US 20180258459A1

(19) **United States**

(12) **Patent Application Publication**

(10) **Pub. No.: US 2018/0258459 A1**

(43) **Pub. Date: Sep. 13, 2018**

TABOR et al.

(54) **NITRATE BIOSENSOR**

(71) Applicant: **William Marsh Rice University**, Houston, TX (US)

(72) Inventors: **Jeffrey J. TABOR**, Houston, TX (US); **Brian LANDRY**, Houston, TX (US); **Nikola DYULGYAROV**, Houston, TX (US)

(21) Appl. No.: **15/760,957**

(22) PCT Filed: **Sep. 16, 2016**

(86) PCT No.: **PCT/US16/52281**

§ 371 (c)(1),
(2) Date: **Mar. 16, 2018**

Related U.S. Application Data

(60) Provisional application No. 62/220,118, filed on Sep. 17, 2015.

Publication Classification

(51) **Int. Cl.**

C12Q 1/12	(2006.01)
C12Q 1/6897	(2006.01)
C12N 9/12	(2006.01)
C12N 9/02	(2006.01)

(52) **U.S. Cl.**

CPC ***C12Q 1/12*** (2013.01); ***C12Q 1/6897*** (2013.01); *C07K 2319/80* (2013.01); ***C12Y 207/13003*** (2013.01); ***C12N 9/0004*** (2013.01); ***C12N 9/12*** (2013.01)

(57) **ABSTRACT**

A nitrate sensing biosensor and bacteria and applications relating to the use of same are described. The nitrate biosensor comprises a two component sensor system (TCS) comprising: a nitrate-sensing sensor kinase (SK) gene comprising a ligand binding domain operably coupled to a kinase domain, and, a cognate response regulator (RR) gene comprising a receiver domain operably coupled to an DNA binding domain (DBD), as well as an output promoter that binds said DBD that is operably coupled to a heterologous reporter gene.

Specification includes a Sequence Listing.

FIGURE 1: TWO COMPNENT SYSTEMS

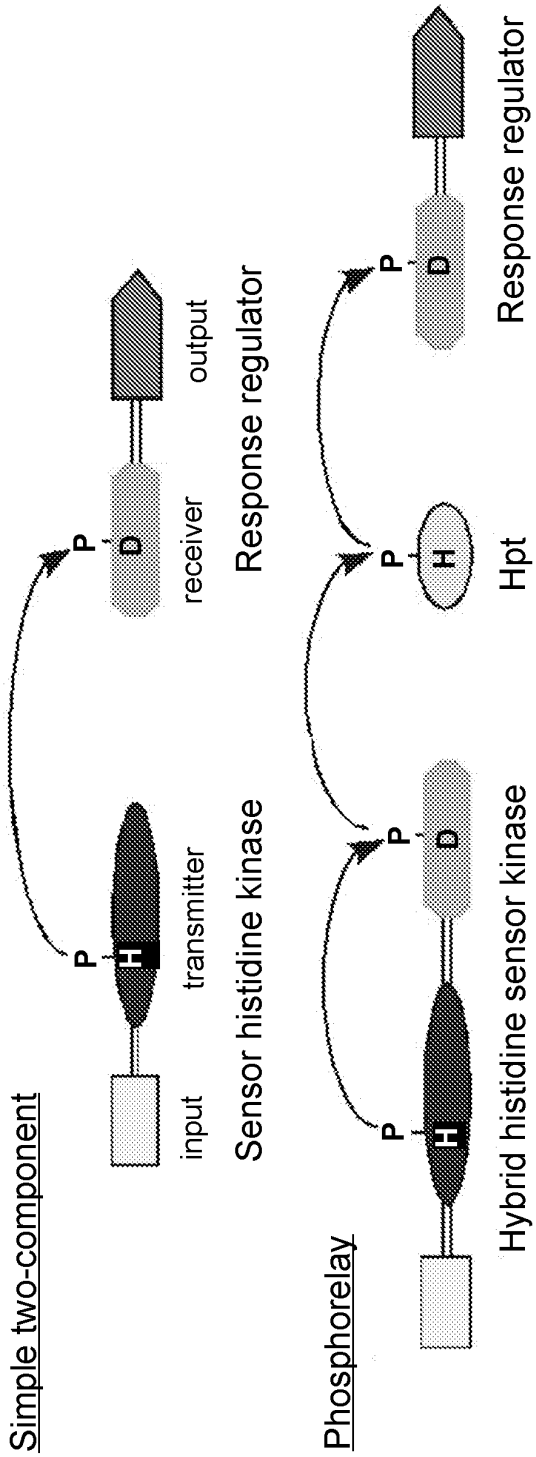


FIGURE 2: GUT MICROFLORA

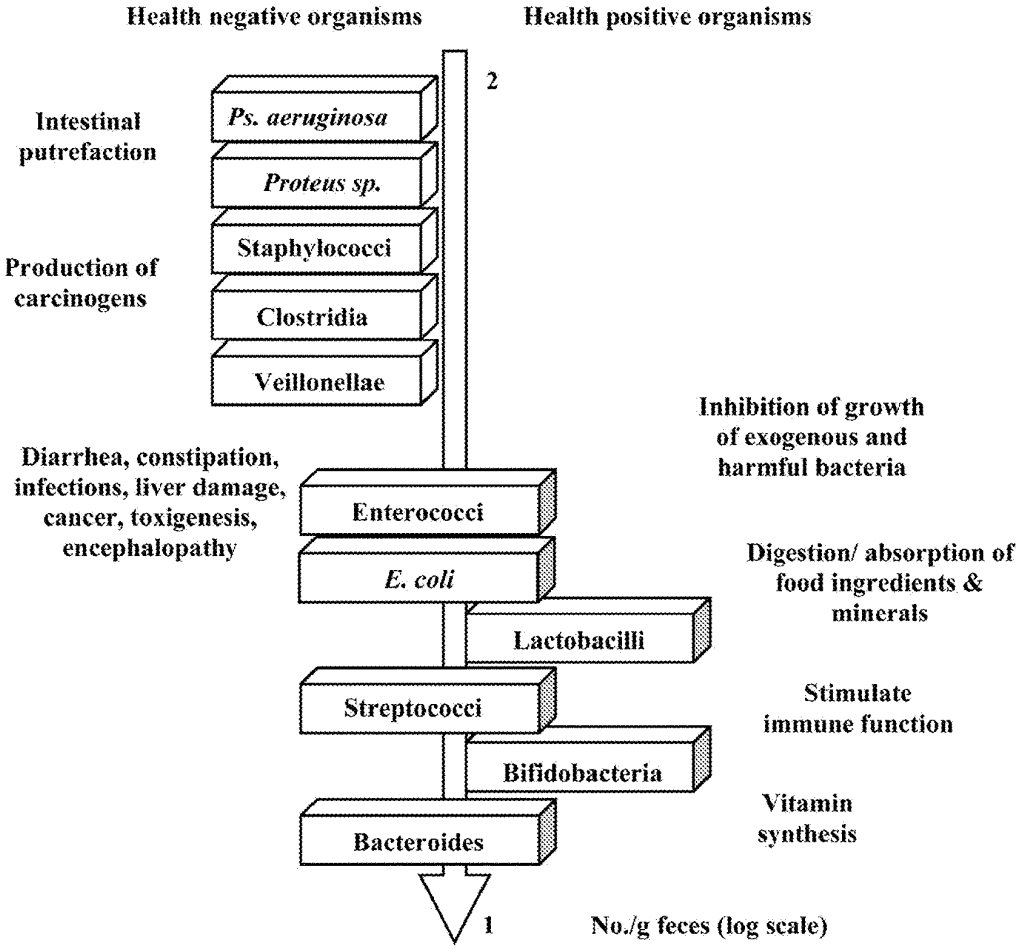


FIGURE 3: REPORTER GENE ACTIVATION BY NITRATE IN BACILLUS

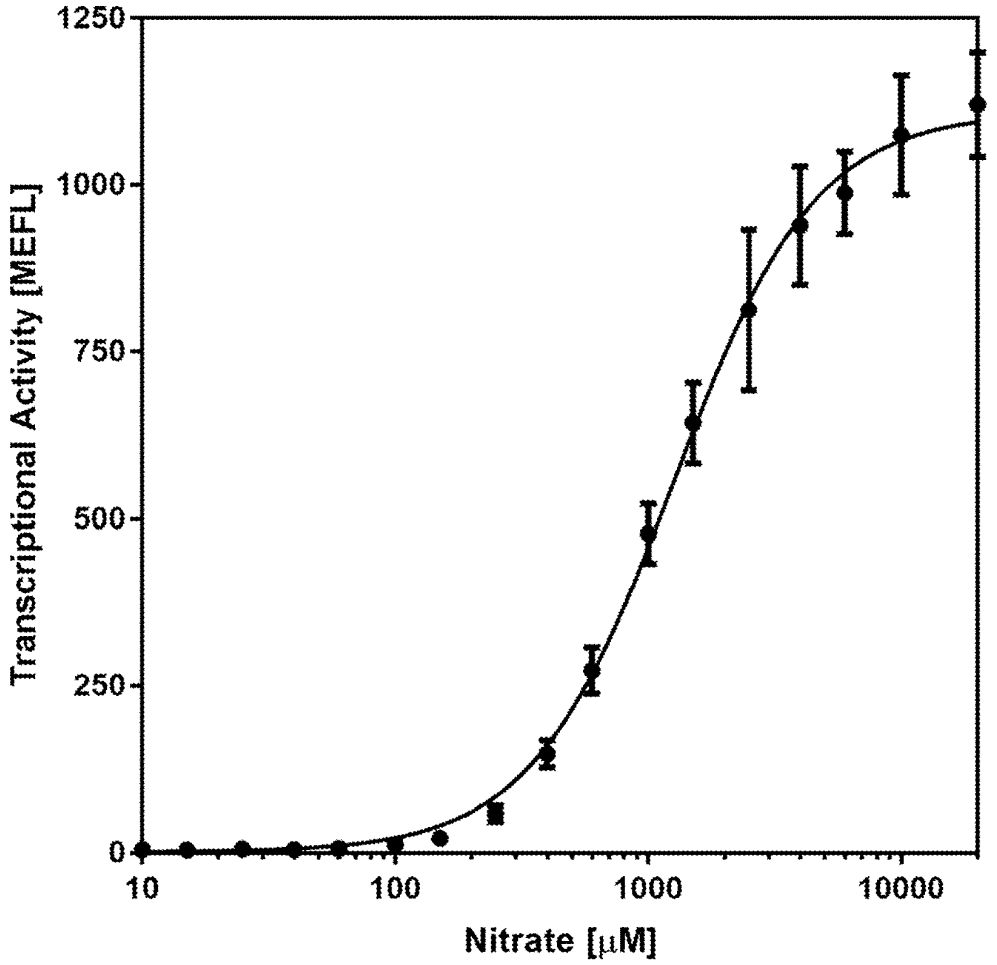


FIGURE 4: RESPONSE TO NITRATE AND FERTILIZER IN SOIL

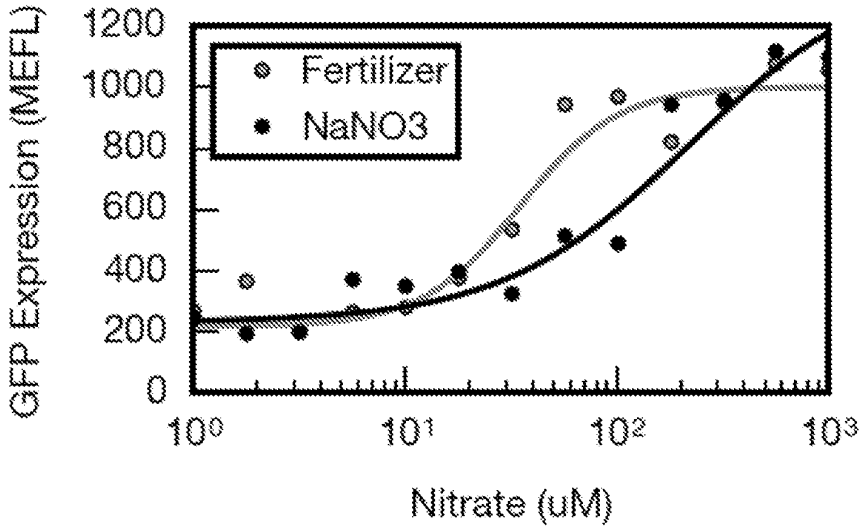


FIGURE 5: NITRATE RESPONSE IN E. COLI

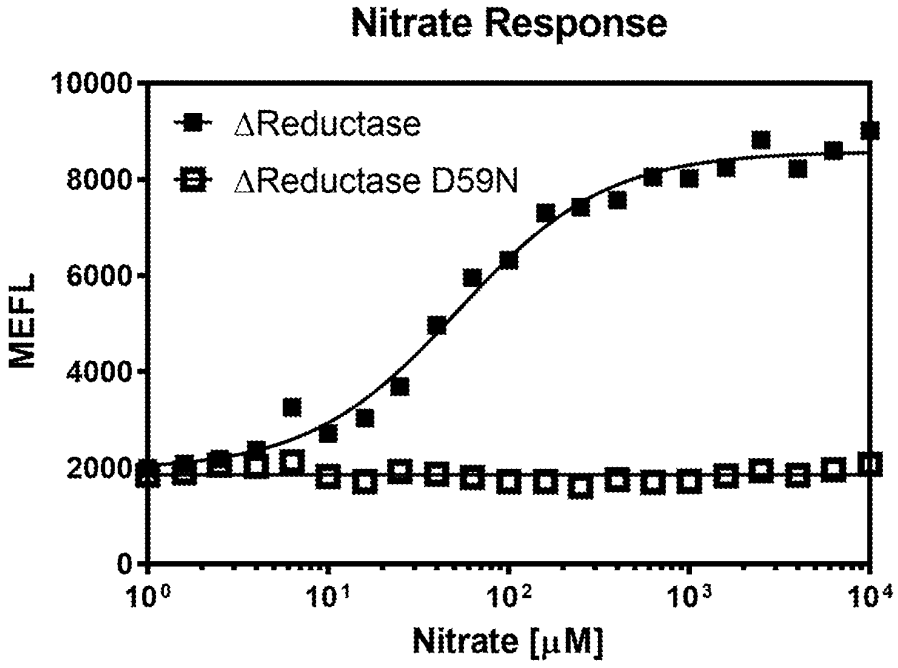
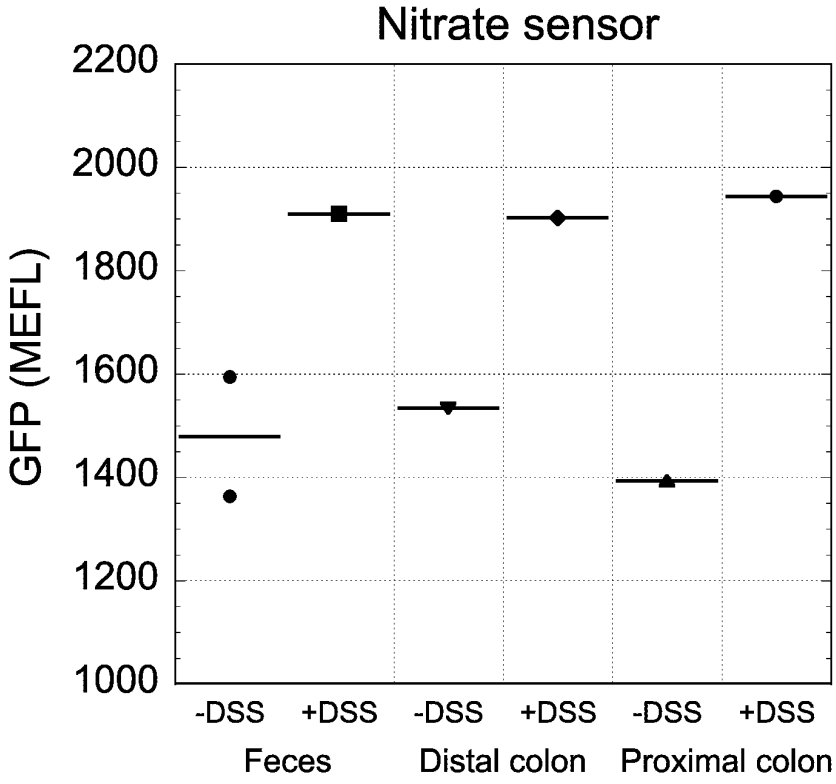


FIGURE 6: IN VIVO USE OF E.COLI NITRATE BIOSENSOR



NITRATE BIOSENSOR

PRIOR RELATED APPLICATIONS

[0001] This application claims priority to U.S. Ser. No. 62/220,118, entitled NITRATE BIOSENSOR FOR DETECTION OF GUT INFLAMMATION, filed Sep. 17, 2015, and incorporated by reference herein in its entirety for all purposes.

FEDERALLY SPONSORED RESEARCH STATEMENT

[0002] This invention was made with government support under N00014-14-1-0487 awarded by the Office of Naval Research. The government has certain rights in the invention.

FIELD OF THE DISCLOSURE

[0003] The invention includes novel materials, methods and systems using a two-component sensor kinase system for detecting nitrate.

BACKGROUND OF THE DISCLOSURE

[0004] A two-component regulatory system serves as a basic stimulus-response coupling mechanism to allow organisms to sense and respond to changes in environmental conditions. Such systems typically consist of a membrane-bound histidine kinase that senses a specific environmental stimulus and a corresponding response regulator that mediates the cellular response, mostly through differential expression of target genes.

[0005] Signal transduction occurs through the transfer of phosphoryl groups from adenosine triphosphate (ATP) to a specific histidine residue in the histidine kinases (HK). This is an autophosphorylation reaction. The response regulators (RRs) were shown to be phosphorylated on an aspartate residue and to be protein phosphatases for the histidine kinases. The response regulators are therefore enzymes with a covalent intermediate that alters response-regulator output function. Phosphorylation causes the response regulator's conformation to change, usually activating an attached output domain, which then leads to the stimulation (or repression) of expression of target genes. The level of phosphorylation of the response regulator controls its activity. Some HKs are bifunctional, catalyzing both the phosphorylation and dephosphorylation of their cognate RR. The input stimuli can regulate either the kinase or phosphatase activity of the bifunctional HK. See e.g., FIG. **1**.

[0006] Two-component signal transduction systems thus enable bacteria to sense, respond, and adapt to a wide range of environments, stressors, and growth conditions. These systems have been adapted to respond to a wide variety of stimuli, including nutrients, cellular redox state, changes in osmolarity, quorum signals, antibiotics, temperature, chemoattractants, pH and more. Some bacteria can contain up to as many as 200 two-component sensor systems that need tight regulation to prevent unwanted cross-talk.

[0007] In *Escherichia coli*, for example, the EnvZ/OmpR osmoregulation system controls the differential expression of the outer membrane porin proteins OmpF and OmpC. The KdpD sensor kinase proteins regulate the kdpFABC operon responsible for potassium transport in bacteria including *E. coli* and *Clostridium acetobutylicum*. The N-terminal domain of this protein forms part of the cytoplasmic region of the protein, which may be the sensor domain responsible for sensing turgor pressure.

[0008] A variant of the two-component system is the phospho-relay system. See FIG. **1**, lower panel. Here a hybrid HK autophosphorylates and then transfers the phosphoryl group to an internal receiver domain, rather than to a separate RR protein. The phosphoryl group is then shuttled to histidine phosphotransferase (HPT) and subsequently to a terminal RR, which can evoke the desired response.

[0009] Signal transducing histidine kinases are the key elements in two-component signal transduction systems. Examples of histidine kinases are EnvZ, which plays a central role in osmoregulation, and CheA, which plays a central role in the chemotaxis system. Histidine kinases usually have an N-terminal ligand-binding domain and a C-terminal kinase domain, but other domains may also be present. The kinase domain is responsible for the autophosphorylation of the histidine with ATP, the phosphotransfer from the kinase to an aspartate of the response regulator, and (with bifunctional enzymes) the phosphotransfer from aspartyl phosphate back to ADP or to water. The kinase core has a unique fold, distinct from that of the Ser/Thr/Tyr kinase superfamily.

[0010] HKs can be roughly divided into two classes: orthodox and hybrid kinases. Most orthodox HKs, typified by the *E. coli* EnvZ protein, function as periplasmic membrane receptors and have a signal peptide and transmembrane segment(s) that separate the protein into a periplasmic N-terminal sensing domain and a highly conserved cytoplasmic C-terminal kinase core. Members of this family, however, have an integral membrane sensor domain. Not all orthodox kinases are membrane bound, e.g., the nitrogen regulatory kinase NtrB (GlnL) is a soluble cytoplasmic HK.

[0011] Hybrid kinases contain multiple phosphodonor and phosphoacceptor sites and use multi-step phospho-relay schemes instead of promoting a single phosphoryl transfer. In addition to the sensor domain and kinase core, they contain a CheY-like receiver domain and a His-containing phosphotransfer (HPt) domain.

[0012] It is possible to identify TCSs from bacterial genome sequences by computational methods, such as homology and/or domain searching. However, such TCSs typically sense unknown inputs and control unknown output genes. Because both key pieces of information are lacking, and the microbes that contain them are often un-culturable or difficult to genetically manipulate in the laboratory, making it very difficult to identify the inputs that they sense. Therefore, while TCSs have tremendous medical, industrial and basic research applications, they have not yet fully been exploited.

[0013] This application explores the use of particular TCSs for use in detecting nitrate, as well as disease such as gut inflammation or infection, but other uses are also described herein.

[0014] The gut nurtures growth of fermentative anaerobes (such as *Clostridia, Bacteroidia*, and the like, see FIG. **2**, that convert complex polysaccharides into simple end products that the host uses for energy and strengthen the intestinal barrier. The various toxins produced by minor gut members (e.g. H_2S) can trigger an inflammatory response, resulting in reactive oxygen (e.g. O_2^-) and nitrogen species (e.g. NO) that generate oxidized compounds such as nitrate (NO_3^-) and tetrathionate ($S_4O_6^{2-}$).

[0015] Rare facultative anaerobes can respire oxidized compounds, resulting in dysbiosis—an unhealthy change in the normal bacterial ecology of e.g., the intestines or the oral cavity. Dysbiosis then results in a weakened intestinal barrier, which permits bacterial toxins (e.g. lipopolysaccharide) to transit to bloodstream, where they can lead to metabolic syndrome, obesity, anxiety, and other symptoms.

[0016] The biological nitrogen cycle involves step-wise reduction of nitrogen oxides to ammonium salts and oxidation of ammonia back to nitrites and nitrates by plants and bacteria. Neither process was thought to have relevance to mammalian physiology. However in recent years, the salivary bacterial reduction of nitrate to nitrite has been recognized as an important metabolic conversion in humans.

[0017] Several enteric bacteria have also shown the ability of catalytic reduction of nitrate to ammonia via nitrite during dissimilatory respiration. However, the importance of this pathway in bacterial species colonizing the human intestine has been little studied. Researchers have found that the presence of 5 mM nitrate provided a growth benefit and induced both nitrite and ammonia generation in *E. coli* and *L. plantarum* bacteria grown at oxygen concentrations compatible with the content in the gastrointestinal tract. Nitrite and ammonia accumulated in the growth medium when at least 2.5 mM nitrate was present. Time-course curves suggest that nitrate is first converted to nitrite and subsequently to ammonia.

[0018] Strains of *L. rhamnosus, L. acidophilus* and *B. longum infantis* grown with nitrate produced minor changes in nitrite or ammonia levels in the cultures. However, when supplied with exogenous nitrite, NO gas was readily produced independently of added nitrate. Bacterial production of lactic acid causes medium acidification that in turn generates NO by non-enzymatic nitrite reduction. In contrast, nitrite was converted to NO by *E. coli* cultures even at neutral pH. It is thus believed that the bacterial nitrate reduction to ammonia, as well as the related NO formation in the gut, could be an important aspect of the overall mammalian nitrate/nitrite/NO metabolism and is yet another way in which the microbiome links diet and health.

[0019] Like *Salmonella*, pathogenic *E. coli*, including EPEC, EHEC, and *C. rodentium*, may benefit from intestinal inflammation. In the inflamed intestine, intestinal epithelium and recruited neutrophils and macrophages that express inducible nitric oxide synthetase (iNOS), upregulate the production of nitrate (NO^{3-}). Obligate anaerobes, such as Bacteroidetes or Firmicutes that are the vast majority of healthy microbial community in the gut, cannot utilize nitrate as an electron acceptor. Rather, nitrate reductase-harboring facultative anaerobes, such as *E. coli*, can utilize NO^{3-} to generate energy for growth, leading to a growth advantage over obligate anaerobes in the inflamed intestine. Although this mechanism of *E. coli* overgrowth within the inflamed gut involves commensal-commensal competition, pathogenic *E. coli* strains, which bear nitrate reductase genes, such as narZ, in their genome, may use a similar mechanism to acquire a growth advantage over the competitive commensal community. Furthermore, the host inflammatory environment can act as a signal to trigger and enhance virulence factor expression. Thus, pathogens can take advantage of the inflammatory response to promote their growth in host tissues.

[0020] This advantage in inflammatory conditions can lead to major blooms of enterobacteria and a dramatic alteration of the gut microbial population or dysbiosis. Therefore, an ability to detect and measure physiological concentrations of inflammatory indicators, such as nitrate, at the site of inflammation would provide a novel measure of gut health. A bacterial sensor is an ideal solution because it can pass through the gut as a non-invasive observer while providing a readout of gut health that does not require removal of human tissue. Also, many usable bacterial strains are already approved for human consumption (probiotics). As an alternative, the gut flora can be sampled and applied to the biosensors in a table top ex vivo experiment, instead of performing the assay in vivo. Purified proteins could also be used in an in vitro experiment.

[0021] Related two-component systems can be mined using bioinformatics, characterized, and incorporated into a gut-friendly host for diagnostic use. We have identified several uncharacterized two-component system sensors in *Shewanella* species and related organisms that are predicted to sense a variety of terminal electron acceptors known to be markers of inflammation. *Shewanella* species demonstrate remarkable versatility in their ability to couple reduction of terminal electron acceptors to energy production and do so using an enhanced collection of reductases and associated transcriptional regulators to fine-tune their metabolic capabilities. They are therefore ideal candidates for sensor mining because these sensors likely demonstrate strong substrate specificity and respond to a broad range of ligand concentrations that correspond to their diverse environmental niches. These sensors could be combined with the nitrate and other existing sensors developed by our lab for implementation in non-invasive diagnostics.

SUMMARY OF THE DISCLOSURE

[0022] This invention relates to nitrate biosensors made of a nitrate-sensing SK and its cognate RR, which can be rewired if needed for compatibility in the host organism or to increase the output signal. These two proteins are combined with an output promoter, responsive to the RR or rewired RR, where that output promoter is operably coupled to a reporter gene for diagnostic uses, or to a gene encoding a therapeutic protein for therapeutic uses.

[0023] In our proof of concept work, we first transported the nitrate binding and signaling protein NarX (UniProt P0AFA2) from its native host *Escherichia coli* to the probiotic host *Bacillus subtilis*.

[0024] NarX sits in the membrane of bacterial cells and binds extracellular nitrate. Upon binding nitrate, NarX undergoes a shape change, which alters its phosphorylation based signaling activity and allows it to activate its cognate RR-NarL or NarQ.

[0025] To transport the NarX signal from the membrane to the genome of *B. subtilis*, we created an engineered protein (NarL/YdfI) composed of the first half of the *E. coli* NarL (UniProt P0AF28) protein and the second half of the *B. subtilis* YdfI protein (GenBank BAA19376.1). This rewired RR is capable of receiving the phosphorylation-based signal from the NarX protein. This was required because previous attempts at transferring the natural NarL protein with its associated promoter failed to result in functional transcription in *B. subtilis*.

[0026] When the hybrid NarL/YdfI receives the signal from the activate SK, it in turn interacts with the third element in our system, the native YdfJ promoter (P_{YdfJ}) in *B. subtilis*, which has been operatively coupled to another gene,

thus changing its expression. P_{ydfJ} is composed of a DNA sequence that can interact with the engineered NarL/YdfI protein, and upon interaction, stimulate production of an arbitrary RNA transcript, which can encode reporter proteins, such as GFP, enabling measurement of nitrate concentration, or therapeutic proteins such as those that make Polymyxin B, an antibiotic.

[0027] Thus, by genetic engineering we are enabling the creation of smart bacteria, which can be administered as therapeutic agents. These bacteria are capable of diagnosing the disease state of the patient inside their own body and upon diagnosis the bacteria produce relevant therapeutic molecule to treat the disease. A key target of these treatments are autoimmune diseases such as arthritis, diabetes, and irritable bowel syndrome, in which up-regulation of immune signaling leads to the production of nitric oxide and its oxidized form nitrate. We have created a novel protein based nitrate sensing system and demonstrated its functionality in bacteria. This will enable smart bacteria to diagnose autoimmune diseases based on the concentration of nitrate present within human gut, enabling non-invasive diagnosis and treatment of a wide variety of highly prevalent diseases.

[0028] Some major potential uses are:

[0029] 1. To create novel therapeutic bacteria, which are capable of sensing nitrate in the gut, to diagnose diabetes and response by treatment with polymyxin B, a molecule that has been shown to ameliorate diabetes symptoms by eliminating the causative bacterial produced chemicals.

[0030] 2. To create novel therapeutic bacteria, which are capable of sensing nitrate in the gut, to diagnose irritable bowel syndrome and respond by producing *Lactobacillus rhamnosus* GG protein p40 which has been shown to activate human cells to increase production of protective mucus coating of the intestine which alleviates disease symptoms.

[0031] 3. Create novel diagnostic bacteria, which are administered orally, measure nitrate concentrations while transiting the gut, and then are collected in fecal matter. Expression of reporter protein in response to nitrate can then be measured to discover the nitrate concentration within the patient's gut, allowing for diagnosis of a wide range of diseases such as diabetes, IBS, or arthritis.

[0032] 4. Create novel bacteria, which can live in plant roots or soil and detect nitrate, a fertilizer component, and when there is a lack of nitrate, supply the plant by producing additional nitrate.

[0033] There are (at least) two novel features of this invention.

[0034] The most prominent source of novelty is the chimeric NarL/YdfI protein, which is a novel synthetic fusion of domains from two natural proteins. This protein is a new, never produced before, molecule with completely novel signally properties enabling nitrate sensing in *B. subtilis*.

[0035] The second novel component of this invention is the expression of the natural NarX protein in conjunction with the previously mentioned chimeric NarL/YdfI protein in *B. subtilis*. This NarX protein, and in fact, the whole family of proteins, have not been previously transported from a gram negative bacteria such as *E. coli* to a gram positive bacteria such as *B. subtilis*.

[0036] There are already documented examples (DeAngelis, 2005) of nitrate-sensing bacteria. However, we are the first to isolate the nitrate-sensing pathway from a range of competing signals such as oxygen and nitrite sensing while still maintaining extremely high change in response to nitrate. This was accomplished by moving the signaling pathway from its natural bacterial host *E. coli* to the probiotic host *B. subtilis* using newly designed signaling proteins.

[0037] The first step in creating this invention was to bioinformatically align protein sequence of the NarL signaling protein with those of similar proteins from *B. subtilis*. This allowed us to select the YdfI (36% identity) as the best target protein for a fusion. We then used the alignment to determine an ideal split point containing the first half of NarL and the second half of YdfI. These were identified by selecting the boundaries of the unstructured linker regions between the α5 and α7 domains.

[0038] We subsequently used DNA manipulation technique to create a series of DNA sequences which allowed use to produce this protein and several others in *B. subtilis*. *B. subtilis* containing this engineered DNA was then grown in our laboratory and the production of a fluorescent protein in response to nitrate was measured. This allowed us to determine the degree of nitrate sensing the engineered bacteria were capable of.

[0039] The original nitrate sensor was only sensitive to nitrate in a narrow range of nitrate concentrations. However, there exist several protein and DNA engineering techniques that enable varying the range of sensitivity of this class of proteins.

[0040] This approach could be used to engineer sensors for other chemicals whose sensing proteins are homologous to the NarX/NarL protein pair. The most likely successful candidate would be the NarQ/NarP nitrite sensing system.

[0041] There are a great variety of reporter genes that can be used herein, and GFP is only one convenient reporter. The amount or activity of the reporter protein produced is taken as a proxy for the cellular response to the target. Importantly, the reporter gene by definition is NOT the wild type downstream target gene, but is artificially coupled to the TCS to provide a more convenient readout.

[0042] Ideal reporter proteins are easy to detect and quantify (preferably noninvasively), highly sensitive and, ideally, not present in the native organism. They can be set up to detect either gene activated or deactivation. Several currently popular reporter proteins and their characteristics are listed in TABLE 1.

TABLE 1

Common spectroscopically active reporter proteins and their detection

Reporter protein	Reporter genes	Origin	Substrate	Detection method	Comments	Refs
Bacterial luciferase	luxAB* or luxCDABE	Bioluminescent bacteria*	O_2, $FMNH_2$ and long-chain aldehydes	Bioluminescence	Requires O_2; aldehyde addition is required if only luxAB is used	94, 95

TABLE 1-continued

Common spectroscopically active reporter proteins and their detection

Reporter protein	Reporter genes	Origin	Substrate	Detection method	Comments	Refs
Firefly luciferase	lucFF	Firefly (*Photinus pyralis*)	O_2, ATP and luciferin	Bioluminescence	Requires O_2	96
Click beetle luciferase	lucGR	Click beetle (*Pyrophorus plagiophthalamus*)	O_2, ATP and pholasin	Bioluminescence	Requires O_2	97
Renilla luciferase	Rluc	*Renilla reniformis*	Coelenterazine and Ca^{2+}	Bioluminescence	Requires O_2	98
β-Galactosidase	lacZ	*Escherichia coli*	Galactopyranosides‡	Chemiluminescence, colorimetry, electrochemistry and fluorescence	External substrate addition (may require cell permeabilization)	1
Fluorescent proteins	gfp. etc.	*Aequorea victoria* and additional marine invertebrates	N/A	Fluorescence	O_2 is required for maturation; different colour varieties exist	99-101
Spheroidene monooxygenase	crtA	*Rhodovulum sulfidophilum*	Spheroidene	Colorimetry	None	102
Infrared fluorescent proteins	Various	Bacteriophytochrome family	N/A	Fluorescence	None	103
FMN-based fluorescent proteins	Various	Engineered from *Bacillus subtilis* and *Pseudomonas putida*	None	Fluorescence	Functional in both oxic and anoxic conditions; requires endogenous FMN	104

N/A, not applicable.

*Most commonly used species inciude *Alivibrio fischeri* (also known as *Vibrio fischeri*), *Vibrio harveyi* and *Photorhabdus luminescens*.

‡For example, O-nitrophenyl-β-D-galactoside (ONPG), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal),4-methylumbelliferyl-β-D-galoctopyranoside, 4-aminophenyl-β-D-galactopyranside and D-luciferin-O-β-galactopyranoside.

[0043] Using the amount of reporter gene as a readout, and using standard high throughput screening methods, such as fluorimetry or flow-cytometry, we can screen potential nitrate sensing TCSs for activity using standard, high throughput laboratory assays. In this way, we can expand the range of nitrate sensor genes that can be employed herein.

[0044] Initial experiments proceeded in *E. coli* and *B. subtillus* for convenience, but the addition of genes to bacteria is of nearly universal applicability, so it will be possible to use a wide variety of organisms with the selection of suitable vectors for same. Furthermore, a number of databases include vector information and/or a repository of vectors. See e.g., Addgene.org, which provides both a repository and a searchable database allowing vectors to be easily located and obtained from colleagues. See also Plasmid Information Database (PlasmID) and DNASU having over 191,000 plasmids. A collection of cloning vectors of *E. coli* is also kept at the National Institute of Genetics as a resource for the biological research community. Furthermore, vectors (including particular ORFS therein) are usually available from colleagues.

[0045] Once an exemplary sequence is obtained, e.g., in *E. coli*, which is completely sequenced and which is the workhorse of genetic engineering and bioproduction, many additional examples proteins of similar activity can be identified by BLAST search or database search. The OMIN database is also a good resource for searching human proteins and has links to the sequences. Further, every protein record is linked to a gene record, making it easy to design genome insertion vectors. Many of the needed sequences are already available in vectors, and can often be obtained from cell depositories or from the researchers who cloned them. But, if necessary, new clones can be prepared based on available sequence information using gene synthesis or PCR techniques. Thus, it should be easily possible to obtain all of the needed sequences.

[0046] Understanding the inherent degeneracy of the genetic code allows one of ordinary skill in the art to design multiple sequences that encode the same amino acid sequence. NCBI® provides codon usage databases for optimizing DNA sequences for protein expression in various species. Using such databases, a gene or cDNA may be "optimized" for expression in probiotic strains, mice, humans, or other species using the codon bias for the species in which the gene will be expressed.

[0047] In calculating "% identity" the unaligned terminal portions of the query sequence are not included in the calculation. The identity is calculated over the entire length of the reference sequence, thus short local alignments with a query sequence are not relevant (e.g., % identity=number of aligned residues in the query sequence/length of reference sequence).

[0048] Alignments are performed using BLAST homology alignment as described by Tatusova T A & Madden T L (1999) FEMS Microbiol. Lett. 174:247-250. The default parameters were used, except the filters were turned OFF. As of Jan. 1, 2001 the default parameters were as follows: BLASTN or BLASTP as appropriate; Matrix=none for BLASTN, BLOSUM62 for BLASTP; G Cost to open gap default=5 for nucleotides, 11 for proteins; E Cost to extend gap [Integer] default=2 for nucleotides, 1 for proteins; q Penalty for nucleotide mismatch [Integer] default=−3; r reward for nucleotide match [Integer] default=1; e expect value [Real] default=10; W word size [Integer] default=11 for nucleotides, 3 for proteins; y Dropoff (X) for blast extensions in bits (default if zero) default=20 for blastn, 7 for other programs; X dropoff value for gapped alignment (in bits) 30 for blastn, 15 for other programs; Z final X dropoff value for gapped alignment (in bits) 50 for blastn, 25 for other programs. This program is available online at NCBI™ (ncbi.nlm.nih.gov/BLAST/). "Positives" includes conservative amino acid changes in addition to identities.

[0049] As used herein, a "two component system" or "two component sensor system" or "TCS" is understood to be a two protein system including a sensor kinase and a response regulator, wherein the sensor kinase when bound to its

cognate ligand, activates the response regulator which then activates the expression of relevant downstream proteins.

[0050] As used herein, a "sensor kinase" or "SK" is a protein understood to have a ligand binding domain ("LBD") operably coupled to a "kinase domain" ("KD"), such that when the LBD binds its cognate ligand (in this application nitrate), the kinase is activated.

[0051] "Cognate" refers to two components systems that function together, such that a SK will bind to its cognate RR and activate it. The SK and RR are thus cognate, meaning they function together, or are related or connected functionally.

[0052] As used herein, a "response regulator" or "RR" typically has a "receiver" or "REC" domain that is activated by the active kinase of the cognate TCS. Typically the REC domain is operably coupled to a "DNA binding domain" or "DBD," which thus can bind to and turn on relevant downstream protein expression, such as a report gene. If the native downstream cognate promoters are not known, or are insufficiently active, the DBD domain can be replaced with a more suitable one, thus "rewiring" the RR.

[0053] As used herein, a "heterologous DBD" means a DBD that comes from another protein, not the response regulator that the REC domain comes from. Typically, the DBD then binds to the DNA it is targeted to, which is itself coupled to a reporter gene that can easily be detected.

[0054] The term "output promoter" means a promoter that is responsive to the TCS used herein. It is operably coupled to a "reporter gene" or a therapeutic protein gene, meaning that the output promoter controls the expression of said gene, typically by binding the DBD of the RR.

[0055] A "reporter gene" is an easily monitored gene that is heterologous to said output promoter (thus the normal downstream target is by definition excluded), and preferably is not present in the host species. Fluorescent proteins make excellent reporters.

[0056] As used herein, reference to cells, bacteria, microbes, microorganisms and like is understood to include progeny thereof having the same genetic modifications. It is also understood that all progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations that have been added to the parent. Mutant progeny that have the same function or biological activity as screened for in the originally transformed cell are included. Where distinct designations are intended, it will be clear from the context.

[0057] The terms "operably associated" or "operably linked," as used herein, refer to functionally coupled nucleic acid sequences.

[0058] As used herein "recombinant" or "engineered" is relating to, derived from, or containing genetically engineered material. In other words, the genome was intentionally manipulated in some way by the hand-of-man.

[0059] "Reduced activity" or "inactivation" is defined herein to be at least a 75% reduction in protein activity, as compared with an appropriate control species.

[0060] Preferably, at least 80, 85, 90, 95% reduction in activity is attained, and in the most preferred embodiment, the activity is eliminated (100%, aka a "knock-out" or "null" mutants which produce undetectable levels of activity). Proteins can be inactivated with inhibitors, by mutation, or by suppression of expression or translation, and the like. Use of a frame shift mutation, early stop codon, point mutations of critical residues, or deletions or insertions, and the like, can completely inactivate (100%) gene product by completely preventing transcription and/or translation of active protein.

[0061] "Overexpression" or "overexpressed" is defined herein to be at least 150% of protein activity as compared with an appropriate control species, and preferably 200, 500, 1000%) or more, or any expression is a species that otherwise lacks the activity. Overexpression can be achieved by mutating the protein to produce a more active form or a form that is resistant to inhibition, by removing inhibitors, or adding activators, and the like. Overexpression can also be achieved by removing repressors, adding multiple copies of the gene to the cell, or up-regulating the endogenous gene, and the like.

[0062] The term "endogenous" or "native" means that a gene originated from the species in question, without regard to subspecies or strain, although that gene may be naturally or intentionally mutated, or placed under the control of a promoter that results in overexpression or controlled expression of said gene. Thus, genes from *Clostridia* would not be endogenous to *Escherichia*, but a plasmid expressing a gene from *E. coli* would be considered to be endogenous to any genus of *Escherichia*, even though it may now be overexpressed. By contrast, "wild type" means the natural functional gene/protein as it exists in nature.

[0063] The invention includes any one or more of the following embodiment(s) in any combination(s) thereof:

[0064] A genetically engineered bacteria, said bacteria overexpressing:

[0065] a two component sensor system (TCS) comprising:

[0066] a nitrate-sensing sensor kinase (SK) gene comprising a ligand binding domain operably coupled to a kinase domain, and,

[0067] a cognate response regulator (RR) gene comprising a receiver domain operably coupled to an DNA binding domain (DBD),

[0068] an output promoter that binds said DBD that is operably coupled to a heterologous reporter gene.

[0069] A genetically engineered bacteria, said bacteria expressing:

[0070] a heterologous two component sensor system (TCS) comprising:

[0071] a nitrate-sensing sensor kinase (SK) gene comprising a ligand binding domain operably coupled to a kinase domain, and,

[0072] a cognate response regulator (RR) gene comprising a receiver domain operably coupled to an DNA binding domain (DBD),

[0073] an output promoter that binds said DBD that is operably coupled to a reporter gene.

[0074] A genetically engineered bacteria, said bacteria expressing:

[0075] a two component sensor system (TCS) comprising:

[0076] a nitrate-sensing sensor kinase (SK) gene comprising a ligand binding domain operably coupled to a kinase domain; and,

[0077] a rewired cognate response regulator (RR) gene comprising a receiver domain operably coupled to a heterologous DNA binding domain (DBD) that is operable in said bacteria;

[0078] a reporter gene comprising an output promoter that is responsive to said DBD that is operably coupled to an open reading frame encoding a reporter protein.

[0079] A genetically engineered bacteria, said bacteria expressing:

[0080] a heterologous two component sensor system (TCS) comprising:

[0081] a nitrate-sensing sensor kinase (SK) gene comprising a ligand binding domain operably coupled to a kinase domain; and,

[0082] a rewired cognate response regulator (RR) gene comprising a receiver domain operably coupled to a heterologous DNA binding domain (DBD) that is operable in said bacteria;

[0083] a reporter gene comprising an output promoter that is responsive to said DBD that is operably coupled to an open reading frame encoding a reporter protein.

[0084] A bacteria as herein described, which is probiotic for use in humans. The bacteria might also be probiotic for use in other species, e.g., companion animals, as appropriate for the species of patient being treated.

[0085] A bacteria as herein described, wherein said SK gene or said RR gene or both genes are encoded on an expression vector, which can be inducible or constitutive.

[0086] A bacteria as herein described, wherein said SK gene or said RR gene or both genes integrated into a genome of said bacteria.

[0087] A bacteria as herein described, wherein said SK gene and said RR gene are encoded in a single operon.

[0088] A bacteria as herein described, wherein said reporter gene is encoded on a plasmid or other expression vector.

[0089] A bacteria as herein described, wherein said reporter gene is integrated into a genome of said bacteria.

[0090] A bacteria as herein described, comprising SEQ ID NO. 1 and an amino terminal portion of SEQ ID NO. 3 operably fused to a carboxy terminal portion of SEQ ID NO 5 containing a DNA binding site. SEQ ID. NO. 6-9 could also be used. Homologs of same are also possible. A bacteria as herein described, wherein said reporter gene encodes a fluorescent protein, such as green fluorescent protein, red fluorescent protein, far red fluorescent protein, blue fluorescent protein, orange fluorescent protein, yellow fluorescent protein, mCHERRY, mORANGE, mCITRINE, VENUS, YPET, EMERALD, or CERULEAN.

[0091] A method of screening for a gut bacteria that produces nitrate, comprising:

[0092] i) applying a bacteria as herein described to a gut or a gut sample; and,

[0093] ii) measuring activity of said reporter gene, wherein activation of said reporter gene indicates that said gut or gut sample harbors a gut bacteria that produces nitrate.

[0094] A method of screening for a nitrate-generating bacteria, comprising:

[0095] i) applying a test sample containing test bacteria to a bacteria as herein described; and, ii) measuring activity of said reporter gene, wherein activation of said reporter gene at a level greater than a normal level indicates that said test sample harbors an excess of a nitrate-generating bacteria.

[0096] A method of detecting nitrate, comprising: i) combining a test sample with the bacteria described; and, ii) measuring activity of said reporter gene, wherein expression of said reporter gene correlates with an amount of nitrate in said sample.

[0097] A method of detecting nitrate, comprising: i) combining a test sample with the bacteria herein described; and, ii) measuring expression of said reporter gene, wherein a change in a level of expression of said reporter gene as compared to a control sample lacking nitrate indicates that said test sample contains nitrate.

[0098] A method of detecting nitrate in soil, comprising: i) combining a test sample of soil with the bacteria herein described; ii) measuring expression of said reporter gene; and iii) correlating a measured level of reporter gene expression with a level of nitrate in said test sample of soil using a standard curve.

[0099] A method of detecting excess nitrate levels in a patient, comprising i) administering the bacteria of claim **5-15** to a patient, ii) collecting a stool sample from said patient; iii) measuring expression of said reporter gene in said stool sample, wherein a change in level of expression of said reporter gene over a normal level in a normal patient indicates that said patient has excess nitrate.

[0100] A method of measuring nitrate levels in a patient, comprising:

[0101] a) combining a gut sample with a nitrate reporter bacteria comprising:

[0102] i) a nitrate-sensing sensor kinase (SK) gene encoding an SK protein comprising a ligand binding domain that binds nitrate and activates a kinase domain,

[0103] ii) a cognate RR gene encoding an RR protein comprising a receiver domain operably coupled to an DNA binding domain (DBD), wherein said cognate RR protein is activated by said activated kinase domain phosphorylating said receiver domain, and

[0104] iii) a reporter gene comprising a DNA binding site that binds said DBD of said cognate activated RR protein operably coupled to an open reading frame encoding a reporter protein;

[0105] b) measuring expression of said reporter gene; and,

[0106] c) correlating a measured level of reporter gene expression with a level of nitrate using a standard curve.

[0107] A method as herein described, wherein said bacteria is a probiotic bacteria and said combining step a is by administering said bacteria to said patient.

[0108] A method as herein described, wherein said combining step a is by collecting a stool sample from said patient and combining said stool sample with said bacteria.

[0109] A genetically engineered probiotic bacteria, said probiotic bacteria overexpressing:

[0110] a heterologous nitrate sensor system comprising:

[0111] i) a nitrate-sensing sensor kinase (SK) comprising a ligand binding domain operably coupled to a kinase domain; and,

[0112] ii) a cognate response regulator (RR) comprising a receiver domain operably coupled to an DNA binding domain (DBD);

[0113] a DNA binding site that binds said DBD that is operably coupled to either a reporter gene or a therapeutic protein gene.

[0114] A bacteria as herein described, wherein said RR GENE is rewired such that said receiver domain is operably coupled to a heterologous DBD from another response regulator.

[0115] A treatment method, comprising administering the probiotic bacteria described herein to a patient having excess nitrate, wherein said DBD is operably coupled to a therapeutic protein.

[0116] A fusion protein comprising the amino terminus of NarL operably fused to the DNA binding site domain of YdfI.

[0117] The fusion protein comprising an amino portion of SEQ ID NO 3 fused to a carboxy portion of SEQ ID NO. 5, or comprising SEQ ID NO. 6-9, or homologs of any of same.

[0118] A bacteria comprising an expression vector encoding the fusion protein herein described.

[0119] The use of the word "a" or "an" when used in conjunction with the term "comprising" in the claims or the specification means one or more than one, unless the context dictates otherwise.

[0120] The term "about" means the stated value plus or minus the margin of error of measurement or plus or minus 10% if no method of measurement is indicated.

[0121] The use of the term "or" in the claims is used to mean "and/or" unless explicitly indicated to refer to alternatives only or if the alternatives are mutually exclusive.

[0122] The terms "comprise", "have", "include" and "contain" (and their variants) are open-ended linking verbs and allow the addition of other elements when used in a claim.

[0123] The phrase "consisting of" is closed, and excludes all additional elements.

[0124] The phrase "consisting essentially of" excludes additional material elements, but allows the inclusions of non-material elements that do not substantially change the nature of the invention, such as instructions for use, buffers, background mutations that do not effect the invention, and the like.

[0125] The following abbreviations are used herein:

ABBREVIATION	TERM
ATP	Adenosine triphosphate
CSE	C minimal medium with sodium succinate (6 g/l) and potassium glutamate (8 g/l), a well known *Bacillus* medium
DBD	DNA binding domain
EAEC	enteroaggregative *E. coli*
EHEC	enterohemorrhagic *E. coli*
EIEC	enteroinvasive *E. coli*
EPEC	enteropathogenic *E. coli*
ETEC	enterotoxigenic *E. coli*
GFP	Green fluorescent protein
HK	Histidine kinases
HPT	Histidine phosphotransferase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KD	Kinase domain
LBD	Ligand binding domain
PBS	Phosphate buffered saline
REC	Receiver domain
RR	Response regulator
SK	Sensor kinase
TCS	Two component sensor system including a KD and an RR

BRIEF DESCRIPTION OF FIGURES

[0126] FIG. **1**. Two component sensor systems.

[0127] FIG. **2**. Overview of the human colonic microflora. Bacterial genera were classified as health positive, health negative, or health neutral. Bacterial enumeration was by selective media. Ps=*Pseudomonas*.

[0128] FIG. **3**. Dose response of the described NarX and NarL/YdfI nitrate sensor in *B. subtilis* to $NaNO_3$.

[0129] FIG. **4**. Demonstration of the use of the *B. subtilis* nitrate sensor to detect both nitrate and fertilizer in a soil sample. Dose response curves of soil with increasing concentrations of $NaNO_3$ and fertilizer are shown.

[0130] FIG. **5**. Dose response of the described NarX and NarL/YdfI nitrate sensor in *E. coli* to $NaNO_3$. An additional inactivated NarL-YdfI D59E mutant (lacking the needed aspartate residue for activation) is shown as a negative control to demonstrate specificity of the response to the described pathway.

[0131] FIG. **6**. Preliminary data demonstrating the use of the *E. coli* nitrate sensor being used in vivo to determine the presence of the dextran sodium sulfate (DSS) inflammation disease model in mice.

DETAILED DESCRIPTION

[0132] In our proof of concept work, we used a NarX SK and a rewired NarL RR with a DBD domain from YdfI and a GFP reporter. These sequence are publically available, and are discussed in more detail below.

[0133] NarX: Acts as a sensor kinase (SK) for nitrate/nitrite and transduces signal of nitrate availability to the NarL protein and of both nitrate/nitrite to the NarP protein. NarX probably activates NarL and NarP by phosphorylation in the presence of nitrate. NarX also plays a negative role in controlling NarL activity, probably through dephosphorylation in the absence of nitrate.

narX from *E. coli* [1](SEQ ID NO. 1):
MLKRCLSPLT LVNQVALIVL LSTAIGLAGM AVSGWLVQGV

QGSAHAINKA GSLRMQSYRL LAAVPLSEKD KPLIKEMEQT

AFSAELTRAA ERDGQLAQLQ GLQDYWRNEL IPALMRAQNR

ETVSADVSQF VAGLDQLVSG FDRTTEMRIE TVVLVHRVMA

VFMALLLVFT IIWLRARLLQ PWRQLLAMAS AVSHRDFTQR

ANISGRNEMA MLGTALNNMS AELAESYAVL EQRVQEKTAG

LEHKNQILSF LWQANRRLHS RAPLCERLSP VLNGLQNLTL

LRDIELRVYD TDDEENHQEF TCQPDMTCDD KGCQLCPRGV

LPVGDRGTTL KWRLADSHTQ YGILLATLPQ GRHLSHDQQQ

LVDTLVEQLT ATLALDRHQE RQQQLIVMEE RATIARELHD

SIAQSLSCMK MQVSCLQMQG DALPESSREL LSQIRNELNA

SWAQLRELLT TFRLQLTEPG LRPALEASCE EYSAKFGFPV

KLDYQLPPRL VPSHQAIHLL QIAREALSNA LKHSQASEVV

VTVAQNDNQV KLTVQDNGCG VPENAIRSNH YGMIIMRDRA

QSLRGDCRVR RRESGGTEVV VTFIPEKTFT DVQGDTHE

NarQ: another nitrate detector from *Shewenella oneidensis*. SEQ ID NO. 2:
MKRGSLTSKI LGLMLVLILL SSSLAIFAII NLSYSLGDAK

AINASGSLRM QSYRLMFYAN SGSEAAQEKI TEFENTLHSE

-continued

ALHPSKSWLS PKKIAAQYQL VIDKWLVMKY YIEQENSRDY

AASLKDFVDT IDLLVLEMEH HAAFKLRLLA ASQIFGLGLM

LSIAFLAVRF TKRKVVVPLQ QLMESANTIS KGNFEIEMPE

TEYIELTALT DALQKTAREL ATLYGNLESQ VAEKTLALTR

ANNELAFLYD TLLTLNAKKL DYKALKAALN QLKDYESIDY

LRLIIQYPEQ ELEMIEANGG WPESADNSTR FPLQFEQANL

GYLELISAQD INTPLFKNFA IMLTRSIVIH NATEQRQQLA

LMEERGVIAR ELHDSLGQVL SFLKIQISLL RKNLDHSCRS

PAVEVQLTEI NEGVSTAYVQ LRELLSTFRL TIKEPNLKNA

MEAMLEQLRA NTDIKIHLDY KLSPQWLEAK QHIHILQITR

EATLNAIKHA NASHINIRCY KDDRGMVNIS VSDNGVGIGH

IKERDQHFGI GIMHERASKL DGEVVFSSND THTNSTATTE

QRHQENPDSP LESHNTSNLS QGTIVTLIFP SQQEPTHG

[0134] Other nitrate SK homologs that can be used include WP 042949651 from *Salmonella* (86% amino acid identity to NarX); WP_042949651 from *Citrobacter* (86%); WP_045142747.1 from *Enterobacter* (94%); and WP_059179795.1 from *Lelliottia* (81%). As can be seen, the degree of homology is quite high, indicating a high likelihood of having the same functionality.

[0135] Additional proteins that can substitute herein can be identified by homology search, and functionality can be confirmed as described herein. These are available by BLAST search of the above sequences at GenBank. Additionally, UniProt and other such databases have links to a large number of variants in the same and different species.

[0136] NarL: This response regulator (RR) protein activates the expression of the nitrate reductase (narGHJI) and formate dehydrogenase-N(fdnGHI) operons and represses the transcription of the fumarate reductase (frdABCD) operon in response to a nitrate/nitrite induction signal transmitted by either the NarX or NarQ proteins. The DNA binding element is 173-192 (underlined).

NarL from *E. coli*. (SEQ ID NO. 3):
MSNQEPATIL LIDDHPMLRT GVKQLISMAP DITVVGEASN

GEQGIELAES LDPDLILLDL NMPGMNGLET LDKLREKSLS

GRIVVFSVSN HEEDVVTALK RGADGYLLKD MEPEDLLKAL

HQAAAGEMVL SEALTPVLAA SLRANRATTE RDVNQLTPRE

RDILKLIAQG LPNKMIARRL DITESTVKVH VKHMLKKMKL

KSRVEAAVWV HQERIF

NarP: another *Shewenella frigidmarina* RR believed to respond to NarX and/or NarQ. SEQ ID NO. 4:
MGKPYSVLVV DDHPLLRRGI CQLITSDGDF SLFGETGTGL

EALTAVAEDE PDIILLDLNM KGMSGLDTLN AMRQEGVTAR

IVILTVSDAK QDVVRLLRAG ADGYLLKDTE PDLLLEQLKK

AMLGHRVISD EVEAYLYELK NTIDDNSWIE NLTPRELQIL

-continued

QELAEGKSNR MIAEDLHISE GTVKVHVKNL LRKANAKSRT

EMAVRYLNN

[0137] Additional nitrate RR homologs that can be used herein include WP_000070489.1 from *Shigella* (99%); WP_045443652.1 from *Citrobacter* (98%); WP_061496301.1 from *Enterobacter* (97%); WP_003856701.1 from *Proteobacter* (96%); WP_032641051.1 from *Enterobacter* (96%); WP_001064598.1 from *Salmonella* (96%); WP 020803248.1 from *Kleibsella* (94%); WP 032611305.1 from *Leclercia* (96%); and WP_035895589.1 from *Kluyvera* (95%).

[0138] YdfI: An RR member of the two-component regulatory system YdfH/YdfI. Regulates the transcription of ydfJ by binding to its promoter region. The DNA binding subsequence is aa 166-186 (underlined).

YdfI from *Bacillus subtilis*. (SEQ ID NO. 5):
MNKVLIVDDH LVVREGLKLL IETNDQYTII GEAENGKVAV

RLADELEPDI ILMDLYMPEM SGLEAIKQIK EKHDTPIIIL

TTYNEDHLMI EGIELGAKGY LLKDTSSETL FHTMDAAIRG

NVLLQPDILK RLQEIQFERM KKQRNETQLT EKEVIVLKAI

AKGLKSKAIA FDLGVSERTV KSRLTSIYNK LGANSRTEAV

TIAMQKGILT IDN

Exemplary NarL-YdfI fusion protein (SEQ ID NO. 6), the NarL split at aa 170 (Ydfl underlined):
MSNQEPATIL LIDDHPMLRT GVKQLISMAP DITVVGEASN

GEQGIELAES LDPDLILLDL NMPGMNGLET LDKLREKSLS

GRIVVFSVSN HEEDVVTALK RGADGYLLKD MEPEDLLKAL

HQAAAGEMVL SEALTPVLAA SLRANRATTE RDVNQLTPRE

RDILKLIAQG AKGLKSKAIA FDLGVSERTV KSRLTSIYNK

LGANSRTEAV TIAMQKGILT IDN

[0139] As of yet, there are no examples of this technique succeeding with a DBD from a non-TCS, but it is possible (albeit unlikely) if the domain structure were such as to be activatable by an active REC domain. However, there are a large number (>10,000 TCS) of proteins available from which to choose, so this limitation is very modest. A homologous DBD from the native RR is predicted to give the best chance of success (>30%, >35%, >40%, or higher), but we have used non-homologous domains too.

[0140] Obviously, the heterologous DBD domain that is rewired to the RR should be functional in the bacterial species in which the nitrate sensor will be hosted. In making the change from disparate species, it may be necessary to select a DBD domain from the host species or a closely related species to ensure operability. In this way, we were able to move a heterologous TCS system from a gram negative (*E. coli*) to a gram positive (*B. subtillus*) species.

[0141] The exact fusion point of the two domains can vary somewhat, provided that the DNA binding subsequence (underlined) of NarL (or a homolog) is replaced with that of YdfI or another suitable DBD from a heterologous RR. By

switching the DBD domains, we are able to transport the nitrate sensor system of *E. coli* into the probiotic strain of *B. Subtilus*.

[0142] Other potential DBDs that can be used herein include LiaR (UniProt 032197) at the linker region in the 20 amino acids surrounding the K120 residue and UhpA (P0AGA6) at the linker region in the 20 amino acids surrounding the T123 residue.

[0143] We have also constructed three other chimera RR proteins herein:

NarL131-Ydfl (SEQ. ID NO. 7) which is split at the 131st amino acid of NarL:
MSNQEPATIL LIDDHPMLRT GVKQLISMAP DITVVGEASN

GEQGIELAES LDPDLILLDL NMPGMNGLET LDKLREKSLS

GRIVVFSVSN HEEDVVTALK RGADGYLLKD MEPEDLLKAL

HQAAAGEMVL SPDILKRLQE IQFERMKKQR NETQLTEKEV

IVLKAIAKGL KSKAIAFDLG VSERTVKSRL TSIYNKLGAN

SRTEAVTIAM QKGILTIDN

[0144] The other two chimeric proteins (SEQ. ID NO. 8 and 9) are split at the nearby amino acids NarL142 and NarL154 and have a similar but slightly decreased functionality.

NarL142-Ydfl (SEQ. ID NO. 8):
MSNQEPATIL LIDDHPMLRT GVKQLISMAP DITVVGEASN

GEQGIELAES LDPDLILLDL NMPGMNGLET LDKLREKSLS

GRIVVFSVSN HEEDVVTALK RGADGYLLKD MEPEDLLKAL

HQAAAGEMVL SEALTPVLAA SLQFERMKKQ RNETQLTEKE

VIVLKAIAKG LKSKAIAFDL GVSERTVKSR LTSIYNKLGA

NSRTEAVTIA MQKGILTIDN

NarL154-Ydfl (SEQ. ID NO. 9):
MSNQEPATIL LIDDHPMLRT GVKQLISMAP DITVVGEASN

GEQGIELAES LDPDLILLDL NMPGMNGLET LDKLREKSLS

GRIVVFSVSN HEEDVVTALK RGADGYLLKD MEPEDLLKAL

HQAAAGEMVL SEALTPVLAA SLRANRATTE RDVNQLTEKE

VIVLKAIAKG LKSKAIAFDL GVSERTVKSR LTSIYNKLGA

NSRTEAVTIA MQKGILTIDN

[0145] For proof of concept experiments to characterize the nitrate sensor in *B. subtilis*, the sensor kinase NarX was expressed under the IPTG inducible Phyper_spank promoter in the AmyE locus and the NarL-YdfI gene was expressed from the xylose inducible PxylA promoter at the LacA locus.

[0146] Growth/assay protocol for in vitro *B. subtilis* experiments:

[0147] Overnight pre-culture (13 hours) in CSE 0.5% glycerol

[0148] Dilute overnight culture in CSE 0.5% glycerol with optimal IPTG and Xylose induction levels

[0149] Grow 90-150 minutes

[0150] Dilute to OD600=0.001 in CSE 0.5% glycerol with optimal IPTG and Xylose induction levels

[0151] Grow shaking at 37° C. until cultures reach OD600=0.1 . . . 0.3

[0152] Put on ice, measure fluorescence by flow cytometry

[0153] Growth/assay protocol for measuring nitrate in soil:

[0154] Overnight pre-culture (13 hours) in CSE 0.5% glycerol

[0155] Dilute overnight culture in 1:100 CSE 0.5% glycerol with optimal IPTG and Xylose induction levels

[0156] Grow for until OD600=0.1

[0157] Add appropriate amounts of either sodium nitrate or commercial fertilizer to the soil

[0158] Add 250 μL of culture to 0.1 g of soil and mix

[0159] Grow standing at 37° C. for 2 hours

[0160] Resuspend in 10× volume PBS

[0161] Filter dirt through a Whatman filter

[0162] Put on ice, measure fluorescence by flow cytometry

[0163] For proof of concept experiments to characterize the use of the engineered nitrate sensor in *E. coli*, the sensor kinase NarX was expressed under the constitutive promoter J23114 and translated with the ribosome binding site (RBS) apFAB655 on a p15a plasmid backbone. The engineered NarL-YdfI response regulator was expressed under the constitutive promoter Bba_J23115 and translated with the RBS BCD24 on a ColE1 plasmid backbone. Transcription of the various genes can terminated by the B0015, T1, or T0 terminators.

[0164] Growth/assay protocol for in vitro *E. coli* experiments:

[0165] Overnight pre-culture (~13 hours) in LB+Cm/Spec.

[0166] Dilute to OD_{600}=0.02 in M9+0.4% glycerol.

[0167] Grow 3 hours to OD_{600}=~0.3.

[0168] Dilute to OD_{600}=0.0001 in M9+0.4% glycerol.

[0169] Add nitrate.

[0170] Grow shaking at 37° C. ~6 hours to OD_{600}=~0.3.

[0171] Put on ice, measure OD, measure fluorescence by flow cytometry (FL1=800, FL3=850).

[0172] Growth Assay protocol for detection of nitrate in inflammation mouse models:

[0173] Treat mice DSS for 5 days to simulate inflammatory bowel disease

[0174] Administer genetically engineered nitrate sensing bacteria prepared according the in vitro protocol above

[0175] At 6 hours collect mouse fecal and organ samples

[0176] Process samples via resuspension in PBS and subsequent filtration through a 10 μM filter

[0177] Put on ice and measure fluorescence by flow cytometry (FL1=800, FL3=850).

[0178] The above described vectors, promoters, terminators and other components of the system are exemplary only, and other components could be used. However, the above assays provided proof of concept and confirmed that the above system is indeed a nitrate two-component nitrate sensor system.

[0179] Although four SK/RR gene pairs were exemplified herein, and at least one pair (SEQ ID NO. 1 and 6) was tested

in two host species, there are two features that indicates broad applicability of the invention. The first feature is tunability, which is particularly important for sensing nitrate because the biological ranges for levels of nitrate in humans has not been studied much. Because this system is tunable, once that range is known the sensor can be easily tuned to sense and provide output at the needed levels.

[0180] The second feature piggybacks on the tunability function but also relies on the fact that the inventors have engineered and characterized a suite of DBD, promoters, and reporters for use in this system (described in 62/157,293). When combined, these features allow the inventors to transfer the system to a broad range of microbial species and strains.

[0181] Each of the following is incorporated by reference herein in its entirety for all purposes:

[0182] Claesen J. & Fischbach M. A., Synthetic Microbes As Drug Delivery Systems, ACS Synthetic Biology 2015 4 (4), 358-364.

[0183] Stewart V., Nitrate- and nitrite-responsive sensors NarX and NarQ of proteobacteria, Biochemical Society Transactions February 2003, 31 (1) 1-10;

[0184] DeAngelis, Kristen M., Pingsheng Ji, Mary K. Firestone, and Steven E. Lindow. "Two Novel Bacterial Biosensors for Detection of Nitrate Availability in the Rhizosphere." Applied and Environmental Microbiology 71, no. 12

[0185] 62/157,293, IDENTIFYING LIGANDS FROM BACTERIAL SENSORS, May 5, 2015

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 9

<210> SEQ ID NO 1
<211> LENGTH: 598
<212> TYPE: PRT
<213> ORGANISM: Escherichia coli

<400> SEQUENCE: 1

```
Met Leu Lys Arg Cys Leu Ser Pro Leu Thr Leu Val Asn Gln Val Ala
1               5                  10                  15

Leu Ile Val Leu Leu Ser Thr Ala Ile Gly Leu Ala Gly Met Ala Val
            20                  25                  30

Ser Gly Trp Leu Val Gln Gly Val Gln Gly Ser Ala His Ala Ile Asn
        35                  40                  45

Lys Ala Gly Ser Leu Arg Met Gln Ser Tyr Arg Leu Leu Ala Ala Val
    50                  55                  60

Pro Leu Ser Glu Lys Asp Lys Pro Leu Ile Lys Glu Met Glu Gln Thr
65                  70                  75                  80

Ala Phe Ser Ala Glu Leu Thr Arg Ala Ala Glu Arg Asp Gly Gln Leu
                85                  90                  95

Ala Gln Leu Gln Gly Leu Gln Asp Tyr Trp Arg Asn Glu Leu Ile Pro
            100                 105                 110

Ala Leu Met Arg Ala Gln Asn Arg Glu Thr Val Ser Ala Asp Val Ser
        115                 120                 125

Gln Phe Val Ala Gly Leu Asp Gln Leu Val Ser Gly Phe Asp Arg Thr
    130                 135                 140

Thr Glu Met Arg Ile Glu Thr Val Val Leu Val His Arg Val Met Ala
145                 150                 155                 160

Val Phe Met Ala Leu Leu Leu Val Phe Thr Ile Ile Trp Leu Arg Ala
                165                 170                 175

Arg Leu Leu Gln Pro Trp Arg Gln Leu Leu Ala Met Ala Ser Ala Val
            180                 185                 190

Ser His Arg Asp Phe Thr Gln Arg Ala Asn Ile Ser Gly Arg Asn Glu
        195                 200                 205

Met Ala Met Leu Gly Thr Ala Leu Asn Asn Met Ser Ala Glu Leu Ala
    210                 215                 220

Glu Ser Tyr Ala Val Leu Glu Gln Arg Val Gln Glu Lys Thr Ala Gly
225                 230                 235                 240

Leu Glu His Lys Asn Gln Ile Leu Ser Phe Leu Trp Gln Ala Asn Arg
                245                 250                 255
```

-continued

```
Arg Leu His Ser Arg Ala Pro Leu Cys Glu Arg Leu Ser Pro Val Leu
            260                 265                 270

Asn Gly Leu Gln Asn Leu Thr Leu Leu Arg Asp Ile Glu Leu Arg Val
        275                 280                 285

Tyr Asp Thr Asp Asp Glu Glu Asn His Gln Glu Phe Thr Cys Gln Pro
    290                 295                 300

Asp Met Thr Cys Asp Asp Lys Gly Cys Gln Leu Cys Pro Arg Gly Val
305                 310                 315                 320

Leu Pro Val Gly Asp Arg Gly Thr Thr Leu Lys Trp Arg Leu Ala Asp
                325                 330                 335

Ser His Thr Gln Tyr Gly Ile Leu Leu Ala Thr Leu Pro Gln Gly Arg
            340                 345                 350

His Leu Ser His Asp Gln Gln Gln Leu Val Asp Thr Leu Val Glu Gln
        355                 360                 365

Leu Thr Ala Thr Leu Ala Leu Asp Arg His Gln Glu Arg Gln Gln Gln
    370                 375                 380

Leu Ile Val Met Glu Glu Arg Ala Thr Ile Ala Arg Glu Leu His Asp
385                 390                 395                 400

Ser Ile Ala Gln Ser Leu Ser Cys Met Lys Met Gln Val Ser Cys Leu
                405                 410                 415

Gln Met Gln Gly Asp Ala Leu Pro Glu Ser Ser Arg Glu Leu Leu Ser
            420                 425                 430

Gln Ile Arg Asn Glu Leu Asn Ala Ser Trp Ala Gln Leu Arg Glu Leu
        435                 440                 445

Leu Thr Thr Phe Arg Leu Gln Leu Thr Glu Pro Gly Leu Arg Pro Ala
    450                 455                 460

Leu Glu Ala Ser Cys Glu Glu Tyr Ser Ala Lys Phe Gly Phe Pro Val
465                 470                 475                 480

Lys Leu Asp Tyr Gln Leu Pro Pro Arg Leu Val Pro Ser His Gln Ala
                485                 490                 495

Ile His Leu Leu Gln Ile Ala Arg Glu Ala Leu Ser Asn Ala Leu Lys
            500                 505                 510

His Ser Gln Ala Ser Glu Val Val Val Thr Val Ala Gln Asn Asp Asn
        515                 520                 525

Gln Val Lys Leu Thr Val Gln Asp Asn Gly Cys Gly Val Pro Glu Asn
    530                 535                 540

Ala Ile Arg Ser Asn His Tyr Gly Met Ile Ile Met Arg Asp Arg Ala
545                 550                 555                 560

Gln Ser Leu Arg Gly Asp Cys Arg Val Arg Arg Arg Glu Ser Gly Gly
                565                 570                 575

Thr Glu Val Val Val Thr Phe Ile Pro Glu Lys Thr Phe Thr Asp Val
            580                 585                 590

Gln Gly Asp Thr His Glu
        595


<210> SEQ ID NO 2
<211> LENGTH: 598
<212> TYPE: PRT
<213> ORGANISM: Shewanella oneidensis

<400> SEQUENCE: 2

Met Lys Arg Gly Ser Leu Thr Ser Lys Ile Leu Gly Leu Met Leu Val
```

-continued

```
1               5                   10                  15

Leu Ile Leu Leu Ser Ser Ser Leu Ala Ile Phe Ala Ile Ile Asn Leu
            20                  25                  30

Ser Tyr Ser Leu Gly Asp Ala Lys Ala Ile Asn Ala Ser Gly Ser Leu
        35                  40                  45

Arg Met Gln Ser Tyr Arg Leu Met Phe Tyr Ala Asn Ser Gly Ser Glu
    50                  55                  60

Ala Ala Gln Glu Lys Ile Thr Glu Phe Glu Asn Thr Leu His Ser Glu
65                  70                  75                  80

Ala Leu His Pro Ser Lys Ser Trp Leu Ser Pro Lys Lys Ile Ala Ala
                85                  90                  95

Gln Tyr Gln Leu Val Ile Asp Lys Trp Leu Val Met Lys Tyr Tyr Ile
            100                 105                 110

Glu Gln Glu Asn Ser Arg Asp Tyr Ala Ala Ser Leu Lys Asp Phe Val
        115                 120                 125

Asp Thr Ile Asp Leu Leu Val Leu Glu Met Glu His His Ala Ala Phe
    130                 135                 140

Lys Leu Arg Leu Leu Ala Ala Ser Gln Ile Phe Gly Leu Gly Leu Met
145                 150                 155                 160

Leu Ser Ile Ala Phe Leu Ala Val Arg Phe Thr Lys Arg Lys Val Val
                165                 170                 175

Val Pro Leu Gln Gln Leu Met Glu Ser Ala Asn Thr Ile Ser Lys Gly
            180                 185                 190

Asn Phe Glu Ile Glu Met Pro Glu Thr Glu Tyr Ile Glu Leu Thr Ala
        195                 200                 205

Leu Thr Asp Ala Leu Gln Lys Thr Ala Arg Glu Leu Ala Thr Leu Tyr
    210                 215                 220

Gly Asn Leu Glu Ser Gln Val Ala Glu Lys Thr Leu Ala Leu Thr Arg
225                 230                 235                 240

Ala Asn Asn Glu Leu Ala Phe Leu Tyr Asp Thr Leu Leu Thr Leu Asn
                245                 250                 255

Ala Lys Lys Leu Asp Tyr Lys Ala Leu Lys Ala Ala Leu Asn Gln Leu
            260                 265                 270

Lys Asp Tyr Glu Ser Ile Asp Tyr Leu Arg Leu Ile Ile Gln Tyr Pro
        275                 280                 285

Glu Gln Glu Leu Glu Met Ile Glu Ala Asn Gly Gly Trp Pro Glu Ser
    290                 295                 300

Ala Asp Asn Ser Thr Arg Phe Pro Leu Gln Phe Glu Gln Ala Asn Leu
305                 310                 315                 320

Gly Tyr Leu Glu Leu Ile Ser Ala Gln Asp Ile Asn Thr Pro Leu Phe
                325                 330                 335

Lys Asn Phe Ala Ile Met Leu Thr Arg Ser Ile Val Ile His Asn Ala
            340                 345                 350

Thr Glu Gln Arg Gln Gln Leu Ala Leu Met Glu Glu Arg Gly Val Ile
        355                 360                 365

Ala Arg Glu Leu His Asp Ser Leu Gly Gln Val Leu Ser Phe Leu Lys
    370                 375                 380

Ile Gln Ile Ser Leu Leu Arg Lys Asn Leu Asp His Ser Cys Arg Ser
385                 390                 395                 400

Pro Ala Val Glu Val Gln Leu Thr Glu Ile Asn Glu Gly Val Ser Thr
                405                 410                 415
```

-continued

```
Ala Tyr Val Gln Leu Arg Glu Leu Leu Ser Thr Phe Arg Leu Thr Ile
            420                 425                 430

Lys Glu Pro Asn Leu Lys Asn Ala Met Glu Ala Met Leu Glu Gln Leu
        435                 440                 445

Arg Ala Asn Thr Asp Ile Lys Ile His Leu Asp Tyr Lys Leu Ser Pro
    450                 455                 460

Gln Trp Leu Glu Ala Lys Gln His Ile His Ile Leu Gln Ile Thr Arg
465                 470                 475                 480

Glu Ala Thr Leu Asn Ala Ile Lys His Ala Asn Ala Ser His Ile Asn
                485                 490                 495

Ile Arg Cys Tyr Lys Asp Asp Arg Gly Met Val Asn Ile Ser Val Ser
            500                 505                 510

Asp Asn Gly Val Gly Ile Gly His Ile Lys Glu Arg Asp Gln His Phe
        515                 520                 525

Gly Ile Gly Ile Met His Glu Arg Ala Ser Lys Leu Asp Gly Glu Val
    530                 535                 540

Val Phe Ser Ser Asn Asp Thr His Thr Asn Ser Thr Ala Thr Thr Glu
545                 550                 555                 560

Gln Arg His Gln Glu Asn Pro Asp Ser Pro Leu Glu Ser His Asn Thr
                565                 570                 575

Ser Asn Leu Ser Gln Gly Thr Ile Val Thr Leu Ile Phe Pro Ser Gln
            580                 585                 590

Gln Glu Pro Thr His Gly
        595


<210> SEQ ID NO 3
<211> LENGTH: 216
<212> TYPE: PRT
<213> ORGANISM: Escherichia coli

<400> SEQUENCE: 3

Met Ser Asn Gln Glu Pro Ala Thr Ile Leu Leu Ile Asp Asp His Pro
1               5                   10                  15

Met Leu Arg Thr Gly Val Lys Gln Leu Ile Ser Met Ala Pro Asp Ile
            20                  25                  30

Thr Val Val Gly Glu Ala Ser Asn Gly Glu Gln Gly Ile Glu Leu Ala
        35                  40                  45

Glu Ser Leu Asp Pro Asp Leu Ile Leu Leu Asp Leu Asn Met Pro Gly
    50                  55                  60

Met Asn Gly Leu Glu Thr Leu Asp Lys Leu Arg Glu Lys Ser Leu Ser
65                  70                  75                  80

Gly Arg Ile Val Val Phe Ser Val Ser Asn His Glu Glu Asp Val Val
                85                  90                  95

Thr Ala Leu Lys Arg Gly Ala Asp Gly Tyr Leu Leu Lys Asp Met Glu
            100                 105                 110

Pro Glu Asp Leu Leu Lys Ala Leu His Gln Ala Ala Ala Gly Glu Met
        115                 120                 125

Val Leu Ser Glu Ala Leu Thr Pro Val Leu Ala Ala Ser Leu Arg Ala
    130                 135                 140

Asn Arg Ala Thr Thr Glu Arg Asp Val Asn Gln Leu Thr Pro Arg Glu
145                 150                 155                 160

Arg Asp Ile Leu Lys Leu Ile Ala Gln Gly Leu Pro Asn Lys Met Ile
```

-continued

```
               165                 170                 175

Ala Arg Arg Leu Asp Ile Thr Glu Ser Thr Val Lys Val His Val Lys
            180                 185                 190

His Met Leu Lys Lys Met Lys Leu Lys Ser Arg Val Glu Ala Ala Val
        195                 200                 205

Trp Val His Gln Glu Arg Ile Phe
    210                 215


<210> SEQ ID NO 4
<211> LENGTH: 209
<212> TYPE: PRT
<213> ORGANISM: Shewanella frigidimarina

<400> SEQUENCE: 4

Met Gly Lys Pro Tyr Ser Val Leu Val Val Asp Asp His Pro Leu Leu
1               5                   10                  15

Arg Arg Gly Ile Cys Gln Leu Ile Thr Ser Asp Gly Asp Phe Ser Leu
            20                  25                  30

Phe Gly Glu Thr Gly Thr Gly Leu Glu Ala Leu Thr Ala Val Ala Glu
        35                  40                  45

Asp Glu Pro Asp Ile Ile Leu Leu Asp Leu Asn Met Lys Gly Met Ser
    50                  55                  60

Gly Leu Asp Thr Leu Asn Ala Met Arg Gln Glu Gly Val Thr Ala Arg
65                  70                  75                  80

Ile Val Ile Leu Thr Val Ser Asp Ala Lys Gln Asp Val Val Arg Leu
                85                  90                  95

Leu Arg Ala Gly Ala Asp Gly Tyr Leu Leu Lys Asp Thr Glu Pro Asp
            100                 105                 110

Leu Leu Leu Glu Gln Leu Lys Lys Ala Met Leu Gly His Arg Val Ile
        115                 120                 125

Ser Asp Glu Val Glu Ala Tyr Leu Tyr Glu Leu Lys Asn Thr Ile Asp
    130                 135                 140

Asp Asn Ser Trp Ile Glu Asn Leu Thr Pro Arg Glu Leu Gln Ile Leu
145                 150                 155                 160

Gln Glu Leu Ala Glu Gly Lys Ser Asn Arg Met Ile Ala Glu Asp Leu
                165                 170                 175

His Ile Ser Glu Gly Thr Val Lys Val His Val Lys Asn Leu Leu Arg
            180                 185                 190

Lys Ala Asn Ala Lys Ser Arg Thr Glu Met Ala Val Arg Tyr Leu Asn
        195                 200                 205

Asn


<210> SEQ ID NO 5
<211> LENGTH: 213
<212> TYPE: PRT
<213> ORGANISM: Bacillus subtilis

<400> SEQUENCE: 5

Met Asn Lys Val Leu Ile Val Asp Asp His Leu Val Val Arg Glu Gly
1               5                   10                  15

Leu Lys Leu Leu Ile Glu Thr Asn Asp Gln Tyr Thr Ile Ile Gly Glu
            20                  25                  30

Ala Glu Asn Gly Lys Val Ala Val Arg Leu Ala Asp Glu Leu Glu Pro
        35                  40                  45
```

-continued

```
Asp Ile Ile Leu Met Asp Leu Tyr Met Pro Glu Met Ser Gly Leu Glu
    50                  55                  60

Ala Ile Lys Gln Ile Lys Glu Lys His Asp Thr Pro Ile Ile Ile Leu
65                  70                  75                  80

Thr Thr Tyr Asn Glu Asp His Leu Met Ile Glu Gly Ile Glu Leu Gly
                85                  90                  95

Ala Lys Gly Tyr Leu Leu Lys Asp Thr Ser Ser Glu Thr Leu Phe His
            100                 105                 110

Thr Met Asp Ala Ala Ile Arg Gly Asn Val Leu Leu Gln Pro Asp Ile
        115                 120                 125

Leu Lys Arg Leu Gln Glu Ile Gln Phe Glu Arg Met Lys Lys Gln Arg
    130                 135                 140

Asn Glu Thr Gln Leu Thr Glu Lys Glu Val Ile Val Leu Lys Ala Ile
145                 150                 155                 160

Ala Lys Gly Leu Lys Ser Lys Ala Ile Ala Phe Asp Leu Gly Val Ser
                165                 170                 175

Glu Arg Thr Val Lys Ser Arg Leu Thr Ser Ile Tyr Asn Lys Leu Gly
            180                 185                 190

Ala Asn Ser Arg Thr Glu Ala Val Thr Ile Ala Met Gln Lys Gly Ile
        195                 200                 205

Leu Thr Ile Asp Asn
    210


<210> SEQ ID NO 6
<211> LENGTH: 223
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: NarL+YdfI

<400> SEQUENCE: 6

Met Ser Asn Gln Glu Pro Ala Thr Ile Leu Leu Ile Asp Asp His Pro
1               5                   10                  15

Met Leu Arg Thr Gly Val Lys Gln Leu Ile Ser Met Ala Pro Asp Ile
            20                  25                  30

Thr Val Val Gly Glu Ala Ser Asn Gly Glu Gln Gly Ile Glu Leu Ala
        35                  40                  45

Glu Ser Leu Asp Pro Asp Leu Ile Leu Leu Asp Leu Asn Met Pro Gly
    50                  55                  60

Met Asn Gly Leu Glu Thr Leu Asp Lys Leu Arg Glu Lys Ser Leu Ser
65                  70                  75                  80

Gly Arg Ile Val Val Phe Ser Val Ser Asn His Glu Glu Asp Val Val
                85                  90                  95

Thr Ala Leu Lys Arg Gly Ala Asp Gly Tyr Leu Leu Lys Asp Met Glu
            100                 105                 110

Pro Glu Asp Leu Leu Lys Ala Leu His Gln Ala Ala Ala Gly Glu Met
        115                 120                 125

Val Leu Ser Glu Ala Leu Thr Pro Val Leu Ala Ala Ser Leu Arg Ala
    130                 135                 140

Asn Arg Ala Thr Thr Glu Arg Asp Val Asn Gln Leu Thr Pro Arg Glu
145                 150                 155                 160

Arg Asp Ile Leu Lys Leu Ile Ala Gln Gly Ala Lys Gly Leu Lys Ser
                165                 170                 175
```

-continued

```
Lys Ala Ile Ala Phe Asp Leu Gly Val Ser Glu Arg Thr Val Lys Ser
            180                 185                 190

Arg Leu Thr Ser Ile Tyr Asn Lys Leu Gly Ala Asn Ser Arg Thr Glu
        195                 200                 205

Ala Val Thr Ile Ala Met Gln Lys Gly Ile Leu Thr Ile Asp Asn
    210                 215                 220


<210> SEQ ID NO 7
<211> LENGTH: 219
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: NarL split at 131 plus the DNA binding
      sequence from YdfI

<400> SEQUENCE: 7

Met Ser Asn Gln Glu Pro Ala Thr Ile Leu Leu Ile Asp Asp His Pro
1               5                   10                  15

Met Leu Arg Thr Gly Val Lys Gln Leu Ile Ser Met Ala Pro Asp Ile
            20                  25                  30

Thr Val Val Gly Glu Ala Ser Asn Gly Glu Gln Gly Ile Glu Leu Ala
        35                  40                  45

Glu Ser Leu Asp Pro Asp Leu Ile Leu Leu Asp Leu Asn Met Pro Gly
    50                  55                  60

Met Asn Gly Leu Glu Thr Leu Asp Lys Leu Arg Glu Lys Ser Leu Ser
65                  70                  75                  80

Gly Arg Ile Val Val Phe Ser Val Ser Asn His Glu Glu Asp Val Val
                85                  90                  95

Thr Ala Leu Lys Arg Gly Ala Asp Gly Tyr Leu Leu Lys Asp Met Glu
            100                 105                 110

Pro Glu Asp Leu Leu Lys Ala Leu His Gln Ala Ala Ala Gly Glu Met
        115                 120                 125

Val Leu Ser Pro Asp Ile Leu Lys Arg Leu Gln Glu Ile Gln Phe Glu
    130                 135                 140

Arg Met Lys Lys Gln Arg Asn Glu Thr Gln Leu Thr Glu Lys Glu Val
145                 150                 155                 160

Ile Val Leu Lys Ala Ile Ala Lys Gly Leu Lys Ser Lys Ala Ile Ala
                165                 170                 175

Phe Asp Leu Gly Val Ser Glu Arg Thr Val Lys Ser Arg Leu Thr Ser
            180                 185                 190

Ile Tyr Asn Lys Leu Gly Ala Asn Ser Arg Thr Glu Ala Val Thr Ile
        195                 200                 205

Ala Met Gln Lys Gly Ile Leu Thr Ile Asp Asn
    210                 215


<210> SEQ ID NO 8
<211> LENGTH: 220
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: NarL split at a.a. 142 plus the DNA binding
      sequence from YdfI

<400> SEQUENCE: 8

Met Ser Asn Gln Glu Pro Ala Thr Ile Leu Leu Ile Asp Asp His Pro
1               5                   10                  15
```

-continued

```
Met Leu Arg Thr Gly Val Lys Gln Leu Ile Ser Met Ala Pro Asp Ile
            20                  25                  30

Thr Val Val Gly Glu Ala Ser Asn Gly Glu Gln Gly Ile Glu Leu Ala
        35                  40                  45

Glu Ser Leu Asp Pro Asp Leu Ile Leu Leu Asp Leu Asn Met Pro Gly
    50                  55                  60

Met Asn Gly Leu Glu Thr Leu Asp Lys Leu Arg Glu Lys Ser Leu Ser
65                  70                  75                  80

Gly Arg Ile Val Val Phe Ser Val Ser Asn His Glu Glu Asp Val Val
                85                  90                  95

Thr Ala Leu Lys Arg Gly Ala Asp Gly Tyr Leu Leu Lys Asp Met Glu
            100                 105                 110

Pro Glu Asp Leu Leu Lys Ala Leu His Gln Ala Ala Ala Gly Glu Met
        115                 120                 125

Val Leu Ser Glu Ala Leu Thr Pro Val Leu Ala Ala Ser Leu Gln Phe
    130                 135                 140

Glu Arg Met Lys Lys Gln Arg Asn Glu Thr Gln Leu Thr Glu Lys Glu
145                 150                 155                 160

Val Ile Val Leu Lys Ala Ile Ala Lys Gly Leu Lys Ser Lys Ala Ile
                165                 170                 175

Ala Phe Asp Leu Gly Val Ser Glu Arg Thr Val Lys Ser Arg Leu Thr
            180                 185                 190

Ser Ile Tyr Asn Lys Leu Gly Ala Asn Ser Arg Thr Glu Ala Val Thr
        195                 200                 205

Ile Ala Met Gln Lys Gly Ile Leu Thr Ile Asp Asn
    210                 215                 220


<210> SEQ ID NO 9
<211> LENGTH: 220
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: NarL split at a.a. 154 plust the DNA binding
      sequence from YdfI

<400> SEQUENCE: 9

Met Ser Asn Gln Glu Pro Ala Thr Ile Leu Leu Ile Asp Asp His Pro
1               5                   10                  15

Met Leu Arg Thr Gly Val Lys Gln Leu Ile Ser Met Ala Pro Asp Ile
            20                  25                  30

Thr Val Val Gly Glu Ala Ser Asn Gly Glu Gln Gly Ile Glu Leu Ala
        35                  40                  45

Glu Ser Leu Asp Pro Asp Leu Ile Leu Leu Asp Leu Asn Met Pro Gly
    50                  55                  60

Met Asn Gly Leu Glu Thr Leu Asp Lys Leu Arg Glu Lys Ser Leu Ser
65                  70                  75                  80

Gly Arg Ile Val Val Phe Ser Val Ser Asn His Glu Glu Asp Val Val
                85                  90                  95

Thr Ala Leu Lys Arg Gly Ala Asp Gly Tyr Leu Leu Lys Asp Met Glu
            100                 105                 110

Pro Glu Asp Leu Leu Lys Ala Leu His Gln Ala Ala Ala Gly Glu Met
        115                 120                 125

Val Leu Ser Glu Ala Leu Thr Pro Val Leu Ala Ala Ser Leu Arg Ala
```

-continued

```
    130                 135                 140

Asn Arg Ala Thr Thr Glu Arg Asp Val Asn Gln Leu Thr Glu Lys Glu
145                 150                 155                 160

Val Ile Val Leu Lys Ala Ile Ala Lys Gly Leu Lys Ser Lys Ala Ile
                165                 170                 175

Ala Phe Asp Leu Gly Val Ser Glu Arg Thr Val Lys Ser Arg Leu Thr
            180                 185                 190

Ser Ile Tyr Asn Lys Leu Gly Ala Asn Ser Arg Thr Glu Ala Val Thr
        195                 200                 205

Ile Ala Met Gln Lys Gly Ile Leu Thr Ile Asp Asn
    210                 215                 220
```

1) A genetically engineered bacteria, said bacteria over-expressing:

a) a two component sensor system (TCS) comprising:

i) a nitrate-sensing sensor kinase (SK) gene comprising a ligand binding domain operably coupled to a kinase domain, and,

ii) a cognate response regulator (RR) gene comprising a receiver domain operably coupled to an DNA binding domain (DBD),

b) an output promoter that binds said DBD that is operably coupled to a heterologous reporter gene.

2) (canceled)

3) (canceled)

4) (canceled)

5) The bacteria of claim **1**, which is gut-adapted for use in humans.

6) The bacteria of claim **5**, wherein said SK gene or said RR gene or both genes are encoded on an expression vector, an inducible expression vector, and/or a constitutive expression vector.

7) (canceled)

8) (canceled)

9) The bacteria of claim **1**, wherein said SK gene or said RR gene or both genes integrated into a genome of said bacteria.

10) The bacteria of claim **1**, wherein said SK gene and said RR gene are encoded in a single operon.

11) The bacteria of claim **1**, wherein said reporter gene is encoded on a plasmid.

12) The bacteria of claim **1**, wherein said reporter gene is integrated into a genome of said bacteria.

13) The bacteria of claim **1**, comprising SEQ ID NO. 1 and an amino terminal portion of SEQ ID NO. 3 operably fused to a carboxy terminal portion of SEQ ID NO 5 containing a DNA binding site.

14) The bacteria of claim **1**, wherein said reporter gene encodes a fluorescent protein.

15) The bacteria of claim **1**, wherein said reporter gene encodes green fluorescent protein, red fluorescent protein, far red fluorescent protein, blue fluorescent protein, orange fluorescent protein, yellow fluorescent protein, mCHERRY, mORANGE, mCITRINE, VENUS, YPET, EMERALD, or CERULEAN.

16) (canceled)

17) (canceled)

18) (canceled)

19) (canceled)

20) (canceled)

21) (canceled)

22) A method of measuring nitrate levels in a patient, comprising:

a) combining a gut sample with a nitrate reporter bacteria comprising:

i) a nitrate-sensing sensor kinase (SK) gene encoding an SK protein comprising a ligand binding domain that binds nitrate and activates a kinase domain,

ii) a cognate RR gene encoding an RR protein comprising a receiver domain operably coupled to an DNA binding domain (DBD), wherein said cognate RR protein is activated by said activated kinase domain phosphorylating said receiver domain, and

iii) a reporter gene comprising a DNA binding site that binds said DBD of said cognate activated RR protein operably coupled to an open reading frame encoding a reporter protein;

b) measuring expression of said reporter gene; and,

c) correlating a measured level of reporter gene expression with a level of nitrate using a standard curve.

23) The method of claim **22**, wherein said bacteria is a gut-adapted bacteria and said combining step is by administering said bacteria to said patient.

24) The method of claim **22**, wherein said combining step a) is by collecting a gut or stool sample from said patient and combining said gut or stool sample with said bacteria.

25) (canceled)

26) (canceled)

27) (canceled)

28) A fusion protein comprising the amino terminus of NarL operably fused to the DNA binding site domain of YdfI.

29) The fusion protein of claim **28**, comprising an amino portion of SEQ ID NO 3 fused to a carboxy portion of SEQ ID NO. 5.

30) The fusion protein of claim **28**, comprising SEQ ID NO 6-9.

31) A bacteria comprising an expression vector encoding the fusion protein of claim **28**.

* * * * *