARNcr HM ou des ARNg dans des cellules hôtes ; dans lequel, ou dans laquelle, au moins 2 ARNc ou ARNg différents sont codés par l'ADN exogène (par exemple, par au moins 2 réseaux CRISPR HM).

15. Vecteur destiné à être utilisé, ou utilisation, selon l'une quelconque des revendications 1, 3 et 6-14 lorsqu'elle dépend de la revendication 1 ou 3, dans lequel le vecteur est dépourvu d'une séquence nucléotidique codant une nucléase Cas qui est apparentée au ou aux ANRc ou au ou aux ARNg.

FIGURE 1

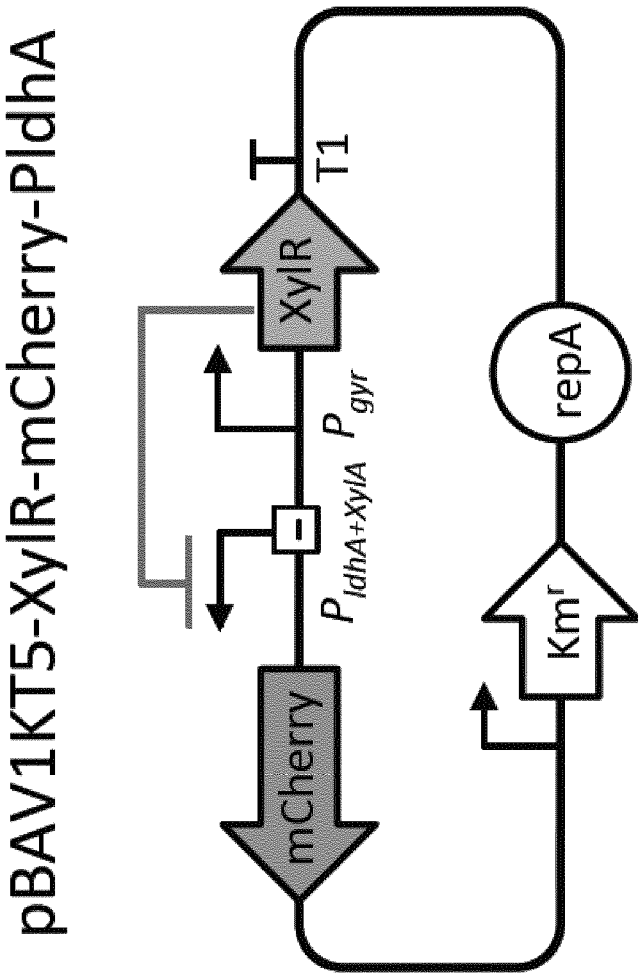

pBAV1KT5-XylR-mCherry-PldhA
T1
XylR
P_{gyr}
$P_{ldhA+XylA}$
repA
Km^r
mCherry

FIGURE 2

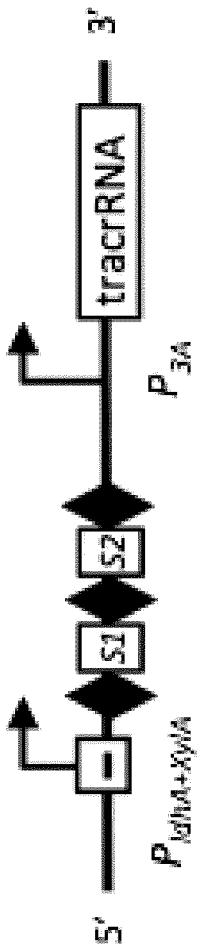


FIGURE 3

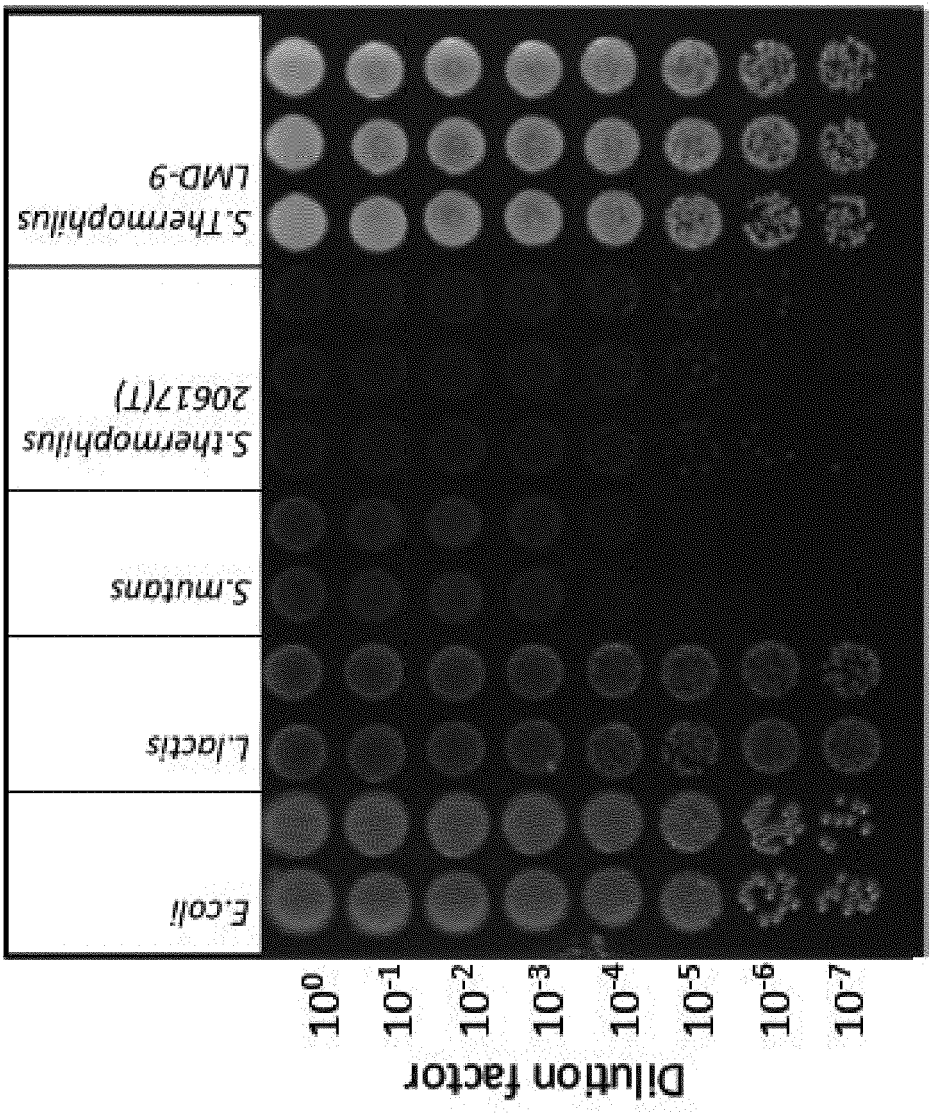

E.coli
L.lactis
S.mutans
S.thermophilus 20617(T)
S.Thermophilus LMD-9
Dilution factor
10^0
10^-1
10^-2
10^-3
10^-4
10^-5
10^-6
10^-7

FIGURE 4

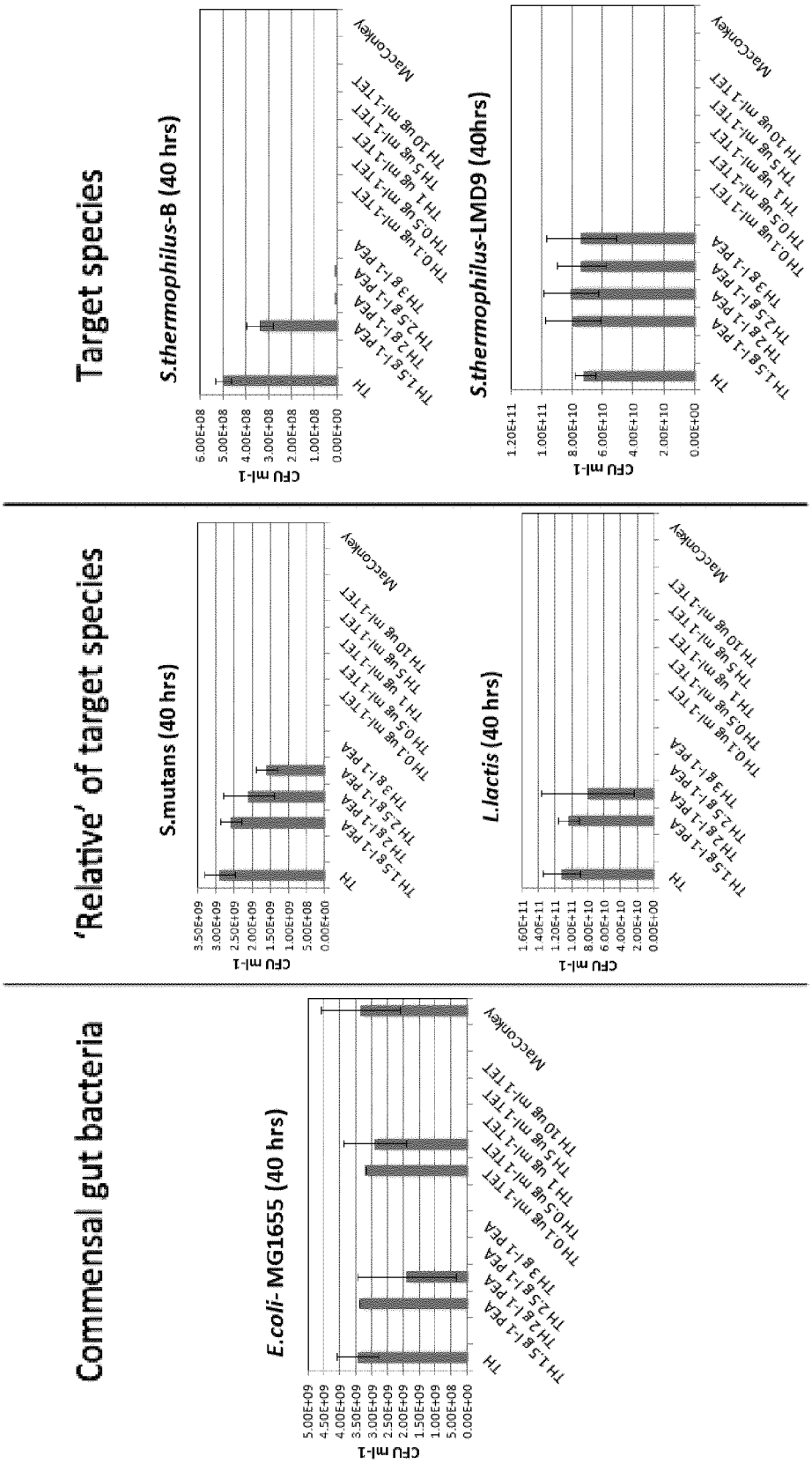

Commensal gut bacteria
'Relative' of target species
Target species
E.coli- MG1655 (40 hrs)
S.mutans (40 hrs)
L.lactis (40 hrs)
S.thermophilus-B (40 hrs)
S.thermophilus-LMD9 (40hrs)
CFU ml-1
TH
TH 1.5 E-1 PEA
TH 2 E-1 PEA
TH 2.5 E-1 PEA
TH 3 E-1 PEA
TH 0.1 ug ml-1 TET
TH 0.5 ug ml-1 TET
TH 1 ug ml-1 TET
TH 5 ug ml-1 TET
TH 10 ug ml-1 TET
MacConkey

FIGURE 5

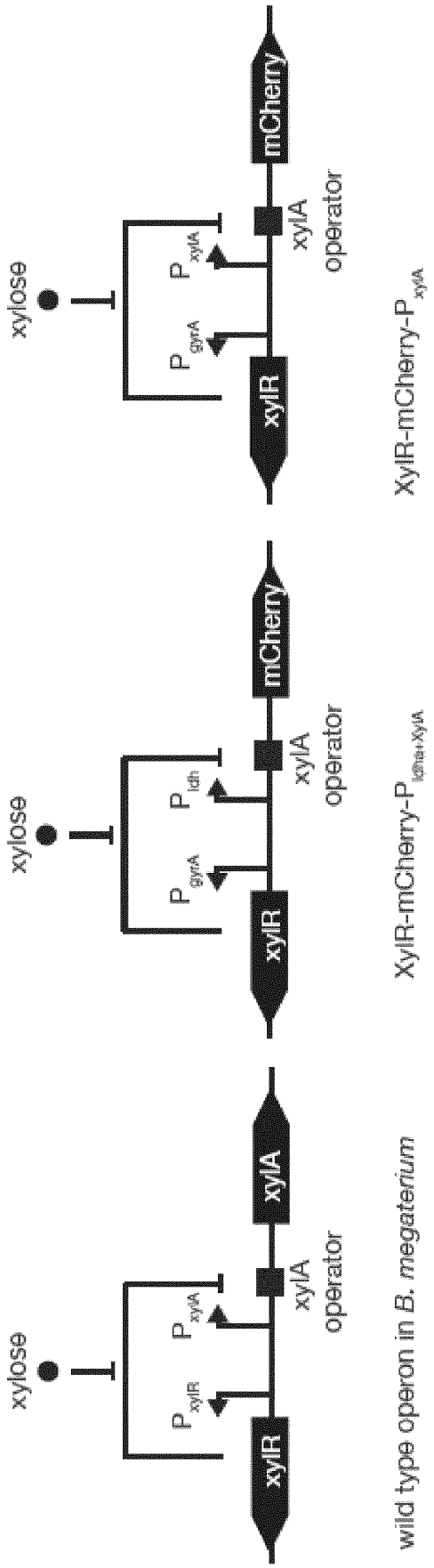

xylose
xylR
P_{xylR}
P_{xylA}
xylA
operator
xylA
wild type operon in B. megaterium
xylose
xylR
P_{gyrA}
P_{ldh}
xylA
operator
mCherry
XylR-mCherry-$P_{ldh+xylA}$
xylose
xylR
P_{gyrA}
P_{xylA}
xylA
operator
mCherry
XylR-mCherry-P_{xylA}

FIGURE 6

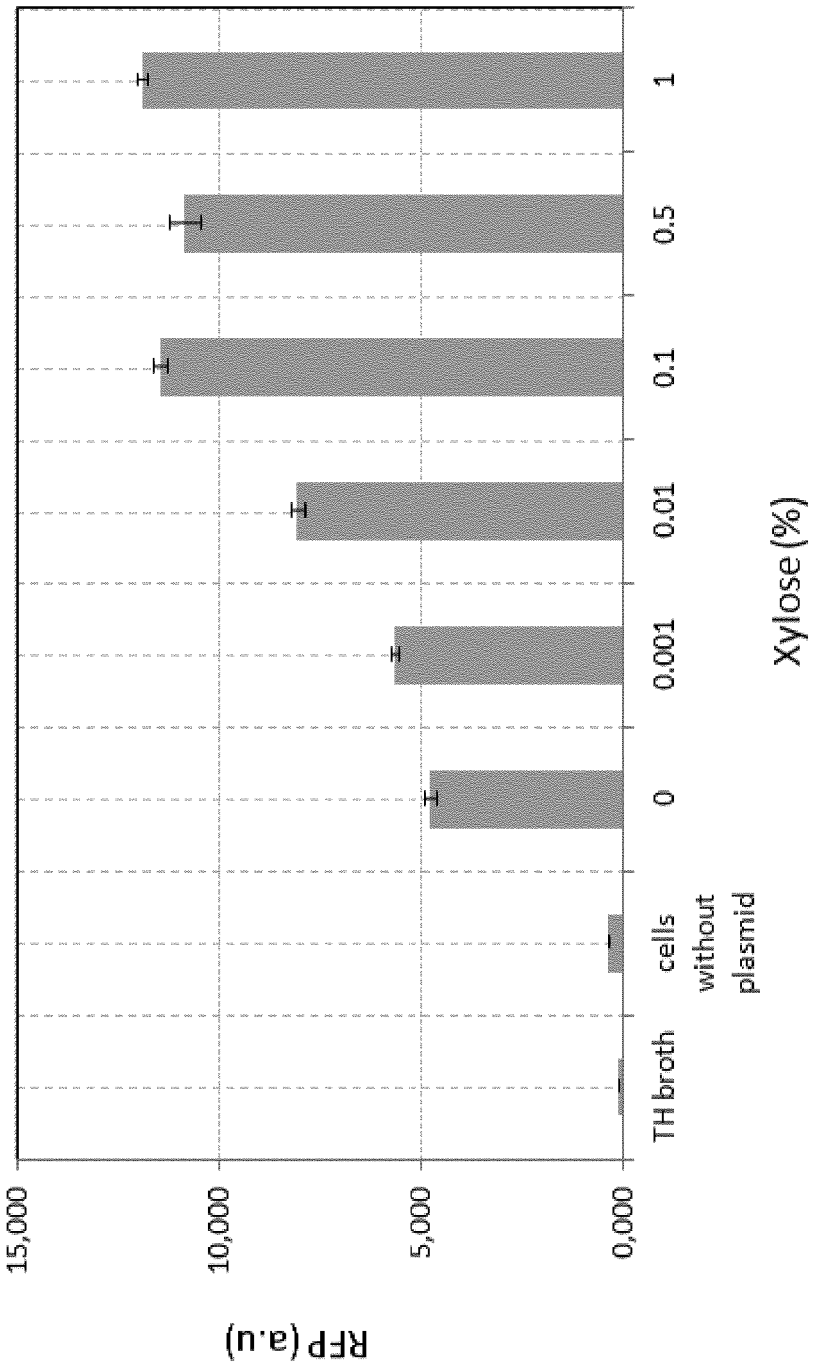

RFP (a.u)
15,000
10,000
5,000
0,000
TH broth
cells without plasmid
0
0.001
0.01
0.1
0.5
1
Xylose (%)

FIGURE 7

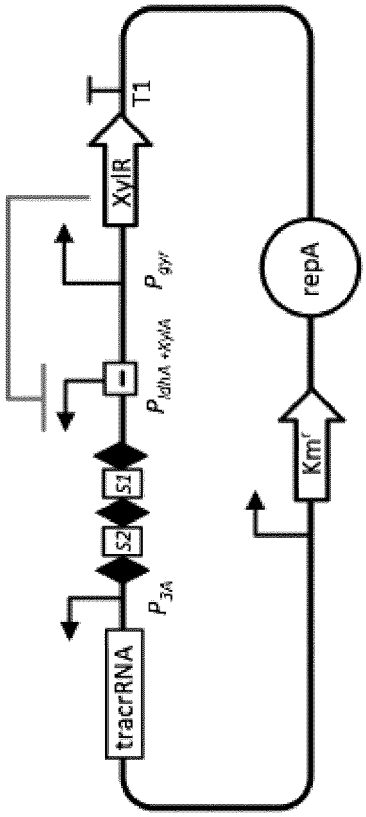


FIGURE 8

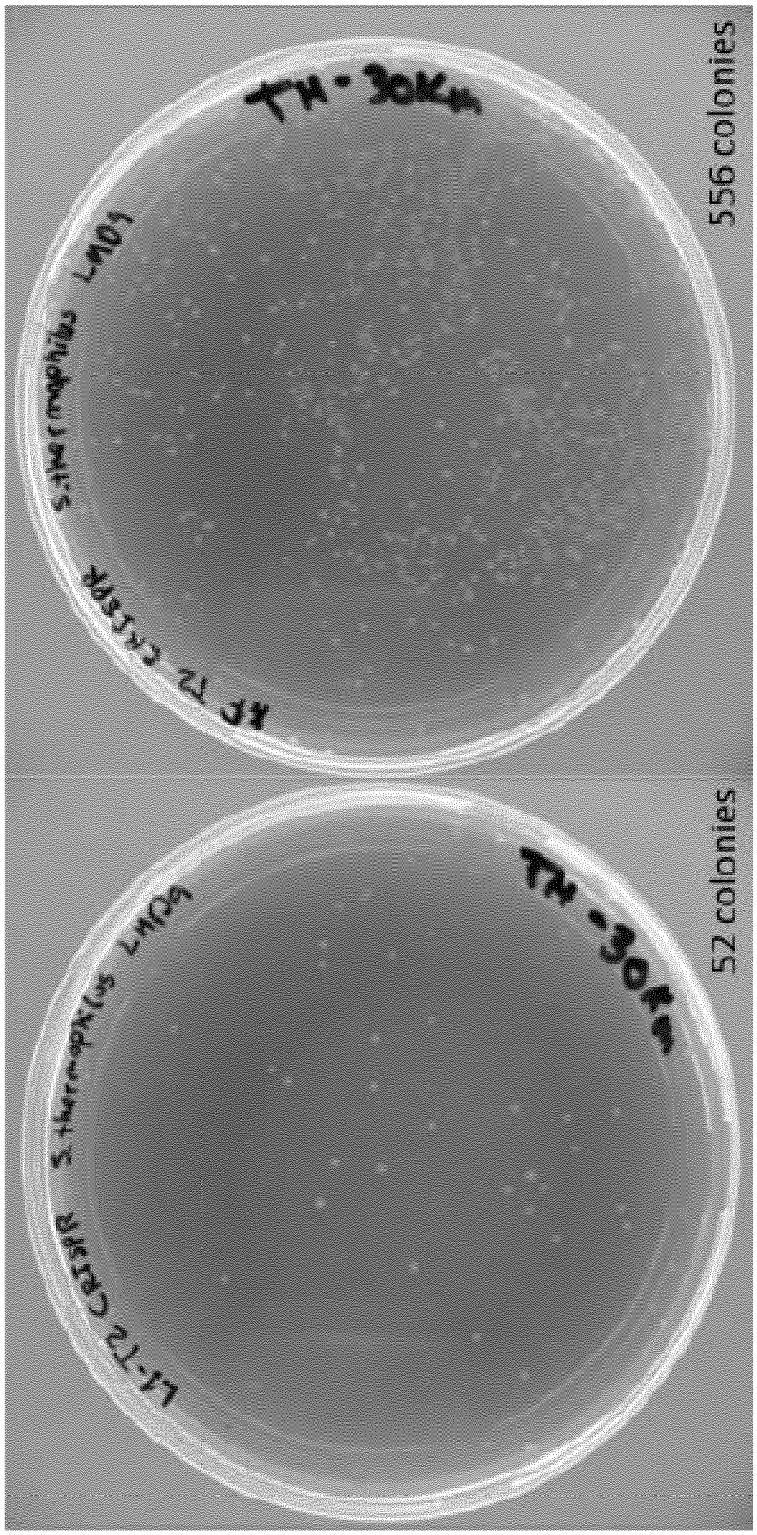


FIGURE 9

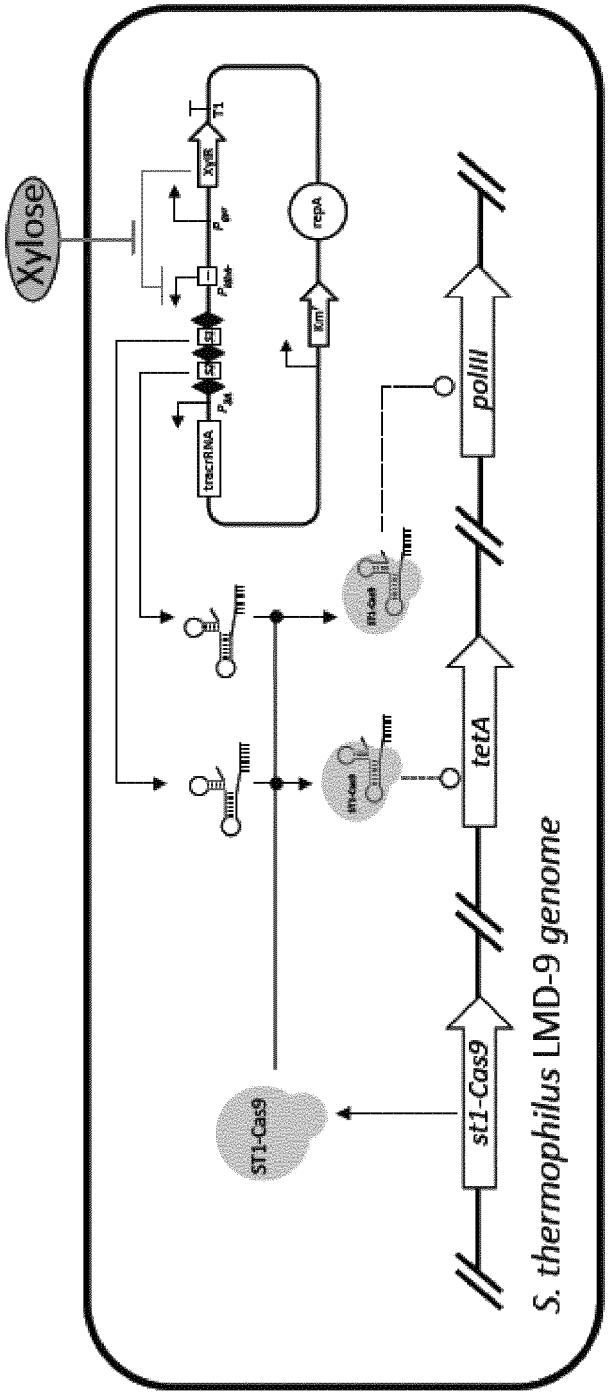

Xylose
XylR
T1
repA
Km
tracrRNA
st1-Cas9
ST1-Cas9
tetA
polIII
S. thermophilus LMD-9 genome

FIGURE 10

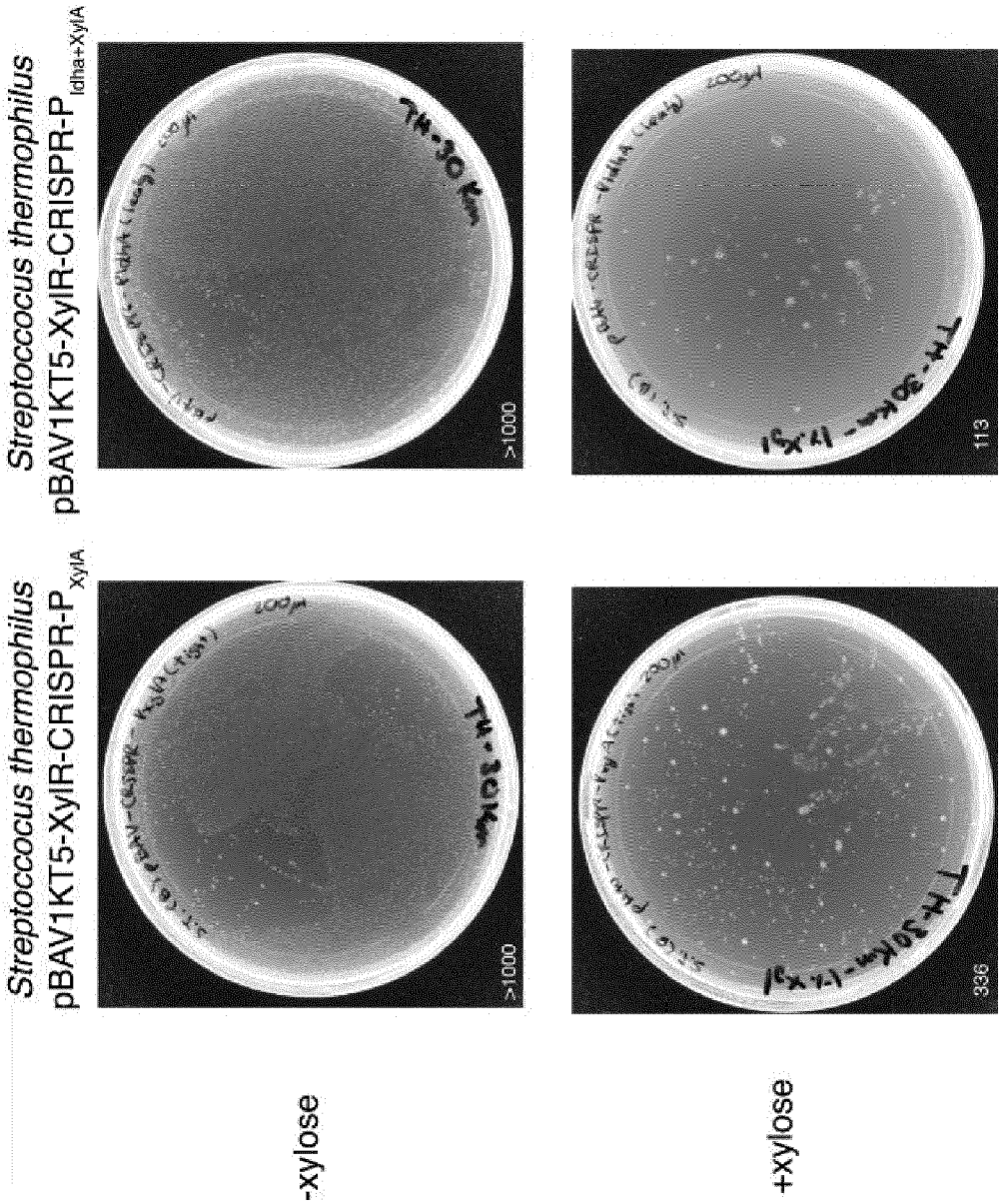

Streptoccocus thermophilus
pBAV1KT5-XylR-CRISPR-P XylA
Streptoccocus thermophilus
pBAV1KT5-XylR-CRISPR-P Idha+XylA
>1000
>1000
113
336
-xylose
+xylose

FIGURE 11

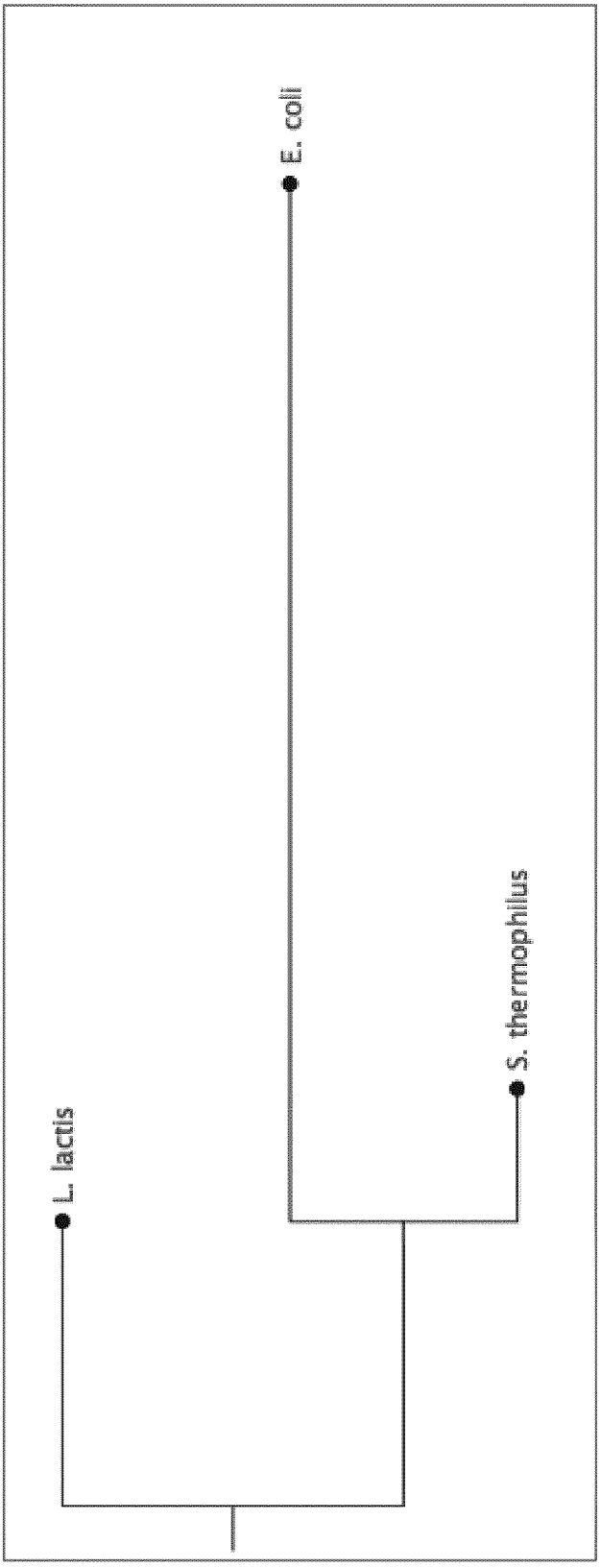
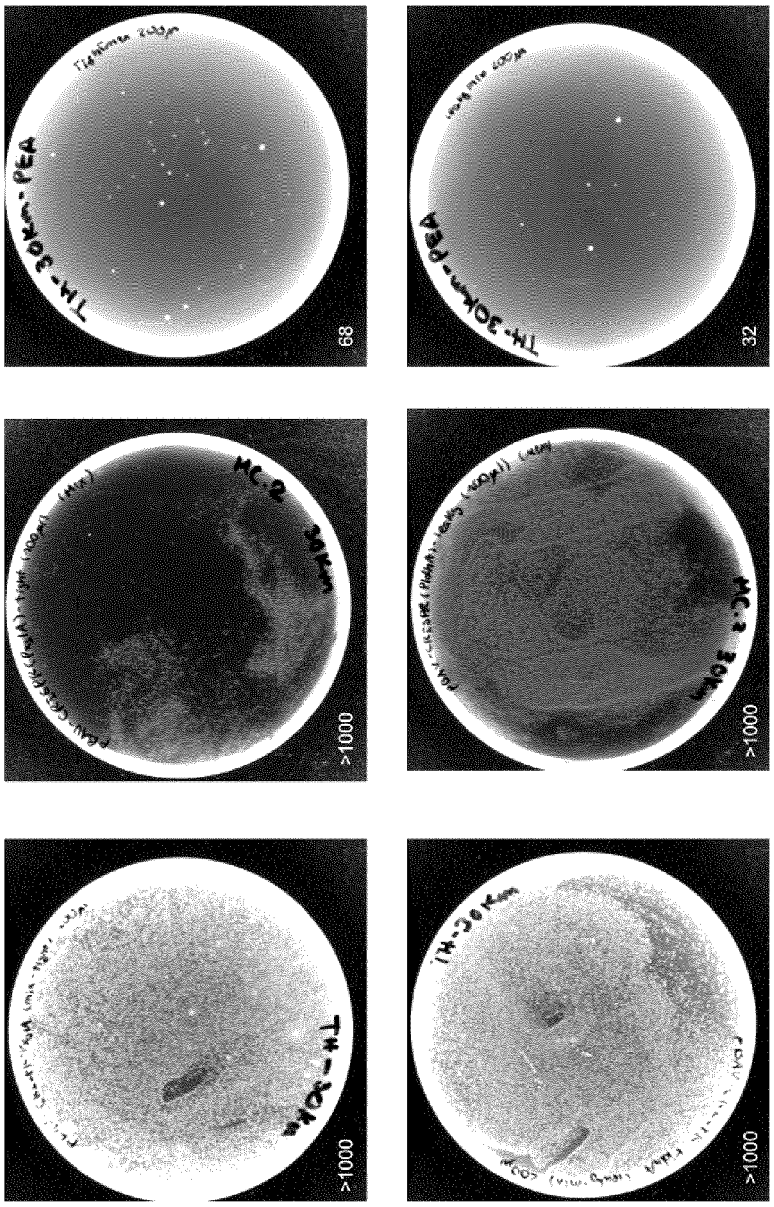


FIGURE 12


Todd Hewitt agar
MacConkey agar
Todd Hewitt agar supplemented with 2.5 g l-1 PEA
S.thermophilus, L.lactis and E.coli with plasmid pBAV1KT5-PXylA
S.thermophilus, L.lactis and E.coli with plasmid pBAV1KT5-PldhA
>1000
>1000
68
>1000
32

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO 2010075424 A **[0004]**
- WO 2014124226 A **[0826] [0864] [0865] [0866]**
- WO 2013176772 A **[0866]**

Non-patent literature cited in the description

- **TURNBAUGH, P. J. ; R. E. LEY ; M. A. MAHOWALD ; V. MAGRINI ; E. R. MARDIS ; J. I. GORDON.** An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature,* 2006, vol. 444, 1027-1131 **[0061]**
- *Nat Commun.,* 24 July 2014, vol. 5, 4498 **[0168]**
- **DUTILH BE.** *A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes* **[0168]**
- *FEBS Lett.,* 03 June 2010, vol. 584 (11), 2421-6 **[0169]**
- **MELLO L.** *Mining metagenomic data for novel domains: BACON, a new carbohydrate-binding module,* 21 April 2010 **[0169]**
- **ROBERTS et al.** *Trends Microbiol.,* June 2009, vol. 17 (6), 251-8 **[0205]**
- **ROBERTS A ; MULLANY P.** *A modular master on the move: the Tn916 family of mobile genetic elements,* 20 May 2009 **[0205]**
- **KAVEH E et al.** Characterization of Bacteria in Ballast Water Using MALDI-TOF Mass Spectrometry. *PLoS One.,* 07 June 2012, vol. 7 (6), e38515 **[0660]**
- **ANSARI ZA.** Marine bioinvasion: Concern for Ecology and Shipping. *Current Science,* 2002, vol. 83 (3), 214-218 **[0667]**
- **AZAM F ; MALFATTI F.** Microbial structuring of marine ecosystems. *Nature Reviews Microbiology,* 2007, vol. 5, 782-791 **[0667]**
- **BELKIN S ; COLWELL RR.** *Ocean and health: pathogens in the marine environment.,* 2005 **[0667]**
- **CARLTON JT ; GELLER JB.** Ecological roulette: the global transfer of non-indigenous marine organisms. *Science,* 1993, vol. 261, 78-82 **[0667]**
- **CARMAN KR ; DOBBS FC.** Epibiotic microorganisms in copepods and other marine crustaceans. *Microscopy Research and Technique,* 1997, vol. 37, 116-135 **[0667]**
- **CARNEY KJ, ; DELANY JE ; SAWANT SS ; MESBAHI E.** The effects of prolonged darkness on temperate and tropical marine phytoplankton, and their implications for ballast water risk management. *Marine Pollution Bulletin,* 2011, vol. 62 (6), 1233-1244 **[0667]**
- **COLWELL RR.** Global climate and infectious disease: the cholera paradigm. *Science,* 1996, vol. 274, 2025-2031 **[0667]**
- **CONWAY DVP ; WHITE RG ; HUGUES-DIT-CILES J ; GALLIENNE CP ; ROBINS DB.** Guide to the coastal and surface zooplankton of the southwestern Indian Ocean. In: Marine Biological Association of the United Kingdom Occasional Publication. *UKDEFRA Darwin Initiative Project 162/09/004 Zooplankton of the Mascarene Plateau,* 2003, vol. 15, 1-354 **[0667]**
- **DALEY RJ ; HOBBIE JE.** Direct counts of aquatic bacteria by a modified epifluorescence technique. *Limnology and Oceanography,* 1975, vol. 20, 875-882 **[0667]**
- Unusual or extreme high-pressure marine environments. **DEMING JW.** In: ASM Manual of Environmental Microbiology. Washington, DC: ASM Press, 1997, 366-376 **[0667]**
- **DRAKE LA ; RUIZ GM ; GALIL BS ; MULLADY TL ; FRIEDMANN DO ; DOBBS FC.** Microbial ecology of ballast water during a transoceanic voyage and the effects of open-ocean exchange. *Marine Ecology Progress Series,* 2002, vol. 233, 13-20 **[0667]**
- **DRAKE LA ; MEYER AE ; FORSBERG RL ; BAIER RE ; DOBLIN MA ; HEINEMANN S ; JOHNSON WP ; KOCH M ; RUBLEE PA ; DOBBS FC.** Potential invasion of microorganisms and pathogens via 'interior hull fouling': biofilms inside ballast water tanks. *Biological Invasions,* 2005, vol. 7, 969-982 **[0667]**
- **DRAKE LA ; DOBLIN MA ; DOBBS FC.** Potential microbial bioinvasions via ship's ballast water, sediment and biofilm. *Marine Pollution Bulletin,* 2007, vol. 55, 333-341 **[0667]**
- **DAWSON MP ; HUMPHREY BA ; MARSHALL KC.** Adhesion: A tactic in the survival strategy of a marine Vibrio during starvation. *Current Microbiology,* 1981, vol. 6, 195-199 **[0667]**
- **FOLADORI P ; BRUNI L ; ANDREOTTOLA G ; ZIGLIO G.** Effects of sonication on bacteria viability in wastewater treatment plants evaluated by flow cytometry-fecal indictors, wastewater and activated sludge. *Water Research,* 2007, vol. 41, 235-243 **[0667]**

- **HALLEGRAEFF GM ; BOLCH CJ.** Transport of diatoms and dinoflagellate resting spores in ships ballast water: implications for plankton biogeography and aquaculture. *Journal of Plankton Research,* 1992, vol. 14 (8), 1067-1084 **[0667]**
- **HARRIS JM.** The presence, nature and role of gut microflora in aquatic invertebrates: a synthesis. *Microbial Ecology,* 1993, vol. 25, 195-231 **[0667]**
- **HEIDELBERG JF ; HEIDELBERG KB ; COLWELL RR.** Bacteria of the gamma-subclass Proteobacteria associated with zooplankton in Chesapeake Bay. *Applied and Environmental Microbiology,* 2002, vol. 68, 5498-5507 **[0667]**
- The ecology of Vibrio cholerae in two Florida estuaries. **HOOD MA ; NESS GE ; RODRICK GE ; BLAKE NJ.** In: Vibrios in the Environment. Wiley, 1984, 399-409 **[0667]**
- **HUQ A ; WEST PA ; SMALL EB ; COLWELL RR.** Influence of water temperature, salinity and pH on survival and growth of toxigenic Vibrio cholerae serovar O1 associated with live copepods in laboratory microcosms. *Applied and Environmental Microbiology,* 1984, vol. 48, 420-424 **[0667]**
- **INTERNATIONAL MARITIME ORGANIZATION (IMO).** International convention for the control and management of ships' ballast water and sediments. *London: International Maritime Organization,* 2004 **[0667]**
- **OACHIMSTHAL EL ; IVANOV V ; TOY ST-L ; TAY J-H.** Bacteriological examination of ballast water in Singapore Harbour by flow cytometry with FISH. *Marine Pollution Bulletin,* 2004, vol. 49, 334-343 **[0667]**
- **JYOTI KK ; PANDIT AB.** Water disinfection by acoustic and hydrodynamic cavitation. *Biochemical Engineering Journal,* 2001, vol. 7, 201-212 **[0667]**
- A key for the identification of the more common planktonic copepoda of Indian coastal waters. **KASTURIRANGAN LR.** In: Indian National Committee on Oceanic Research. New Delhi: Council of Scientific and Industrial Research, 1963, 87 **[0667]**
- **KHANDEPARKER L ; ANIL AC.** Association of bacteria with marine invertebrates: implications for ballast water management. *Ecohealth.,* 12 July 2013, vol. 10 (3), 268-76 **[0667]**
- **KRIEG NR.** Bergey's manual of systematic bacteriology. Williams & Wilkins, 1984, vol. 1 **[0667]**
- **LEE BG ; FISHER NS.** Decomposition and release of elements from zooplankton debris. *Marine Ecology Progress Series,* 1992, vol. 88, 117-128 **[0667]**
- **LLOYD'S REGISTER.** *Ballast water treatment technology guide,* 2010 **[0667]**
- **MCFALL-NGAI MJ ; RUBY EG.** Symbiont recognition and subsequent morphogenesis as early events in an animal-bacterial mutualism. *Science,* 1991, vol. 254, 1491-1494 **[0667]**
- **MIMURA H ; KATAKURA R ; ISHIDA H.** Changes of microbial populations in a ship's ballast water and sediments on a voyage from Japan to Qatar. *Marine Pollution Bulletin,* 2005, vol. 50, 751-757 **[0667]**
- **MUNRO PM ; COLWELL RR.** Fate of Vibrio cholerae O1 in sea water microcosms. *Water Research,* 1996, vol. 30 (1), 47-50 **[0667]**
- **PFEFFER C ; OLIVER DJ.** A comparison of thiosulphate-citrate-bile salts-sucrose (TCBS) agar and thiosulphate-chloride-iodide (TCI) agar for the isolation of Vibrio species from estuarine environments. *Letters in Applied Microbiology,* 2003, vol. 36, 150-151 **[0667]**
- **PETER H ; SOMMARUGA R.** An evaluation of methods to study the gut bacterial community composition of fresh water zooplankton. *Journal of Plankton Research,* 2008, vol. 30, 997-1006 **[0667]**
- **POLZ MF ; DISTEL DL ; ZARDA B ; AMANN R ; FELBECK H ; OTTJA ; CAVANAUGH CM.** Phylogenetic analysis of a highly specific association between ectosymbiotic, sulfuroxidizing bacteria and a marine nematode. *Applied and Environmental Microbiology,* 1994, vol. 60, 4461-4467 **[0667]**
- **PRUZZO C, ; VEZZULLI L ; COLWELL RR.** Global impact of Vibrio cholerae interactions with chitin. *Environmental Microbiology,* 2008, vol. 10 (6), 1400-1410 **[0667]**
- **REHNSTAM-HOLM A-S ; GODHE A ; HÄRNSTRÖM K ; RAGHUNATH P ; SARAVANAN V ; COLLIN B ; KARUNASAGAR I ; KARUNASAGAR I.** Association between phytoplankton and Vibrio spp. along the southwest coast of India: a mesocosm experiment. *Aquatic Microbial Ecology,* 2010, vol. 58, 127-139 **[0667]**
- **RICHARD C.** Chromobacterium violaceum, opportunistic pathogenic bacteria in tropical and subtropical regions. *Bulletin de la Société de Pothologie exotique,* 1993, vol. 86, 169-173 **[0667]**
- **ROSZAK DB ; GRIMES DJ ; COLWELL RR.** Viable but nonrecoverable stage of Salmonella enteritidis aquatic systems. *Canadian Journal of Microbiology,* 1983, vol. 30, 334-338 **[0667]**
- **RUIZ GM ; RAWLINGS TK ; DOBBS FC ; DRAKE LA ; MULLADY T ; HUQ A ; COLWELL RR.** Global spread of microorganisms by ships—Ballast water discharged from vessels harbours a cocktail of potential pathogens. *Nature,* 2000, vol. 408, 49-50 **[0667]**
- **SAWANT SS ; ANIL AC ; KRISHNAMURTHY V ; GAONKAR C ; KOLWALKAR J ; KHANDEPARKER L ; DESAI DV ; MAHULKAR AV ; RONADE VV ; PANDIT AB.** Effect of hydrodynamic cavitation on zooplankton: A tool for disinfection. *Biochemical Engineering Journal,* 2008, vol. 42, 320-328 **[0667]**

- **SETH N ; CHAKRAVARTY P ; KHANDEPARKER L ; ANIL AC ; PANDIT AB.** Quantification of the energy required for the destruction of Balanus amphitrite larva by ultrasonic treatment. *Journal of the Marine Biological Association of the United Kingdom,* 2010, vol. 90 (7), 1475-1482 **[0667]**
- Biological Invasions by Nonindigenous Species in United States Waters: Quantifying the Role of Ballast Water and Sediments. **SMITH LD ; WONHAM MJ ; MCCANN LD ; REID DM ; CARLTON JT ; RUIZ GM.** Report Number CG-D-02-97. US Coast Guard Research and Development Center, 1996 **[0667]**
- **TANG KW.** Copepods as microbial hotspots in the ocean: effects of host feeding activities on attached bacteria. *Aquatic Microbial Ecology,* 2005, vol. 38, 31-40 **[0667]**
- **TANG KW ; FREUND CS ; SCHWEITZER CL.** Occurrence of copepod carcasses in the lower Chesapeake Bay and their decomposition by ambient microbes. *Estuarine, Coastal and Shelf Science,* 2006, vol. 68, 499-508 **[0667]**
- **TANG KW ; DZIALLAS C ; HUTALLE-SCHMELZER K ; GROSSART HP.** Effects of food on bacterial community composition associated with the copepod Acartia tonsa Dana. *Biological Letters,* 2009, vol. 5, 549-553 **[0667]**
- **THOMPSON CC ; THOMPSON FL ; VANDEMEULEBROECKE K ; HOSTE B ; DAWYNDT P ; SWINGS J.** Use of recA as an alternative phylogenetic marker in the family Vibrionaceae. *International Journal of Systematic and Evolutionary Microbiology,* 2004, vol. 54, 919-929 **[0667]**
- **TODD CD ; LAVERACK MS ; BOXSHALL GA.** Coastal marine zooplankton: A practical manual for students. *The Natural History Museums,* 1966, 106 **[0667]**
- **VIÑAS MD ; DIOVISOLVI NR ; CEPEDA GD.** Individual biovolume of some dominant copepod species in coastal waters off Buenos Aires Province, Argentine sea. *Brazilian Journal of Oceanography,* 2010, vol. 58 (2), 177-181 **[0667]**
- **WILLIAMS RJ ; GRIFFITHS FB, ; VAN DER WAL EJ ; KELLY J.** Cargo vessel ballast water as a vector for the transport of nonindigenous marine species. *Estuarine, Coastal and Shelf Science,* 1988, vol. 26, 409-420 **[0667]**
- **KHALIFA L.** Targeting Enterococcus faecalis biofilms with phage therapy. *Appl Environ Microbiol.,* 06 February 2015, vol. 81 (8), 2696-705 **[0687]**
- **F.A. RAN et al.** In vivo genome editing using Staphylococcus aureus Cas9. *Nature,* 2015 **[0765]**
- **JIANG W.** Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. *PLoS Genet.,* 26 September 2013, vol. 9 (9), e1003844 **[0800]**
- **KUGELBERG E et al.** Establishment of a superficial skin infection model in mice by using Staphylococcus aureus and Streptococcus pyogenes. *Antimicrob Agents Chemother.,* 2005, vol. 49, 3435-3441 **[0865]**
- **PASTAGIA M et al.** A novel chimeric lysin shows superiority to mupirocin for skin decolonization of methicillin-resistant and -sensitive Staphylococcus aureus strains. *Antimicrob Agents Chemother.,* 2011, vol. 55, 738-744 **[0865]**
- **KRUPOVIC M.** Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. *BMC Biol.,* 19 May 2014, vol. 12, 36 **[0875]**
- **WALL, J. D. ; B. J. RAPP-GILES ; M. ROUSSET.** Characterization of a small plasmid from Desulfovibrio desulfuricans and its use for shuttle vector construction. *J. Bacteriol.,* 1993, vol. 175, 4121-4128 **[0884]**
- **SEYEDIRASHTI S et al.** Molecular characterization of two bacteriophages isolated from Desulfovibrio vulgaris NCIMB 8303 (Hildenborough). *J Gen Microbiol.,* July 1992, vol. 138 (7), 1393-7 **[0884] [0978]**
- **DRIGGERS ; SCHMIDT.** Induction of Defective and Temperate Bacteriophages in Caulobacter. *J. gen. Virol.,* 1970, vol. 6, 421-427 **[0884]**
- **SEYEDIRASHTI S et al.** Induction and partial purification of bacteriophages from Desulfovibrio vulgaris (Hildenborough) and Desulfovibrio desulfuricans ATCC 13541. *J Gen Microbiol.,* July 1991, vol. 137 (7), 1545-9 **[0978]**
- **WALKER CB et al.** Recovery of temperate Desulfovibrio vulgaris bacteriophage using a novel host strain. *Environ Microbiol.,* November 2006, vol. 8 (11), 1950-9 **[0978]**
- **MIRANDA et al.** Biocorrosion of carbon steel alloys by an hydrogenotrophic sulfate-reducing bacterium Desulfovibrio capillatus isolated from a Mexican oil field separator. *Corrosion Science,* 2006, vol. 48, 2417-2431 **[0978]**
- **EYDAL et al.** Bacteriophage lytic to Desulfovibrio aespoeensis isolated from deep groundwater. *The ISME Journal,* 11 June 2009, vol. 3, 1139-1147 **[0978]**
- **WALKER CB et al.** Contribution of mobile genetic elements to Desulfovibrio vulgaris genome plasticity. *Environ Microbiol.,* September 2009, vol. 11 (9), 2244-52 **[0978]**
- **LIANG et al.** Inhibiting the growth of methicillinresistant Staphylococcus aureus in vitro with antisense peptide nucleic acid conjugates targeting the ftsZ gene. *Int J Infect Dis.,* 05 November 2014, vol. 30, 1-6 **[0983]**
- **JIANG W et al.** Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res.,* 02 September 2013, vol. 41 (20), e188 **[0985]**

- **SEED KD et al.** A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature,* 28 February 2013, vol. 494 (7438), 489-91 **[0985]**
- **SEMENOVA E et al.** Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc Natl Acad Sci U S A.,* 06 June 2011, vol. 108 (25), 10098-103 **[0985]**
- **HELER R et al.** Adapting to new threats: the generation of memory by CRISPR-Cas immune systems. *Mol Microbiol.,* 04 June 2014, vol. 93 (1), 1-9 **[0985]**
- **GOMAA A et al.** Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. *MBio.,* 28 January 2014, vol. 5 (1), e00928-13 **[0985]**
- **FINERAN PC et al.** Degenerate target sites mediate rapid primed CRISPR adaptation. *Proc Natl Acad Sci U S A.,* 07 April 2014, vol. 111 (16), E1629-38 **[0985]**
- **WIEDENHEFT et al.** Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature,* 21 September 2011, vol. 477 (7365), 486-9 **[0985]**
- **BONDY-DENOMY et al.** Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature,* 17 January 2013, vol. 493, 429-432 **[0985]**
- **NUÑEZ JK et al.** Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature,* 18 February 2015, vol. 519 (7542), 193-8 **[0985]**
- A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature,* 28 February 2013, vol. 494 (7438), 489-91 **[0996]**
- High-resolution definition of the Vibrio cholerae essential gene set with hidden Markov model-based analyses of transposon-insertion sequencing data. *Nucleic Acids Res.,* 30 July 2013, vol. 41 (19), 9033-48 **[1007]**
- A Genome-Wide Screen Reveals that the Vibrio cholerae Phosphoenolpyruvate Phosphotransferase System Modulates Virulence Gene Expression. *Infect Immun.,* 08 June 2015, vol. 83 (9), 3381-95 **[1011]**
- **BARRANGOU, R. ; FREMAUX, C. ; DEVEAU, H. ; RICHARDS, M. ; PATRICK BOYAVAL ; MOINEAU, S. ; HORVATH, P.** CRISPRProvides Acquired Resistance Against Viruses in Prokaryotes. *Science,* March 2007, vol. 315, 1709-1712, http://doi.org/10.1126/science.1138140 **[1061]**
- **BRYKSIN, A. V ; MATSUMURA, I.** Rational design of a plasmid origin that replicates efficiently in both gram-positive and gram-negative bacteria. *PloS One,* 2010, vol. 5 (10), e13244, http://doi.org/10.1371/journal.pone.0013244 **[1061]**
- **CHAN CTY ; LEE JW ; CAMERON DE ; BASHOR CJ ; COLLINS JJ.** Deadman'' and ''Passcode'' microbial kill switches for bacterial containment. *Nat Chem Biol,* December 2015, vol. 12, 1-7 **[1061]**
- **HORVATH, P. ; ROMERO, D. A. ; COÛTÉ-MONVOISIN, A.-C. ; RICHARDS, M. ; DEVEAU, H. ; MOINEAU, S. ; BARRANGOU, R.** Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. *Journal of Bacteriology,* 2008, vol. 190 (4), 1401-12, http://doi.org/10.1128/JB.01415-07 **[1061]**
- **LUDWIG, E. S. ; KLIPPER, R. ; MAGRUM L. ; WOSE C. ; STACKEBRANDT, E.** The phylogenetic position of Streptococcus and Enterococcus. *Journul of Gencwl Microhiologj.,* 1985, vol. 131, 543-55 1 **[1061]**
- **MERCENIER, A.** Molecular genetics of Streptococcus thermophilus. *FEMS Microbiology Letters,* 1990, vol. 87 (1-2), 61-77, http://doi.org/10.1016/0378-1097(90)90697-0 **[1061]**
- **ANTUNAC, N. ; HAVRANEK, J.** Taxonomy, physiology and growth of Lactococcus lactis: a review. *Mljekarstvo,* 2001, vol. 51 (1), 35-48, http://hrcak.srce.hr/index.php?show=clanak_download&id_clanakJezik=2828 **[1061]**
- **SAPRANAUSKAS, R. ; GASIUNAS, G. ; FREMAUX, C. ; BARRANGOU, R. ; HORVATH, P. ; SIKSNYS, V.** The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic Acids Research,* 2011, vol. 39 (21), 9275-9282 **[1061]**
- **SOMKUTI, G. A. ; STEINBERG, D. H.** Genetic transformation of Streptococcus thermophilus by electroporation. *Biochimie,* 1988, vol. 70 (4), 579-585 **[1061]**
- **SORG, R. A. ; KUIPERS, O. P. ; VEENING, J.-W.** Gene expression platform for synthetic biology in the human pathogen Streptococcus pneumoniae. *ACS Synthetic Biology,* 2014, vol. 4 (3), 228-239 **[1061]**
- **SUVOROV, A.** Transformation of group A streptococci by electroporation. *FEMS Microbiology Letters,* 1988, vol. 56 (1), 95-99 **[1061]**
- **XIE, Z. ; QI, F. ; MERRITT, J.** Development of a tunable wide-range gene induction system useful for the study of streptococcal toxin-antitoxin systems. *Applied and Environmental Microbiology,* 2013, vol. 79 (20), 6375-84 **[1061]**
- **ZHANG, X. Z. ; ZHANG, Y. H. P.** Simple, fast and high-efficiency transformation system for directed evolution of cellulase in Bacillus subtilis. *Microbial Biotechnology,* 2011, vol. 4 (1), 98-105 **[1061]**