



(12) **DEMANDE DE BREVET CANADIEN**  
**CANADIAN PATENT APPLICATION**  
(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2018/02/23  
(87) Date publication PCT/PCT Publication Date: 2018/08/30  
(85) Entrée phase nationale/National Entry: 2019/06/25  
(86) N° demande PCT/PCT Application No.: US 2018/019577  
(87) N° publication PCT/PCT Publication No.: 2018/156998  
(30) Priorités/Priorities: 2017/02/24 (US62/463,451);  
2017/11/09 (US62/583,990)

(51) Cl.Int./Int.Cl. *A01N 63/00* (2006.01),  
*C12N 15/113* (2010.01), *C12N 15/63* (2006.01)  
(71) Demandeur/Applicant:  
FLAGSHIP PIONEERING INNOVATIONS V, INC., US  
(72) Inventeurs/Inventors:  
MARTINEZ, IGNACIO, US;  
ARMEN, ZACHARY GARO, US;  
MARTIN, BARRY ANDREW, US;  
AVENDANO AMADO, MAIER STEVE, US  
(74) Agent: GOWLING WLG (CANADA) LLP

(54) Titre : COMPOSITIONS ET PROCEDES ASSOCIES DESTINES A LA MODULATION D'ENDOSYMBIONTES  
(54) Title: COMPOSITIONS AND RELATED METHODS FOR MODULATING ENDOSYMBIONTS

(57) **Abrégé/Abstract:**

Provided herein are methods and compositions for modulating the fitness of a host invertebrate (e.g., insect, mollusk, or nematode) by altering interactions between the host and one or more microorganisms resident in the host. The invention features a composition including a modulating agent (e.g., a polypeptide, nucleic acid, small molecule, or combinations thereof) that can induce changes in the host's microbiota in a manner that modulates (e.g., increases or decreases) host fitness. The modulating agent described herein may modulate the fitness of a variety of invertebrates that are important for agriculture, commerce, and/or public health.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property  
Organization  
International Bureau

(43) International Publication Date  
30 August 2018 (30.08.2018)



(10) International Publication Number  
WO 2018/156998 A1

(51) International Patent Classification:

A01N 63/00 (2006.01) C12N 15/113 (2010.01)  
C12N 15/63 (2006.01)

(21) International Application Number:

PCT/US2018/019577

(22) International Filing Date:

23 February 2018 (23.02.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/463,451 24 February 2017 (24.02.2017) US  
62/583,990 09 November 2017 (09.11.2017) US

(71) Applicant: FLAGSHIP PIONEERING, INC. [US/US];  
55 Cambridge Parkway, 8th Floor 800E, Cambridge, MA  
02142 (US).

(72) Inventors: MARTINEZ, Ignacio; 36 Moon Hill Rd,  
Lexington, MA 02421 (US). ARMEN, Zachary, Garo;  
1200 Washington St., Unit 412, Boston, MA 02118 (US).  
MARTIN, Barry, Andrew; 150 Mount Vernon St, Boston,  
MA 02108 (US). AVENDANO AMADO, Maier, Steve;  
10 Glassworks Ave, Apt 411, Cambridge, MA 02141 (US).

(74) Agent: ELBING, Karen, L. et al.; Clark & Elbing LLP,  
101 Federal Street, 15th Floor, Boston, MA 02110 (US).

(81) Designated States (unless otherwise indicated, for every  
kind of national protection available): AE, AG, AL, AM,  
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,  
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,  
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,  
HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP,  
KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,  
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,  
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,  
SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,  
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every  
kind of regional protection available): ARIPO (BW, GH,  
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ,  
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,  
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,  
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,  
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,  
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,  
KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a  
patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the  
earlier application (Rule 4.17(iii))

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: COMPOSITIONS AND RELATED METHODS FOR MODULATING ENDOSYMBIONTS

(57) Abstract: Provided herein are methods and compositions for modulating the fitness of a host invertebrate (e.g., insect, mollusk, or nematode) by altering interactions between the host and one or more microorganisms resident in the host. The invention features a composition including a modulating agent (e.g., a polypeptide, nucleic acid, small molecule, or combinations thereof) that can induce changes in the host's microbiota in a manner that modulates (e.g., increases or decreases) host fitness. The modulating agent described herein may modulate the fitness of a variety of invertebrates that are important for agriculture, commerce, and/or public health.

WO 2018/156998 A1

## COMPOSITIONS AND RELATED METHODS FOR MODULATING ENDOSYMBIONTS

### CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional Application No. 62/463,451, filed on February 24, 2017, and U.S. Provisional Application No. 62/583,990, filed on November 9, 2017, the contents of which are hereby incorporated herein by reference in their entireties.

### BACKGROUND

Invertebrate organisms (e.g., insects, mollusks, or nematodes) are pervasive in the human environment. In some instances, invertebrates serve beneficial roles, such as nematodes or arthropods utilized in agriculture for pollination efforts and pest control or in commerce for the production of commercial products (e.g., honey or silk). In other instances, invertebrates can be detrimental, including some species of mollusks (e.g., snails and slugs), nematodes, or insects that can be serious crop pests or carriers of disease. Thus, there is need in the art for methods and compositions to modulate the fitness of invertebrates that are important in agriculture, commerce, or public health.

### SUMMARY OF THE INVENTION

Disclosed herein are compositions and methods for modulating the fitness of invertebrates, including insects, nematodes, or mollusks, by altering the interactions between the host and one or more microorganisms resident in the host.

In one aspect, provided herein is a method for decreasing the fitness of an insect, the method including delivering to the insect an effective amount of a polynucleotide that includes a dsRNA that decreases expression of an insect bacteriocyte regulatory gene or an insect immunoregulatory gene in the insect relative to an insect that has not been administered the dsRNA.

In some embodiments, the gene encodes a protein from the bacteriocyte-specific cysteine rich proteins BCR family, a protein from the secreted proteins SP family, BicD (Protein bicaudal D), Cact (cactus), DIF (Dorsal related immunity factor), Toll (Toll Interacting Protein), or imd (immune deficiency protein). In some embodiments, the gene encodes a protein having at least 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100% amino acid sequence identity to a protein listed in Table 5, Table 8, or Table 9. In some embodiments, the gene encodes a functional homolog of a protein listed in Table 5, Table 8, or Table 9. For example, the gene may encode a cactus-like protein in aphids (e.g., any one of the proteins described by GenBank Accession Nos: XP\_022175228.1, XP\_016656687.1, NP\_001156668.1, XP\_008179071.1, or XP\_016656686.1, the associated amino acid and nucleotide sequences of which are incorporated by reference).

In some embodiments, the dsRNA is complementary to 10 to 30 nucleotides of the gene in the insect (e.g., 10 to 30 nucleotides, 12 to 28 nucleotides, 14 to 26 nucleotides, 16 to 24 nucleotides, 14 to 22 nucleotides, or 18 to 20 nucleotides). In some embodiments, the dsRNA is complementary to 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides of the gene in the insect. In some embodiments, the entire length of the dsRNA is complementary to the gene. In alternative embodiments, only a portion of the dsRNA is complementary to the gene.

In some embodiments, the method is effective to decrease expression of the gene in the insect, e.g., by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% or greater relative to an insect that has not been administered the polynucleotide. In some embodiments, the method is effective to decrease expression of the gene in the insect, e.g., by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% or greater as compared to a reference level (e.g., as compared to expression of one or more control genes (e.g., a housekeeping gene), expression of the same gene in a different sample (e.g., one or more control samples), or expression of the same gene in the same sample at one or more earlier time points).

In some embodiments, the method is effective to decrease expression of the gene in the insect, e.g., by about 1.5x, 1.75x, 2x, 3x, 4x, 5x, 6x, 7x, 8x, 9x, 10x, 25x, 50x, 75x, or 100x fold less relative to an insect that has not been administered the polynucleotide. In some embodiments, the method is effective to decrease expression of the gene in the insect, e.g., by about 1.5x, 1.75x, 2x, 3x, 4x, 5x, 6x, 7x, 8x, 9x, 10x, 25x, 50x, 75x, or 100x fold less as compared to a reference level (e.g., as compared to expression of one or more control genes (e.g., a housekeeping gene), expression of the same gene in a different sample (e.g., one or more control samples), or expression of the same gene in the same sample at one or more earlier time points).

In some embodiments, the method is effective to inhibit expression of the gene in the insect or to decrease expression of the gene to an undetectable level.

In some embodiments, the method is effective to decrease the level, diversity, or metabolism of one or more microorganisms resident in the insect relative to an insect that has not been delivered the polynucleotide. In some embodiments, the method is effective to decrease the level, diversity, or metabolism of one or more microorganisms resident in the insect, e.g., by about 1.5x, 1.75x, 2x, 3x, 4x, 5x, 6x, 7x, 8x, 9x, 10x, 25x, 50x, 75x, or 100x fold less relative to an insect that has not been delivered the polynucleotide. In some embodiments, the method is effective to decrease the level, diversity, or metabolism of one or more microorganisms resident in the insect, e.g., by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% less relative to an insect that has not been delivered the polynucleotide. In certain embodiments, the one or more microorganisms is a *Buchnera* spp.

In some embodiments, the method is effective to decrease the fitness of the insect relative to an insect that has not been delivered the polynucleotide. In some embodiments, the method is effective to decrease the fitness of the insect, e.g., by about 1.5x, 1.75x, 2x, 3x, 4x, 5x, 6x, 7x, 8x, 9x, 10x, 25x, 50x, 75x, or 100x fold less relative to an insect that has not been delivered the polynucleotide. In some embodiments, the method is effective to decrease the fitness of the insect, e.g., by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% less relative to an insect that has not been delivered the polynucleotide.

In some embodiments, the polynucleotide is delivered in a composition formulated for delivery to insects. In some embodiments, the delivery includes delivering the polynucleotide to at least one habitat where the insect pest grows, lives, reproduces, feeds, or infests. In some embodiments, the delivery comprises spraying the antimicrobial peptide on an agricultural crop. In some embodiments, the polynucleotide is delivered as an insect comestible composition for ingestion by the insect.

In some embodiments, the polynucleotide is formulated with an agriculturally acceptable carrier as a liquid, a solid, an aerosol, a paste, a gel, or a gas composition.

In some embodiments, the insect is an aphid.

In another aspect, provided herein is a composition including a polynucleotide that includes a  
 5 dsRNA formulated for delivery to an insect, wherein the dsRNA is complementary to 15 to 30 nucleotides of an insect bacteriocyte regulatory gene or an insect immunoregulatory gene. In some embodiments, the gene encodes a protein selected from the group consisting of a protein from the bacteriocyte-specific cysteine rich proteins BCR family, a protein from the secreted proteins SP family, BicD (Protein bicaudal D), Cact (cactus), DIF (Dorsal related immunity factor), Toll (Toll Interacting Protein), and imd (immune  
 10 deficiency protein). In some embodiments, the gene encodes a protein having at least 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100% amino acid sequence identity to a protein listed in Table 5, Table 8, or Table 9. In some embodiments, the gene encodes a functional homolog of a protein listed in Table 5, Table 8, or Table 9. For example, the gene may encode a cactus-like protein in aphids (e.g., any one of the proteins described by GenBank Accession Nos:  
 15 XP\_022175228.1, XP\_016656687.1, NP\_001156668.1, XP\_008179071.1, or XP\_016656686.1, the associated amino acid and nucleotide sequences of which are incorporated by reference). In some embodiments, the dsRNA is complementary to 10 to 30 nucleotides of the gene in the insect (e.g., 10 to 30 nucleotides, 12 to 28 nucleotides, 14 to 26 nucleotides, 16 to 24 nucleotides, 14 to 22 nucleotides, or 18 to 20 nucleotides). In some embodiments, the dsRNA is complementary to 10, 11, 12, 13, 14, 15, 16,  
 20 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides of the gene in the insect. In some embodiments, the entire length of the dsRNA is complementary to the gene. In alternative embodiments, only a portion of the dsRNA is complementary to the gene.

In a further aspect, provided herein is a plant comprising a topical application of the compositions described herein.

In yet another aspect, provided herein is a transgenic plant cell having in its genome a  
 25 recombinant DNA construct, wherein the recombinant DNA construct includes a heterologous promoter operably linked to a DNA encoding a RNA including at least one double-stranded RNA region, at least one strand of which includes a nucleotide sequence that is complementary to 15 to 30 nucleotides of an insect bacteriocyte regulatory gene or an insect immunoregulatory gene. In some embodiments, the  
 30 gene encodes a protein selected from the group consisting of a protein from the bacteriocyte-specific cysteine rich proteins BCR family, a protein from the secreted proteins SP family, BicD (Protein bicaudal D), Cact (cactus), DIF (Dorsal related immunity factor), Toll (Toll Interacting Protein), and imd (immune deficiency protein). In some embodiments, the gene encodes a protein having at least 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100% amino acid sequence identity to a protein listed in Table 5, Table 8, or Table 9. In some embodiments, the gene encodes a functional  
 35 homolog of a protein listed in Table 5, Table 8, or Table 9. For example, the gene may encode a cactus-like protein in aphids (e.g., any one of the proteins described by GenBank Accession Nos:  
 XP\_022175228.1, XP\_016656687.1, NP\_001156668.1, XP\_008179071.1, or XP\_016656686.1, the associated amino acid and nucleotide sequences of which are incorporated by reference). In some  
 40 embodiments, the dsRNA is complementary to 10 to 30 nucleotides of the gene in the insect (e.g., 10 to

30 nucleotides, 12 to 28 nucleotides, 14 to 26 nucleotides, 16 to 24 nucleotides, 14 to 22 nucleotides, or 18 to 20 nucleotides). In some embodiments, the dsRNA is complementary to 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides of the gene in the insect. In some embodiments, the entire length of the dsRNA is complementary to the gene. In alternative embodiments, only a portion of the dsRNA is complementary to the gene.

In yet another aspect, provided herein are compositions that include a modulating agent (e.g., a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi, shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or a combination thereof) that modulates (e.g., increases or decreases) the fitness of an invertebrate host (e.g., insect, mollusk, or nematode), wherein the modulating agent alters interactions between the host and one or more microorganisms resident in the host.

In some embodiments, the modulating agent (e.g., a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi, shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or a combination thereof) targets one or more host pathways that mediate interactions between the host and the one or more microorganisms resident in the host (e.g., host-microbiota interactions). In certain embodiments, the targeting (e.g., upregulation, downregulation, or inhibition) of the one or more host pathways alters the level, diversity, or function of the one or more microorganisms resident in the host in comparison to a host organism to which the modulating agent has not been administered. In certain embodiments, the targeting (e.g., upregulation, downregulation, or inhibition) of the one or more host pathways increases the level, diversity, or function of the one or more microorganisms resident in the host in comparison to a host organism to which the modulating agent has not been administered. In alternative embodiments, the targeting (e.g., upregulation, downregulation, or inhibition) of the one or more host pathways decreases the level, diversity, or function of the one or more microorganisms resident in the host in comparison to a host organism to which the modulating agent has not been administered.

In some embodiments, the host pathway is a pathway that regulates bacteriocyte function or development. In some embodiments, the targeting of bacteriocyte function or development may increase and/or decrease the level, diversity, and/or function of one or more microorganisms resident in the bacteriocyte in comparison to a host organism to which the modulating agent has not been administered. In certain embodiments, the targeting of bacteriocyte function or development decreases the level, diversity, or function of one or more microorganisms resident in the bacteriocyte (e.g., a bacteriocyte of an aphid) in comparison to a host organism to which the modulating agent has not been administered. In certain embodiments, the targeting of bacteriocyte function or development increases the level, diversity, or function of one or more microorganisms resident in the bacteriocyte (e.g., a bacteriocyte of an aphid) in comparison to a host organism to which the modulating agent has not been administered.

In some embodiments, the host pathway is a pathway that regulates the host's immune system. For example, in some embodiments, the modulating agent activates an immune response against the one or more microorganisms resident in the host, thereby decreasing the level, diversity, and/or function of the

one or more microorganisms in comparison to a host organism to which the modulating agent has not been administered. Alternatively, in some embodiments, the modulating agent suppresses an immune response against the one or more microorganisms resident in the host, thereby increasing the level, diversity, and/or function of the one or more microorganisms in comparison to a host organism to which the modulating agent has not been administered.

In some embodiments, the modulating agent (e.g., a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi, shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or a combination thereof) targets one or more host pathways by altering gene expression in the host in comparison to a host organism to which the modulating agent has not been administered. For example, the modulating agent may increase and/or decrease gene expression in the host in comparison to a host organism to which the modulating agent has not been administered. In some embodiments, the modulating agent alters expression of a gene that encodes a protein listed in Table 3, Table 4, Table 5, Table 7, Table 8, or Table 9 in comparison to a host organism to which the modulating agent has not been administered. In some embodiments, the modulating agent decreases expression of a gene that encodes a protein listed in Table 3, Table 4, Table 5, Table 7, Table 8, or Table 9 in comparison to a host organism to which the modulating agent has not been administered. In some embodiments, the modulating agent decreases expression of a gene that encodes a protein listed in Table 3, Table 4, Table 5, Table 7, Table 8, or Table 9 in comparison to a host organism to which the modulating agent has not been administered. In some embodiments, the gene encodes a bacteriocyte regulatory peptide. For example, the bacteriocyte regulatory peptide may be one listed in Table 5 or Table 8 (e.g., BCR1). In some embodiments, the gene encodes an immune system component. For example, the immune system component may be one listed in Table 9. In some embodiments, the modulating agent targets a polypeptide in the host. In some embodiments, the polypeptide is an enzyme or cell receptor. In some embodiments, the modulating agent increases and/or decreases enzyme activity in comparison to a host organism to which the modulating agent has not been administered. In some embodiments, the modulating agent increases and/or decreases cell receptor signaling in comparison to a host organism to which the modulating agent has not been administered. In some embodiments, the host protein is one listed in Table 4, Table 5, Table 8, or Table 9.

The modulating agent (e.g., a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi, shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or a combination thereof) may additionally or alternatively target one or more microbial pathways that mediate interactions between the host and one or more microorganisms resident in the host. In some embodiments, the modulating agent alters gene expression in one or more microorganisms resident in the host in comparison to a host organism to which the modulating agent has not been administered. For example, the modulating agent may increase and/or decrease gene expression in the one or more microorganisms in comparison to a host organism to which the modulating agent has not been administered. In some embodiments, the modulating agent alters (e.g., increases or decreases) the expression of a gene that encodes a protein listed in Table 3 or Table 7 in comparison to a host organism

to which the modulating agent has not been administered. In some embodiments, the modulating agent targets (e.g., binds, antagonizes, and/or agonizes) a polypeptide in one or more microorganisms resident in the host (e.g., a protein listed in Table 3 or Table 7).

In some embodiments, the one or more microorganisms resident in the host is an endosymbiotic microorganism. In some embodiments, the one or more microorganisms is resident in the host's gut. In some embodiments, the one or more microorganisms is resident in a bacteriocyte in the host. In some embodiments, the one or more microorganisms resident in the host is a fungus or bacterium. In some embodiments, the bacterium resident in the host is at least one selected from the group consisting of *Candidatus spp*, *Buchenera spp*, *Blattabacterium spp*, *Baumania spp*, *Wigglesworthia spp*, *Wolbachia spp*, *Rickettsia spp*, *Orientia spp*, *Sodalis spp*, *Burkholderia spp*, *Cupriavidus spp*, *Frankia spp*, *Snirhizobium spp*, *Streptococcus spp*, *Wolinella spp*, *Xylella spp*, *Erwinia spp*, *Agrobacterium spp*, *Bacillus spp*, *Paenibacillus spp*, *Streptomyces spp*, *Micrococcus spp*, *Corynebacterium spp*, *Acetobacter spp*, *Cyanobacteria spp*, *Salmonella spp*, *Rhodococcus spp*, *Pseudomonas spp*, *Lactobacillus spp*, *Enterococcus spp*, *Alcaligenes spp*, *Klebsiella spp*, *Paenibacillus spp*, *Arthrobacter spp*, *Corynebacterium spp*, *Brevibacterium spp*, *Thermus spp*, *Pseudomonas spp*, *Clostridium spp*, and *Escherichia spp*. In some embodiments, the fungus resident in the host is at least one selected from the group consisting of *Candida*, *Metschnikowia*, *Debaromyces*, *Starmerella*, *Pichia*, *Cryptococcus*, *Pseudozyma*, *Symbiotaphrina buchneri*, *Symbiotaphrina kochii*, *Scheffersomyces shehatae*, *Scheffersomyces stipites*, *Cryptococcus*, *Trichosporon*, *Amylostereum areolatum*, *Epichloe spp*, *Pichia pinus*, *Hansenula capsulate*, *Daldinia decipien*, *Ceratocytis spp*, *Ophiostoma spp*, and *Attamyces bromatificus*. In some embodiments, the modulating agent alters the growth, division, viability, metabolism, and/or longevity of the microorganism resident in the host. In some embodiments, the modulating agent decreases the growth, division, viability, metabolism, and/or longevity of the one or more microorganisms. In some embodiments, the modulating agent increases the growth, division, viability, metabolism, and/or longevity of the one or more microorganisms.

In some embodiments, the modulating agent is a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi, shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or any combination thereof.

In some embodiments, the modulating agent is a nucleic acid. The nucleic acid may be a DNA molecule, a RNA molecule (e.g., double-stranded RNA (dsRNA) or single-stranded RNA (ssRNA)), or a hybrid DNA-RNA molecule. In some embodiments, the RNA is a messenger RNA (mRNA), a guide RNA (gRNA), or an inhibitory RNA. In some embodiments, the inhibitory RNA is RNAi, shRNA, or miRNA. In some embodiments, the nucleic acid encodes a polypeptide. In some embodiments, the nucleic acid is an expression vector encoding a polypeptide. In some embodiments, the nucleic acid is a CRISPR nucleic acid.

In some embodiments, the modulating agent is a small molecule. In some embodiments, the small molecule is an agonist, antagonist, inhibitor, or an activator of a component of a host immune system pathway or bacteriocyte regulatory pathway. In some embodiments, the small molecule is prostaglandin.

In some embodiments, the modulating agent is a polypeptide. In some embodiments, the polypeptide is an antibody or an antibody fragment. For example, the antibody or antibody fragment may be an agonist or antagonist of an enzyme in the host (e.g., an immune system or bacteriocyte-regulatory enzyme) or in the microorganism resident in the host, including any of the proteins listed in Table 5, Table 7, Table 8, or Table 9.

In some embodiments, the modulating agent (e.g., a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi, shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or a combination thereof) modulates the host's fitness by increasing or decreasing the host's susceptibility to a pesticide (e.g., a pesticide listed in Table 11). In some embodiments, the pesticide is a bactericide or fungicide. In some embodiments, the pesticide is an insecticide, molluscicide, or nematocide.

In some embodiments, the composition includes a plurality (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, or more than 10) of different modulating agents (e.g., a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi, shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or a combination thereof). In some embodiments, the composition includes a modulating agent and a pesticide (e.g., a pesticide listed in Table 11). In some embodiments, the pesticide is a bactericide or fungicide. In some embodiments, the pesticide is an insecticide, molluscicide, or nematocide. In some embodiments, the composition includes a modulating agent and an agent that increases crop growth.

In some embodiments, the modulating agent (e.g., a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi, shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or a combination thereof) is linked to a second moiety. In some embodiments, the second moiety is selected from the group consisting of a modulating agent, peptide nucleic acid, cell penetrating peptide (CPP), and targeting domain. In some embodiments, the modulating agent includes a linker. In some embodiments, the linker is a cleavable linker. In some embodiments, the CPP is any one listed in Table 10.

In some embodiments, the composition further includes a carrier. In some embodiments, the carrier is an agriculturally acceptable carrier. In some embodiments, the composition further includes a host bait, a sticky agent, or a combination thereof. In some embodiments, the host bait is a comestible agent. In some embodiments, the host bait is a chemoattractant.

In some embodiments, the composition is at a dose effective to modulate host fitness. In some embodiments, the composition is at a dose effective to increase host fitness. In alternative embodiments, the composition is at a dose effective to decrease host fitness. In some embodiments, host fitness is measured by survival, lifespan, reproduction, or metabolism of the host.

In some embodiments, the composition is formulated for delivery to a microorganism inhabiting the gut of the host. In some embodiments, the composition is formulated for delivery to a microorganism inhabiting a bacteriocyte of the host. In some embodiments, the composition is formulated for delivery to

a plant. In some embodiments, the composition is formulated for use in a host feeding station. In some  
embodiments, the composition is formulated as a liquid, a powder, granules, or nanoparticles. In some  
embodiments, the composition is formulated as one selected from the group consisting of a liposome,  
polymer, bacteria secreting peptide, and synthetic nanocapsule. In some embodiments, the synthetic  
5 nanocapsule delivers the composition to a target site in the host. In some embodiments, the target site is  
the gut of the host. In some embodiments, the target site is a bacteriocyte in the host.

In another aspect, provided herein are plants including any of the previous compositions (e.g., a  
modulating agent (e.g., a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte  
regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi,  
10 shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or a combination  
thereof)). In some embodiments, the plant includes a nucleic acid integrated into the plant genome,  
wherein the nucleic acid encodes any of the previous modulating agents (e.g., a polypeptide (e.g., an  
antibody, a bacteriocin (e.g., colA), an antimicrobial peptide, a bacteriocyte regulatory peptide, a nucleic  
acid, or a small molecule). The modulating agent may be non-endogenous to the plant. In some  
15 embodiments, the plant further includes a comestible agent for invertebrates (e.g., insect, mollusk, or  
nematode), wherein the comestible agent produces and/or carries the modulating agent. In some  
embodiments, the comestible agent includes one or more components of the plant. In some  
embodiments, the one or more components of the plant includes a root, stem, leaf, flower, sap, bark,  
wood, spine, pollen, nectar, seed, fruit, or any combination thereof. For example, in some embodiments,  
20 the plant produces a modulating agent that the insect ingests by eating one or more components of the  
plant.

In yet another aspect, provided herein are hosts including any of the previous compositions (e.g.,  
a modulating agent (e.g., a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte  
regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi,  
25 shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or a combination  
thereof)). The host may be an invertebrate (e.g., insect, mollusk, or nematode). In some embodiments,  
the invertebrate is an insect. In some embodiments, the insect is a bacteriocyte-containing insect. For  
example, in certain embodiments, the bacteriocyte-containing insect may be an aphid (e.g., a corn leaf  
aphid or green peach aphid). In further embodiments, the insect is a beetle, weevil, fly, aphid, whitefly,  
30 leafhopper, scale, moth, butterfly, grasshopper, cricket, thrip, or mite. In other embodiments, the  
invertebrate is a mollusk. In some embodiments, the mollusk is a species belonging to Veronicellidae,  
Ampullariidae, Achatinidae, Helicidae, Hydromiidae, Planorbidae, Lymnaeidae, Urocyclidae,  
Bradybaenidae, Agriolimacidae, Arionidae, or Milacidae. In another embodiment, the invertebrate may be  
a nematode. In some embodiments, the nematode is a species belonging to Criconematidae,  
35 Belonolaimidae, Hoplolaimidae, Heteroderidae, Longidoridae, Pratylenchidae, Trichodoridae, or  
Anguinidae.

In another aspect, provided herein is a system for modulating (e.g., increasing or decreasing) a  
host's fitness. The system includes a modulating agent (e.g., a polypeptide (e.g., antibody, bacteriocin,  
antimicrobial peptide, or bacteriocyte regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA,  
40 gRNA, or inhibitory RNA (e.g., RNAi, shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g.,

prostaglandin), or a combination thereof) that alters interactions between the host and one or more microorganisms resident in the host, wherein the system is effective to modulate (e.g., increasing or decreasing) the host's fitness, and wherein the host is an invertebrate (e.g., insect (e.g., an aphid), mollusk, or nematode). In some embodiments, the modulating agent of the system is any of the previous compositions. In some embodiments, the modulating agent is formulated as a powder. In some  
 5       embodiments, the modulating agent is formulated as a solvent. In some embodiments, the modulating agent is formulated as a concentrate. In some embodiments, the modulating agent is formulated as a diluent. In some embodiments, the modulating agent is prepared for delivery by combining any of the previous compositions with a carrier.

10       In another aspect, provided herein are methods of modulating a host-microbiota interaction that includes delivering any of the compositions described herein (e.g., a modulating agent (e.g., a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi, shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or a combination thereof)) to the host,  
 15       wherein the modulating agent modulates one or more interactions between the host and one or more microorganisms resident in the host.

20       In another aspect, provided herein are methods of modulating the fitness of an invertebrate host (e.g., insect (e.g., an aphid), mollusk, or nematode), wherein the method includes delivering any of the compositions described herein (e.g., a modulating agent (e.g., a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte regulatory peptide), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi, shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or a combination thereof)) to the host and wherein the modulating agent alters interactions between the host and one or more microorganisms resident in the host.

25       In some embodiments of any of the above methods, the one or more microorganisms resident in the host may be a fungus or bacterium. In some embodiments, the one or more microorganisms is an endosymbiotic microorganism. In some embodiments, the one or more microorganisms is resident in the host's gut. In some embodiments, the one or more microorganisms is resident in a bacteriocyte in the host. In some embodiments, the one or more microorganisms are required for host fitness or host survival.

30       In some embodiments of any of the above methods, the modulating agent may target one or more host pathways that mediate interactions between the host and the one or more microorganisms. In some embodiments, the host pathway is a pathway that regulates insect (e.g., an aphid) bacteriocyte function or development. In some embodiments, the targeting of the host bacteriocyte function or development decreases the level, diversity, and/or function of one or more microorganisms resident in the  
 35       bacteriocyte. Alternatively, the targeting of the host bacteriocyte function or development increases the level, diversity, and/or function of one or more microorganisms resident in the bacteriocyte. In some embodiments, the host pathway is a pathway that regulates the host's immune system.

40       In some embodiments, the modulating agent activates an immune response against the one or more microorganisms resident in the host, thereby decreasing the level, diversity, and/or function of the one or more microorganisms. In some embodiments, the modulating agent suppresses an immune

response against the one or more microorganisms resident in the host, thereby increasing the level, diversity, and/or function of the one or more microorganisms.

In some embodiments, the modulating agent targets one or more microbial pathways that mediate interactions between the host and the one or more microorganisms.

5 In some embodiments, the delivering step includes providing the modulating agent at a dose and time sufficient to effect the one or more microorganisms, thereby modulating microbial diversity in the host. In some embodiments, the delivering step includes topical application of any of the previous compositions to a plant. In some embodiments, the delivering step includes providing the modulating agent through a genetically modified, engineered, or transgenic plant (e.g., any of the plants described  
10 herein). In other embodiments, the delivering step includes providing the modulating agent to the host as a comestible agent for invertebrates (e.g., insect, mollusk, or nematode). In further embodiments, the delivering step includes providing a host carrying the modulating agent. In some embodiments, the host carrying the modulating agent can transmit the modulating agent to one or more additional hosts.

Also provided herein are screening assays to identify a modulating agent that modulates (e.g.,  
15 increases or decreases) the fitness of a host. The screening assay may include the steps of (a) exposing a microorganism that can be resident in the host to one or more candidate modulating agents and (b) identifying a modulating agent that increases or decreases the fitness of the host. In some embodiments, the modulating agent is a microorganism resident in the host. In some embodiments, the microorganism is a bacterium. In some embodiments, the bacterium, when resident in the host, increases host fitness.  
20 Alternatively, the bacterium, when resident in the host, decreases host fitness. In some embodiments, the modulating agent is any of the modulating agents described herein (e.g., a modulating agent (e.g., a polypeptide (e.g., antibody, bacteriocin, antimicrobial peptide, or bacteriocyte regulatory peptide (e.g., Coleopteracin A), a nucleic acid (e.g., DNA, RNA (e.g., mRNA, gRNA, or inhibitory RNA (e.g., RNAi, shRNA, miRNA)), CRISPR nucleic acid), a small molecule (e.g., prostaglandin), or a combination  
25 thereof)). In some embodiments, the modulating agent is provided by a genetically modified phage or bacteria. In some embodiments, the host's fitness is modulated by modulating the host microbiota.

### *Definitions*

As used herein, the term "bacteriocyte" refers to a specialized cell found in invertebrates, e.g.,  
30 insects, nematodes, or mollusks, where intracellular bacteria reside with symbiotic bacterial properties. In some instances, the bacteriocyte may be clustered with other bacteriocytes to form a bacteriome.

As used herein, the term "effective amount" refers to an amount of a modulating agent (e.g., a polypeptide, nucleic acid, small molecule, or combinations thereof) or composition including said agent sufficient to effect the recited result, e.g., to increase or decrease the fitness of a host organism (e.g.,  
35 insect, nematode, or mollusk); to reach a target level (e.g., a predetermined or threshold level) of a modulating agent concentration inside a target host; to reach a target level (e.g., a predetermined or threshold level) of a modulating agent concentration inside a target host gut; to reach a target level (e.g., a predetermined or threshold level) of a modulating agent concentration inside a target host bacteriocyte; to modulate the level, or an activity, of one or more microorganisms (e.g., endosymbiont) in the target  
40 host.

As used herein, the term “fitness” refers to the ability of a host invertebrate (e.g., insect, mollusk, or nematode) to survive, and/or to produce surviving offspring. Fitness of a host (e.g., insect, mollusk, or nematode) may be measured by one or more parameters, including, but not limited to, life span, reproductive rate, mobility, body weight, or metabolic rate. Depending on the host, fitness may additionally be measured based on measures of activity (e.g., biting animals or humans, plant pollination), disease transmission (e.g., vector-vector transmission or vector-animal transmission), or production (e.g., honey or silk).

As used herein, the term “gut” refers to any portion of a host’s gut, including, the foregut, midgut, or hindgut of the host.

As used herein, the term “host” refers to an organism, such as an invertebrate (e.g., insect, mollusk, or nematode) carrying resident microorganisms (e.g., endogenous microorganisms, endosymbiotic microorganisms (e.g., primary or secondary endosymbionts), commensal organisms, and/or pathogenic microorganisms).

As used herein “decreasing host fitness” or “reducing host fitness” refers to any disruption to host physiology, or any activity carried out by said host, as a consequence of administration of a modulating agent, including, but not limited to, any one or more of the following desired effects: (1) decreasing a population of a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (2) decreasing the reproductive rate of a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (3) decreasing the mobility of a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (4) decreasing the body weight of a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (5) decreasing the metabolic rate or activity of a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (6) decreasing plant infestation by a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (7) decrease disease transmission (e.g., of a plant, animal, or human pathogen) by a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; or (8) decrease growth, increase nymphal mortality, and/or increase adult sterility of a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more. A decrease in host fitness can be determined in comparison to a host organism to which the modulating agent has not been administered.

As used herein “increasing host fitness” or “promoting host fitness” refers to any favorable alteration in host physiology, or any activity carried out by said host, as a consequence of administration of a modulating agent, including, but not limited to, any one or more of the following desired effects: (1) increasing a population of a host by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (2) increasing the reproductive rate of a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (3) increasing the mobility of a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (4) increasing the body weight of a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (5)

increasing the metabolic rate or activity of a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (6) increasing pollination (e.g., number of plants pollinated in a given amount of time) by a host (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (7) increasing  
 5 production of host (e.g., insect, mollusk, or nematode) byproducts (e.g., honey or silk) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; (8) increasing nutrient content of the host (e.g., insect, mollusk, or nematode) (e.g., protein, fatty acids, or amino acids) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100% or more; or (9) increasing host resistance to pesticides (e.g., insect, mollusk, or nematode) by about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%,  
 10 90%, 95%, 99%, 100% or more. An increase in host fitness can be determined in comparison to a host organism to which the modulating agent has not been administered.

As used herein, "interactions between a host and microorganisms resident in the host" or "host-microbiota interactions" refer to (i) any pathways (e.g., metabolic, gene regulation, cell signaling, or immune-inflammatory pathways) in the host that directly or indirectly influences the survival, growth, or  
 15 metabolism of microorganisms resident in the host (e.g., endosymbiotic microorganisms), (ii) any pathways (e.g., metabolic or cell signaling pathways) in a resident microorganism that directly or indirectly influences the fitness of the host invertebrate (e.g., insect, nematode, or mollusk), and/or (iii) any pathways (e.g., metabolic, cell signaling, or immune-inflammatory pathways) in a resident microorganism that directly or indirectly influences survival, growth or metabolism of the host, and/or (iv) any pathways  
 20 (e.g., metabolic, gene regulation, cell signaling, or immune-inflammatory pathways) in the host that directly or indirectly influences the fitness of the resident microorganism.

The term "insect" includes any organism belonging to the phylum *Arthropoda* and to the class *Insecta* or the class *Arachnida*, in any stage of development, i.e., immature and adult insects.

The term "mollusk" includes any organism belonging to the phylum Mollusca, including organisms  
 25 of the class Gastropoda (e.g., snails and slugs), in any stage of development, i.e., immature and adult mollusks.

The term "nematode" includes any organism belonging to the phylum Nematoda (e.g., nematodes) in any stage of development, i.e., immature and adult nematodes.

As used herein, the term "microorganism" or "microbiota" refers to bacteria or fungi.  
 30 Microorganisms may refer to microorganisms resident in a host organism (e.g., endogenous microorganisms, endosymbiotic microorganisms (e.g., primary or secondary endosymbionts)) or microorganisms exogenous to the host, including those that may act as modulating agents. As used herein, the term "target microorganism" refers to a microorganism that is resident in the host and impacted by a modulating agent, either directly or indirectly.

35 As used herein, the term "modulating agent" or "agent" refers to an agent that is capable of altering the levels and/or functioning of microorganisms resident in a host organism (e.g., invertebrate, e.g., insect, mollusk, or nematode), and thereby modulate (e.g., increase or decrease) the fitness of the host organism.

As defined herein, the term "nucleic acid" and "polynucleotide" are interchangeable and refer to  
 40 RNA or DNA that is linear or branched, single or double stranded, or a hybrid thereof, regardless of length

(e.g., at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 150, 200, 250, 500, 1000, or more nucleic acids). The term also encompasses RNA/DNA hybrids. Nucleotides are typically linked in a nucleic acid by phosphodiester bonds, although the term “nucleic acid” also encompasses nucleic acid analogs having other types of linkages or backbones (e.g., phosphoramidate, phosphorothioate, phosphorodithioate, O-methylphosphoroamidate, morpholino, locked nucleic acid (LNA), glycerol nucleic acid (GNA), threose nucleic acid (TNA), and peptide nucleic acid (PNA) linkages or backbones, among others). The nucleic acids may be single-stranded, double-stranded, or contain portions of both single-stranded and double-stranded sequence. A nucleic acid can contain any combination of deoxyribonucleotides and ribonucleotides, as well as any combination of bases, including, for example, adenine, thymine, cytosine, guanine, uracil, and modified or non-canonical bases (including, e.g., hypoxanthine, xanthine, 7-methylguanine, 5,6-dihydrouracil, 5-methylcytosine, and 5 hydroxymethylcytosine).

As used herein, the term “pest” refers to invertebrates (e.g., insects, nematodes, or mollusks) that cause damage to plants or other organisms, or otherwise are detrimental to humans, for example, human agricultural methods or products.

As used herein, the term “pesticide” or “pesticidal agent” refers to a substance that can be used in the control of agricultural, environmental, and domestic/household pests, such as insects, mollusks, nematodes, fungi, bacteria, and viruses. The term “pesticide” is understood to encompass naturally occurring or synthetic insecticides (larvicides or adulticides), insect growth regulators, nematocides, molluscicides, acaricides (miticides), nematocides, ectoparasiticides, bactericides, fungicides, or herbicides (substance which can be used in agriculture to control or modify plant growth). Further examples of pesticides or pesticidal agents are listed in Table 11. In some instances, the pesticide is an allelochemical. As used herein, “allelochemical” or “allelochemical agent” is a substance produced by an organism that can effect a physiological function (e.g., the germination, growth, survival, or reproduction) of another organism (e.g., an insect, mollusk, or nematode).

As used herein, the term “peptide,” “protein,” or “polypeptide” encompasses any chain of naturally or non-naturally occurring amino acids (either D- or L-amino acids), regardless of length (e.g., at least 2, 3, 4, 5, 6, 7, 10, 12, 14, 16, 18, 20, 25, 30, 40, 50, 100, or more amino acids), the presence or absence of post-translational modifications (e.g., glycosylation or phosphorylation), or the presence of, e.g., one or more non-amino acyl groups (for example, sugar, lipid, etc.) covalently linked to the peptide, and includes, for example, natural proteins, synthetic, or recombinant polypeptides and peptides, hybrid molecules, peptoids, or peptidomimetics.

As used herein, “percent identity” between two sequences is determined by the BLAST 2.0 algorithm, which is described in Altschul et al., *J. Mol. Biol.* 215:403-410, 1990. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information.

As used herein, the term “plant” refers to whole plants, plant organs, plant tissues, seeds, plant cells, seeds, or progeny of the same. Plant cells include, without limitation, cells from seeds, suspension cultures, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, or microspores. Plant parts include differentiated and undifferentiated tissues including, but not limited to the following: roots, stems, shoots, leaves, pollen, seeds, tumor tissue, or various forms of cells

and culture (e.g., single cells, protoplasts, embryos, and callus tissue). The plant tissue may be in a plant or in a plant organ, tissue, or cell culture.

As used herein a “transgenic plant,” “genetically engineered plant,” or “genetically modified plant” refers to a plant whose genome (e.g., chromosomal DNA, chloroplast DNA, or mitochondrial DNA) has been altered by the stable integration of recombinant DNA. A transgenic plant includes a plant regenerated from an originally-transformed plant cell and progeny transgenic plants from later generations or crosses of a transformed plant. For example, a transgenic plant may be genetically engineered to produce a heterologously (e.g., non-endogenous) or non-heterologously (e.g., endogenous) encoded protein or RNA, for example, of any of the modulating agents in the methods or compositions described herein. Any plant species may be transformed to create a transgenic plant. The transgenic plant may be a dicotyledonous plant or a monocotyledonous plant. For example and without limitation, transgenic plants of the compositions and methods described herein may be derived from any of the following dicotyledonous plant families: Leguminosae, including plants such as pea, alfalfa and soybean; Umbelliferae, including plants such as carrot and celery; Solanaceae, including the plants such as tomato, potato, aubergine, tobacco, and pepper; Cruciferae, particularly the genus Brassica, which includes plant such as oilseed rape, beet, cabbage, cauliflower and broccoli); and Arabidopsis thaliana; Compositae, which includes plants such as lettuce; Malvaceae, which includes cotton; Fabaceae, which includes plants such as peanut, and the like. Transgenic plants of the invention may be derived from monocotyle- donous plants, such as, for example, wheat, barley, sorghum, millet, rye, triticale, maize, rice, oats, switchgrass, miscanthus, and sugarcane. Transgenic plants of the invention are also embodied as trees such as apple, pear, quince, plum, cherry, peach, nectarine, apricot, papaya, mango, and other woody species including coniferous and deciduous trees such as poplar, pine, sequoia, cedar, oak, willow, and the like.

Other features and advantages of the invention will be apparent from the following Detailed Description and the Claims.

### BRIEF DESCRIPTION OF THE FIGURES

The figures are meant to be illustrative of one or more features, aspects, or embodiments of the invention and are not intended to be limiting.

Fig. 1 is a panel of graphs showing that treatment with *P. pastoris* delayed aphid development. First and second instar LSR-1 aphids were placed on leaves perfused with water (negative control) or with a solution of *P. pastoris* in water and developmental stage was monitored at each indicated time point during the experiment. Shown are the mean percentages of aphids in each group  $\pm$  SD.

Fig. 2 is a graph showing that *P. pastoris* treatment resulted in aphid death. First and second instar LSR-1 aphids were treated with water (control) or with *P. pastoris* via leaf perfusion and survival was monitored daily during the experiment. N=62-63 aphids/group. Statistically significant differences were determined using Log-Rank (Mantel Cox) test. \*\*\*\*,  $p < 0.0001$ .

Fig. 3 is a panel of graphs showing that *P. pastoris* treatment via leaf spraying did not affect aphid development. First and second instar LSR-1 aphids were treated with water (control) or *P. pastoris* via airbrush spraying and developmental stage was monitored over time. Shown are the mean percentages

± SD of dead or live aphids in each developmental stage (1st, 2nd, 3rd, 4th, or 5th instar) at each time point. N=two replicates of 30 aphids/group.

Fig. 4 is a graph showing that *P. pastoris* treatment increased aphid mortality. First and second instar LSR-1 aphids were placed on leaves sprayed with water (control) or *P. pastoris* and survival was monitored over time. N=60 aphids/treatment group.

Figs. 5A and 5B are graphs showing that spraying *P. pastoris* on fava bean leaves reduced endosymbiotic Buchnera in aphids feeding on the leaves. Symbiont titer was determined at 6 (A) and 9 (B) days post-treatment with *P. pastoris*. Shown are the mean Buchnera/aphid copies ± SD. The number in the box above the indicated dataset represents the median value of that group. Each dot represents a single aphid.

Fig. 6 is a graph showing microinjection of BCR-4 PNA reduced BCR-4 expression. Fourth and fifth instar *A. pisum* aphids were injected with 20 nl water or 321 ng/ul of BCR-4 PNA, RNA was extracted from aphids after 7 days, and RT-qPCR was performed to measure expression of BCR-4. Shown are the mean BCR-4/Actin copies ± SD. Each data point represents a single aphid. The number in the box above each dataset represents the median of the data.

Fig. 7 is a graph showing the decrease in insect survival after treatment with a PNA to BCR-4. Fourth and fifth instar *A. pisum* aphids were injected with water or with a PNA to BCR-4. Survival was monitored daily over the course of the experiment. N=20 aphids per treatment group.

Fig. 8 is a graph showing injection of aphids with a PNA to BCR-4 resulted in decreased fecundity. Fourth and fifth instar *A. pisum* aphids were injected with water or with a PNA to BCR-4. Fecundity was measured by counting the number of offspring produced by each group at each time point and is represented as the number of F1's (first generation offspring) produced by F0 (adults) per day. N=20 aphids per treatment group.

Fig. 9 is a panel of graphs showing treatment with BCR-4 PNA delayed aphid development. First and second instar LSR-1 aphids were placed on leaves perfused with water (negative control) or with a solution of BCR-4 PNA in water and developmental stage was monitored at each indicated time point during the experiment. Shown are the mean percentages of aphids in each group ± SD.

Fig. 10 is a graph showing BCR-4 treatment resulted in increased aphid death. First and second instar LSR-1 aphids were treated with water (control) or with BCR-4 PNA via leaf perfusion and survival was monitored daily during the experiment. N=60 aphids/group. Statistically significant differences were determined using Log-Rank (Mantel Cox) test.

Fig. 11 is a graph showing treatment with BCR-4 PNA delivered via leaf perfusion increased Buchnera titers. First and second instar LSR-1 aphids were treated with water (control) or BCR4 PNA via leaf perfusion and dead aphids were collected on days 5 and 6 after treatment. DNA was extracted, and qPCR was performed to determine the number of Buchnera/aphid DNA copies. Shown are the mean number of Buchnera/aphid DNA copies ± SD of 6-7 aphids/group.

Fig. 12 is a graph showing treatment of aphids with a PNA against BCR-4 via leaf perfusion resulted in a reduction of BCR-4 expression. First and second instar LSR-1 aphids were treated with water (control) or BCR-4 PNA via leaf perfusion and on day 7, RNA was extracted from living aphids and

RT-qPCR was performed to quantify expression of BCR-4 relative to actin expression. The number in the box represents the median of the dataset.

Fig. 13 is a graph showing treatment with dsRNA-ApGLNT1 knocked down the expression of ApGLNT1. Fifth instar *A. pisum* aphids were injected with water or dsRNA-ApGLNT1 in water. At 2 days post-treatment, total RNA was extracted and RT-qPCR was performed to determine ApGLNT1 gene relative expression (Actin as internal reference gene). Shown is the mean ratio of relative expression of ApGLNT1/Actin  $\pm$  SD of 5-7 aphids/group. Statistically significant differences were determined by Student's T-test (\*\*,  $p < 0.01$ ).

Fig. 14 is a graph showing treatment with dsRNA-ApGLNT1 increased aphid mortality. LSR-1 *A. pisum* aphids were injected with water or dsRNA-ApGLNT1 in water and survival was monitored over the course of the experiment. N=40 aphids/group. A statistically significant difference was identified between the two groups as determined using a Log-Rank (Mantel Cox) test.

Fig. 15 is a graph showing treatment with dsRNA-ApGLNT1 resulted in decreased Buchnera titers. LSR-1 *A. pisum* aphids were injected with water or dsRNA-ApGLNT1 in water, DNA was extracted from aphids at 5 days post-injection, and qPCR was performed to quantify Buchnera. Shown are the mean copies of Buchnera/aphid DNA  $\pm$  SD. Each dot represents an individual aphid. The number in the box above each data set represents the median of the group.

Figs. 16A and 16B are a panel of graphs showing offspring from aphids microinjected with dsRNA-ApGLNT1 displayed delayed development. Fig. 16A: LSR-1 *A. pisum* aphids were injected with water or dsRNA-ApGLNT1 in water and life stages were monitored at the indicated time point after injection. Shown is the mean percent  $\pm$  of aphids that were dead or alive at each instar stage throughout the experiment. N=40 aphids/group. Fig. 16B: 4 days after offspring were collected, images were taken of each aphid to determine the area of each aphid. Shown is the mean area  $\pm$  SD of offspring taken from aphids injected with either water or dsRNA-ApGLNT1. Statistically significant differences were determined using a Student's t-test. Each data point represents one individual aphid.

Fig. 17 is an illustration showing the dsRNA expression cassette. pCaMV 35S promoter is placed upstream of the dsRNA expressing sequence. The sense and the antisense strands of a region of the target aphid gene are placed in tandem with a small spacer which will act as the hairpin loop. Once expressed, the RNA formed will assume a double stranded configuration due to the complementarity of the sequence.

Fig. 18 is an illustration of the shuttle vector for the constructs for expressing dsRNA in *N. tabacum*. The plasmid includes origins of replications compatible with *E. coli* and *A. tumefaciens*, kanamycin and gentamycin resistance markers, green fluorescence expression cassette under a parsley ubiquitin promoter, and finally the dsRNA expression cassette driven by the pCaMV 35S.

Fig. 19 is a panel of images showing GFP expression in *N. tabacum* plant infiltrated by *A. tumefaciens*. The top panels are *N. tabacum* infiltrated with *A. tumefaciens* containing a plasmid that can constitutively drive the expression of GFP in *N. tabacum* (Top left is brightfield, and top right is green channel). The bottom panels are negative control leaves not infiltrated by *A. tumefaciens*.

## DETAILED DESCRIPTION

Provided herein are methods and compositions for modulating the fitness of a host invertebrate (e.g., insect, mollusk, or nematode) by altering interactions between the host and one or more microorganisms resident in the host. The invention features a composition including a modulating agent (e.g., a polypeptide, nucleic acid, small molecule, or combinations thereof) that can indirectly induce changes in the host's microbiota in a manner that modulates (e.g., increases or decreases) host fitness. For example, the modulating agent may target host pathways (e.g., immune system or bacteriocyte pathways) or microbial pathways that alter (e.g., increase or decrease) microbial levels, microbial activity, microbial metabolism, and/or microbial diversity, and in turn modulates (e.g., increase or decrease) the fitness of a variety of invertebrates (e.g., insect, mollusk, or nematode) that are important for agriculture, commerce, and/or public health.

The methods and compositions described herein are based, in part, on the examples which illustrate how different agents, for example, small compounds (e.g., prostaglandin), inhibitory RNA (e.g., dsRNA or PNAs), and microorganisms (fungi or bacteria) can be used to alter the host's immune system response towards microorganisms resident in the host. The methods and compositions described herein can also be used to alter the function of host organs or cells in which microorganisms typically reside. For example, RNA may be used to impair bacteriocyte function in an aphid, thereby disrupting endosymbiotic microorganism populations resident in the bacteriocyte of the aphid. Disruption of endosymbiotic populations of microorganisms (e.g., *Buchnera* spp.) in the aphid, in turn, decreases the fitness of the aphid. Nucleic acids, such as RNAs (e.g., dsRNA) or PNAs, or small molecules (e.g., prostaglandin) are representative of modulating agents useful in the invention, and other modulating agents of this type may be useful in the invention. On this basis, the present disclosure describes a variety of different approaches for the use of agents that modulates (e.g., increases or decreases) the fitness of an invertebrate host (e.g., insect, mollusk, or nematode), wherein the modulating agent alters interactions between the host and one or more microorganisms resident in the host.

### I. Hosts

#### *i. Insect hosts*

In some instances, the host described herein is an organism belonging to the phylum Arthropoda. In some instances, the insect is considered a pest, e.g., an agricultural pest. In some instances, the insect carries a bacterium or virus that is considered a plant pest that causes disease in a plant (e.g., *Agrobacterium* or tomato yellow leaf curl virus (TYLCV)). The host may be at any stage developmentally. For instance, the host may be an embryo, a larva, a pupa, or an adult.

In some instances, the insect may belong to the following orders: Acari, Araneae, Anoplura, Coleoptera, Collembola, Dermaptera, Dictyoptera, Diplura, Diptera (e.g., spotted-wing *Drosophila*), Embioptera, Ephemeroptera, Grylloblatodea, Hemiptera (e.g., an aphid, Greenhouse whitefly, or stinkbug), Homoptera, Hymenoptera, Isoptera, Lepidoptera, Mallophaga, Mecoptera, Neuroptera, Odonata, Orthoptera, Phasmida, Plecoptera, Protura, Psocoptera, Siphonaptera, Siphunculata, Thysanura, Strepsiptera, Thysanoptera, Trichoptera, or Zoraptera.

In some instances, the insect is from the class Arachnida, for example, *Acarus* spp., *Aceria sheldoni*, *Aculops* spp., *Aculus* spp., *Amblyomma* spp., *Amphitetranynchus viennensis*, *Argas* spp., *Boophilus* spp., *Brevipalpus* spp., *Bryobia graminum*, *Bryobia praetiosa*, *Centruroides* spp., *Chorioptes* spp., *Dermanyssus gallinae*, *Dermatophagoides pteronyssinus*, *Dermatophagoides farinae*, *Dermacentor* spp., *Eotetranychus* spp., *Epitrimerus pyri*, *Eutetranychus* spp., *Eriophyes* spp., *Glycyphagus domesticus*, *Halotydeus destructor*, *Hemitarsonemus* spp., *Hyalomma* spp., *Ixodes* spp., *Latrodectus* spp., *Loxosceles* spp., *Metatetranychus* spp., *Neutrombicula autumnalis*, *Nuphessa* spp., *Oligonychus* spp., *Ornithodoros* spp., *Ornithonyssus* spp., *Panonychus* spp., *Phyllocoptruta oleivora*, *Polyphagotarsonemus latus*, *Psoroptes* spp., *Rhipicephalus* spp., *Rhizoglyphus* spp., *Sarcoptes* spp., *Scorpio maurus*, *Steneotarsonemus* spp., *Steneotarsonemus pinki*, *Tarsonemus* spp., *Tetranychus* spp., *Trombicula alfreddugesi*, *Vaejovis* spp., *Vasates lycopersici*.

In some instances, the insect is from the class Chilopoda, for example, *Geophilus* spp., *Scutigera* spp.

In some instances, the insect is from the order Collembola, for example, *Onychiurus armatus*.

In some instances, the insect is from the class Diplopoda, for example, *Blaniulus guttulatus*.

In some instances, the insect is from the class Insecta, e.g. from the order Blattodea, for example, *Blattella asahinai*, *Blattella germanica*, *Blatta orientalis*, *Leucophaea maderae*, *Panchlora* spp., *Parcoblatta* spp., *Periplaneta* spp., *Supella longipalpa*.

In some instances, the insect is from the order Coleoptera, for example, *Acalymma vittatum*, *Acanthoscelides obtectus*, *Adoretus* spp., *Agelastica alni*, *Agriotes* spp., *Alphitobius diaperinus*, *Amphimallon solstitialis*, *Anobium punctatum*, *Anoplophora* spp., *Anthonomus* spp., *Anthrenus* spp., *Apion* spp., *Apogonia* spp., *Atomaria* spp., *Attagenus* spp., *Bruchidius obtectus*, *Bruchus* spp., *Cassida* spp., *Cerotoma trifurcata*, *Ceutorrhynchus* spp., *Chaetocnema* spp., *Cleonus mendicus*, *Conoderus* spp., *Cosmopolites* spp., *Costelytra zealandica*, *Ctenicera* spp., *Curculio* spp., *Cryptolestes ferrugineus*, *Cryptorhynchus lapathi*, *Cylindrocopturus* spp., *Dendroctonus* spp. (e.g., *Dendroctonus ponderosae*), *Dermestes* spp., *Diabrotica* spp. (e.g., corn rootworm), *Dichocrocis* spp., *Dicladispa armigera*, *Diloboderus* spp., *Epilachna* spp., *Epitrix* spp., *Faustinus* spp., *Gibbium psylloides*, *Gnathocerus cornutus*, *Hellula undalis*, *Heteronychus arator*, *Heteronyx* spp., *Hylamorpha elegans*, *Hylotrupes bajulus*, *Hypera postica*, *Hypomeces squamosus*, *Hypothenemus* spp. (*Hypothenemus hampei*), *Lachnosterna consanguinea*, *Lasioderma serricorne*, *Latheticus oryzae*, *Lathridius* spp., *Lema* spp., *Leptinotarsa decemlineata*, *Leucoptera* spp., *Lissorhoptrus oryzophilus*, *Lixus* spp., *Luperodes* spp., *Lyctus* spp., *Megascelis* spp., *Melanotus* spp., *Meligethes aeneus*, *Melolontha* spp., *Migdolus* spp., *Monochamus* spp., *Naupactus xanthographus*, *Necrobia* spp., *Niptus hololeucus*, *Oryctes rhinoceros*, *Oryzaephilus surinamensis*, *Oryzaphagus oryzae*, *Otiorrhynchus* spp., *Oxycetonia jucunda*, *Phaedon cochleariae*, *Phyllophaga* spp., *Phyllophaga helleri*, *Phyllotreta* spp., *Popillia japonica*, *Premnotrypes* spp., *Prostephanus truncatus*, *Psylliodes* spp., *Ptinus* spp., *Rhizobius ventralis*, *Rhizopertha dominica*, *Sitophilus* spp., *Sitophilus oryzae*, *Sphenophorus* spp., *Stegobium paniceum*, *Sternechus* spp., *Symphyletes* spp., *Tanymecus* spp., *Tenebrio molitor*, *Tenebrioides mauretanicus*, *Tribolium* spp., *Trogoderma* spp., *Tychius* spp., *Xylotrechus* spp., *Zabrus* spp.;

from the order Diptera, for example, *Aedes* spp., *Agromyza* spp., *Anastrepha* spp., *Anopheles* spp.,

*Asphondylia* spp., *Bactrocera* spp., *Bibio hortulanus*, *Calliphora erythrocephala*, *Calliphora vicina*,  
*Ceratitis capitata*, *Chironomus* spp., *Chrysomyia* spp., *Chrysops* spp., *Chrysozona pluvialis*, *Cochliomyia*  
spp., *Contarinia* spp., *Cordylobia anthropophaga*, *Cricotopus sylvestris*, *Culex* spp., *Culicoides* spp.,  
*Culiseta* spp., *Cuterebra* spp., *Dacus oleae*, *Dasyneura* spp., *Delia* spp., *Dermatobia hominis*, *Drosophila*  
5 spp., *Echinocnemus* spp., *Fannia* spp., *Gasterophilus* spp., *Glossina* spp., *Haematopota* spp., *Hydrellia*  
spp., *Hydrellia griseola*, *Hylemya* spp., *Hippobosca* spp., *Hypoderma* spp., *Liriomyza* spp., *Lucilia* spp.,  
*Lutzomyia* spp., *Mansonina* spp., *Musca* spp. (e.g., *Musca domestica*), *Oestrus* spp., *Oscinella frit*,  
*Paratanytarsus* spp., *Paralauterborniella subcincta*, *Pegomyia* spp., *Phlebotomus* spp., *Phorbia* spp.,  
*Phormia* spp., *Piophilina casei*, *Prodiptosis* spp., *Psila rosae*, *Rhagoletis* spp., *Sarcophaga* spp., *Simulium*  
10 spp., *Stomoxys* spp., *Tabanus* spp., *Tetanops* spp., *Tipula* spp.

In some instances, the insect is from the order Heteroptera, for example, *Anasa tristis*,  
*Antestiopsis* spp., *Boisea* spp., *Blissus* spp., *Calocoris* spp., *Campylomma livida*, *Cavelerius* spp., *Cimex*  
spp., *Collaria* spp., *Creontiades dilutus*, *Dasynus piperis*, *Dichelops furcatus*, *Diconocoris hewetti*,  
*Dysdercus* spp., *Euschistus* spp., *Eurygaster* spp., *Heliopeltis* spp., *Horcias nobilellus*, *Leptocoris* spp.,  
15 *Leptocoris varicornis*, *Leptoglossus phyllopus*, *Lygus* spp., *Macropes excavatus*, *Miridae*, *Monalonion*  
*atratum*, *Nezara* spp., *Oebalus* spp., *Pentomidae*, *Piesma quadrata*, *Piezodorus* spp., *Psallus* spp.,  
*Pseudacysta perseae*, *Rhodnius* spp., *Sahlbergella singularis*, *Scaptocoris castanea*, *Scotinophora* spp.,  
*Stephanitis nashi*, *Tibraca* spp., *Triatoma* spp.

In some instances, the insect is from the order Hemiptera or suborder Homoptera, for example,  
20 *Acizzia acaciaebaileyanae*, *Acizzia dodonaeae*, *Acizzia uncatoides*, *Acrida turrita*, *Acyrtosipon* spp.,  
*Acrogonia* spp., *Aeneolamia* spp., *Agonosceana* spp., *Aleyrodes proletella*, *Aleurolobus barodensis*,  
*Aleurothrixus floccosus*, *Allocaridara malayensis*, *Amrasca* spp., *Anuraphis cardui*, *Aonidiella* spp.,  
*Aphanostigma pini*, *Aphis* spp. (e.g., *Apis gossypii*), *Arboridia apicalis*, *Arytainilla* spp., *Aspidiella* spp.,  
*Aspidiotus* spp., *Atanus* spp., *Aulacorthum solani*, *Bemisia tabaci*, *Blastopsylla occidentalis*,  
25 *Boreioglycaspis melaleucae*, *Brachycaudus helichrysi*, *Brachycolus* spp., *Brevicoryne brassicae*,  
*Cacopsylla* spp., *Calligypona marginata*, *Carneocephala fulgida*, *Ceratovacuna lanigera*, *Cercopidae*,  
*Ceroplastes* spp., *Chaetosiphon fragaefolii*, *Chionaspis tegalensis*, *Chlorita onukii*, *Chondracris rosea*,  
*Chromaphis juglandicola*, *Chrysomphalus ficus*, *Cicadulina mbila*, *Coccoxystus halli*, *Coccus* spp.,  
*Cryptomyzus ribis*, *Cryptoneossa* spp., *Ctenarytaina* spp., *Dalbulus* spp., *Dialeurodes citri*, *Diaphorina*  
30 *citri*, *Diaspis* spp., *Drosicha* spp., *Dysaphis* spp., *Dysmicoccus* spp., *Empoasca* spp., *Eriosoma* spp.,  
*Erythroneura* spp., *Eucalyptolyma* spp., *Euphyllura* spp., *Euscelis bilobatus*, *Ferrisia* spp., *Geococcus*  
*coffaeae*, *Glycaspis* spp., *Heteropsylla cubana*, *Heteropsylla spinulosa*, *Homalodisca coagulata*,  
*Homalodisca vitripennis*, *Hyalopterus arundinis*, *Icerya* spp., *Idiocerus* spp., *Idioscopus* spp., *Laodelphax*  
*striatellus*, *Lecanium* spp., *Lepidosaphes* spp., *Lipaphis erysimi*, *Macrosiphum* spp., *Macrosteles*  
35 *facifrons*, *Mahanarva* spp., *Melanaphis sacchari*, *Metcalfiella* spp., *Metopolophium dirhodum*, *Monellia*  
*costalis*, *Monelliopsis pecanis*, *Myzus* spp., *Nasonovia ribisnigri*, *Nephotettix* spp., *Nettigoniclla spectra*,  
*Nilaparvata lugens*, *Oncometopia* spp., *Orthezia praelonga*, *Oxya chinensis*, *Pachypsylla* spp.,  
*Parabemisia myricae*, *Paratrioza* spp., *Parlatoria* spp., *Pemphigus* spp., *Peregrinus maidis*, *Phenacoccus*  
40 spp., *Phloeomyzus passerinii*, *Phorodon humuli*, *Phylloxera* spp., *Pinnaspis aspidistrae*, *Planococcus*  
spp., *Prosopidopsylla flava*, *Protopulvinaria pyriformis*, *Pseudaulacaspis pentagona*, *Pseudococcus* spp.,

*Psyllopsis* spp., *Psylla* spp., *Pteromalus* spp., *Pyrilla* spp., *Quadraspidotus* spp., *Quesada* gigas, *Rastrococcus* spp., *Rhopalosiphum* spp., *Saissetia* spp., *Scaphoideus titanus*, *Schizaphis graminum*, *Selenaspidus articulatus*, *Sogata* spp., *Sogatella furcifera*, *Sogatodes* spp., *Stictocephala festina*, *Siphoninus phillyreae*, *Tenalaphara malayensis*, *Tetragonocephala* spp., *Tinocallis caryaefoliae*,  
 5 *Tomaspis* spp., *Toxoptera* spp., *Trialeurodes vaporariorum*, *Trioza* spp., *Typhlocyba* spp., *Unaspis* spp., *Viteus vitifolii*, *Zygina* spp.; from the order Hymenoptera, for example, *Acromyrmex* spp., *Athalia* spp., *Atta* spp., *Diprion* spp., *Hoplocampa* spp., *Lasius* spp., *Monomorium pharaonis*, *Sirex* spp., *Solenopsis invicta*, *Tapinoma* spp., *Urocerus* spp., *Vespa* spp., *Xeris* spp. In certain instances, the insect is an aphid (e.g., *Rhopalosiphum maidis* or *Myzus persicae*).

10 In some instances, the insect is from the order Isopoda, for example, *Armadillidium vulgare*, *Oniscus asellus*, *Porcellio scaber*.

In some instances, the insect is from the order Isoptera, for example, *Coptotermes* spp., *Cornitermes cumulans*, *Cryptotermes* spp., *Incisitermes* spp., *Microtermes obesi*, *Odontotermes* spp., *Reticulitermes* spp.

15 In some instances, the insect is from the order Lepidoptera, for example, *Achroia grisella*, *Acronicta major*, *Adoxophyes* spp., *Aedia leucomelas*, *Agrotis* spp., *Alabama* spp., *Amyelois transitella*, *Anarsia* spp., *Anticarsia* spp., *Argyroplote* spp., *Barathra brassicae*, *Borbo cinnara*, *Bucculatrix thurberiella*, *Bupalus piniarius*, *Busseola* spp., *Cacoecia* spp., *Caloptilia theivora*, *Capua reticulana*, *Carpocapsa pomonella*, *Carposina niponensis*, *Cheimatobia brumata*, *Chilo* spp., *Choristoneura* spp.,  
 20 *Clysia ambiguella*, *Cnaphalocerus* spp., *Cnaphalocrocis medinalis*, *Cnephasia* spp., *Conopomorpha* spp., *Conotrachelus* spp., *Copitarsia* spp., *Cydia* spp., *Dalaca noctuides*, *Diaphania* spp., *Diatraea saccharalis*, *Earias* spp., *Ecdytolopha aurantium*, *Elasmopalpus lignosellus*, *Eldana saccharina*, *Ephestia* spp., *Epinotia* spp., *Epiphyas postvittana*, *Etiella* spp., *Eulia* spp., *Eupoecilia ambiguella*, *Euproctis* spp., *Euxoa* spp., *Feltia* spp., *Galleria mellonella*, *Gracillaria* spp., *Grapholitha* spp., *Hedylepta* spp., *Helicoverpa* spp.,  
 25 *Heliothis* spp., *Hofmannophila pseudospretella*, *Homoeosoma* spp., *Homona* spp., *Hyponomeuta padella*, *Kakivoria flavofasciata*, *Laphygma* spp., *Laspeyresia molesta*, *Leucinodes orbonalis*, *Leucoptera* spp., *Lithocolletis* spp., *Lithophane antennata*, *Lobesia* spp., *Loxagrotis albicosta*, *Lymantria* spp., *Lyonetia* spp., *Malacosoma neustria*, *Maruca testulalis*, *Mamstra brassicae*, *Melanitis leda*, *Mocis* spp., *Monopis obviella*, *Mythimna separata*, *Nemapogon cloacellus*, *Nymphula* spp., *Oiketicus* spp., *Oria* spp., *Orthaga* spp., *Ostrinia* spp., *Oulema oryzae*, *Panolis flammea*, *Parnara* spp., *Pectinophora* spp., *Perileucoptera* spp., *Phthorimaea* spp., *Phyllocnistis citrella*, *Phyllonorycter* spp., *Pieris* spp., *Platynota stultana*, *Plodia interpunctella*, *Plusia* spp., *Plutella xylostella*, *Prays* spp., *Prodenia* spp., *Protoparce* spp., *Pseudaletia* spp., *Pseudaletia unipuncta*, *Pseudoplusia includens*, *Pyrausta nubilalis*, *Rachiplusia nu*, *Schoenobius* spp., *Scirpophaga* spp., *Scirpophaga innotata*, *Scotia segetum*, *Sesamia* spp., *Sesamia inferens*,  
 35 *Sparganothis* spp., *Spodoptera* spp., *Spodoptera praefica*, *Stathmopoda* spp., *Stomopteryx subsecivella*, *Synanthedon* spp., *Tecia solanivora*, *Thermesia gemmatilis*, *Tinea cloacella*, *Tinea pellionella*, *Tineola bisselliella*, *Tortrix* spp., *Trichophaga tapetzella*, *Trichoplusia* spp., *Tryporyza incertulas*, *Tuta absoluta*, *Virachola* spp.

In some instances, the insect is from the order Orthoptera or Saltatoria, for example, *Acheta domesticus*, *Dichroplus* spp., *Gryllotalpa* spp., *Hieroglyphus* spp., *Locusta* spp., *Melanoplus* spp., *Schistocerca gregaria*.

In some instances, the insect is from the order Phthiraptera, for example, *Damalinia* spp.,  
5 *Haematopinus* spp., *Linognathus* spp., *Pediculus* spp., *Ptirus pubis*, *Trichodectes* spp.

In some instances, the insect is from the order Psocoptera for example *Lepinatus* spp., *Liposcelis* spp.

In some instances, the insect is from the order Siphonaptera, for example, *Ceratophyllus* spp., *Ctenocephalides* spp., *Pulex irritans*, *Tunga penetrans*, *Xenopsylla cheopsis*.

10 In some instances, the insect is from the order Thysanoptera, for example, *Anaphothrips obscurus*, *Baliothrips biformis*, *Drepanothrips reuteri*, *Enneothrips flavens*, *Frankliniella* spp., *Heliothrips* spp., *Hercinothrips femoralis*, *Rhipiphorothrips cruentatus*, *Scirtothrips* spp., *Taeniothrips cardamomi*, *Thrips* spp.

In some instances, the insect is from the order Zygentoma (=Thysanura), for example,  
15 *Ctenolepisma* spp., *Lepisma saccharina*, *Lepismodes inquilinus*, *Thermobia domestica*.

In some instances, the insect is from the class Symphyla, for example, *Scutigerella* spp.

In some instances, the insect is a mite, including but not limited to, Tarsonemid mites, such as *Phytonemus pallidus*, *Polyphagotarsonemus latus*, *Tarsonemus bilobatus*, or the like; Eupodid mites, such as *Penthaleus erythrocephalus*, *Penthaleus major*, or the like; Spider mites, such as *Oligonychus*  
20 *shinkajii*, *Panonychus citri*, *Panonychus mori*, *Panonychus ulmi*, *Tetranychus kanzawai*, *Tetranychus urticae*, or the like; Eriophyid mites, such as *Acaphylla theavagrans*, *Aceria tulipae*, *Aculops lycopersici*, *Aculops pelekassi*, *Aculus schlechtendali*, *Eriophyes chibaensis*, *Phyllocoptruta oleivora*, or the like; Acarid mites, such as *Rhizoglyphus robini*, *Tyrophagus putrescentiae*, *Tyrophagus similis*, or the like; Bee brood mites, such as *Varroa jacobsoni*, *Varroa destructor* or the like; Ixodides, such as *Boophilus*  
25 *microplus*, *Rhipicephalus sanguineus*, *Haemaphysalis longicornis*, *Haemaphysalis flava*, *Haemaphysalis campanulata*, *Ixodes ovatus*, *Ixodes persulcatus*, *Amblyomma* spp., *Dermacentor* spp., or the like; Cheyletidae, such as *Cheyletiella yasguri*, *Cheyletiella blakei*, or the like; Demodicidae, such as *Demodex canis*, *Demodex cati*, or the like; Psoroptidae, such as *Psoroptes ovis*, or the like; Sarcoptidae, such as *Sarcoptes scabiei*, *Notoedres cati*, *Knemidocoptes* spp., or the like.

30 The methods and compositions provided herein may be used with any insect host that is considered a vector for a pathogen that is capable of causing disease in animals. For example, the insect host may include, but is not limited to those with piercing-sucking mouthparts, as found in Hemiptera and some Hymenoptera and Diptera such as mosquitoes, bees, wasps, midges, lice, tsetse fly, fleas and ants, as well as members of the Arachnidae such as ticks and mites; order, class or family of Acarina  
35 (ticks and mites) e.g. representatives of the families Argasidae, Dermanyssidae, Ixodidae, Psoroptidae or Sarcoptidae and representatives of the species *Amblyomma* spp., *Anocenton* spp., *Argas* spp., *Boophilus* spp., *Cheyletiella* spp., *Chorioptes* spp., *Demodex* spp., *Dermacentor* spp., *Dermanyssus* spp., *Haemaphysalis* spp., *Hyalomma* spp., *Ixodes* spp., *Lynxacarus* spp., *Mesostigmata* spp., *Notoedres* spp., *Ornithodoros* spp., *Ornithonyssus* spp., *Otobius* spp., *otodectes* spp., *Pneumonyssus* spp.,  
40 *Psoroptes* spp., *Rhipicephalus* spp., *Sarcoptes* spp., or *Trombicula* spp.; Anoplura (sucking and biting

lice) e.g. representatives of the species *Bovicola* spp., *Haematopinus* spp., *Linognathus* spp., *Menopon* spp., *Pediculus* spp., *Pemphigus* spp., *Phylloxera* spp., or *Solenopotes* spp.; Diptera (flies) e.g. representatives of the species *Aedes* spp., *Anopheles* spp., *Calliphora* spp., *Chrysomyia* spp., *Chrysops* spp., *Cochliomyia* spp., *Cw/ex* spp., *Culicoides* spp., *Cuterebra* spp., *Dermatobia* spp., *Gastrophilus* spp.,  
5 *Glossina* spp., *Haematobia* spp., *Haematopota* spp., *Hippobosca* spp., *Hypoderma* spp., *Lucilia* spp., *Lyperosia* spp., *Melophagus* spp., *Oestrus* spp., *Phaenicia* spp., *Phlebotomus* spp., *Phormia* spp., *Sarcophaga* spp., *Simulium* spp., *Stomoxys* spp., *Tabanus* spp., *Tannia* spp. or *Zzpu/alpha* spp.; Mallophaga (biting lice) e.g. representatives of the species *Damalina* spp., *Felicola* spp., *Heterodoxus* spp. or *Trichodectes* spp.; or Siphonaptera (wingless insects) e.g. representatives of the species  
10 *Ceratophyllus* spp., *Xenopsylla* spp.; Cimicidae (true bugs) e.g. representatives of the species *Cimex* spp., *Tritominae* spp., *Rhodnius* spp., or *Triatoma* spp. In some instances, the insect is a blood-sucking insect from the order Diptera (e.g., suborder Nematocera, e.g., family Colicidae). In some instances, the insect is from the subfamilies Culicinae, Corethrinae, Ceratopogonidae, or Simuliidae. In some instances, the insect is of a *Culex* spp., *Theobaldia* spp., *Aedes* spp., *Anopheles* spp., *Aedes* spp., *Forcipomyia* spp.,  
15 *Culicoides* spp., or *Helea* spp.

## ii. Mollusk hosts

In some instances, the host described herein may be an organism belonging to the phylum Mollusca. In some instances, the mollusk is considered a pest, e.g., an agricultural pest. For example,  
20 the methods and compositions are suitable for controlling terrestrial Gastropods (e.g., slugs and snails) in agriculture and horticulture. They include all terrestrial slugs and snails which mostly occur as polyphagous pests on agricultural and horticultural crops.

In some instances, the mollusk belongs to the family Achatinidae, Agriolimacidae, Ampullariidae, Arionidae, Bradybaenidae, Helicidae, Hydromiidae, Lymnaeidae, Milacidae, Urocyclidae, or  
25 Veronicellidae.

For example, in some instances, the mollusk is *Achatina* spp., *Agriolimax* spp., *Arion* spp. (e.g., *A. ater*, *A. circumscriptus*, *A. distinctus*, *A. fasciatus*, *A. hortensis*, *A. intermedius*, *A. rufus*, *A. subfuscus*, *A. silvaticus*, *A. lusitanicus*), *Biomphalaria* spp., *Bradybaena* spp. (e.g., *B. fruticum*), *Bulinus* spp., *Cantareus* spp. (e.g., *C. asperses*), *Cepaea* spp. (e.g., *C. hortensis*, *C. nemoralis*), *Cernuella* spp.,  
30 *Cochlicella* spp., *Cochlodina* spp. (e.g., *C. laminata*), *Deroceras* spp. (e.g., *D. agrestis*, *D. empiricum*, *D. laeve*, *D. panormitatum*, *D. reticulatum*), *Discus* spp. (e.g., *D. rotundatus*), *Euomphalia* spp., *Galba* spp. (e.g., *G. trunculata*), *Helicella* spp. (e.g., *H. itala*, *H. obvia*), *Helicigona* spp. (e.g., *H. arbustorum*), *Helicodiscus* spp., *Helix* spp. (e.g., *H. aperta*, *H. aspersa*, *H. pomatia*), *Limax* spp. (e.g., *L. cinereoniger*, *L. flavus*, *L. marginatus*, *L. maximus*, *L. tenellus*), *Lymnaea* spp. (e.g., *L. stagnalis*), *Milax* spp. (e.g., *M. gagates*, *M. marginatus*, *M. sowerbyi*, *M. budapestensis*), *Oncomelania* spp., *Opeas* spp., *Oxyloma* spp. (e.g., *O. pfeifferi*), *Pomacea* spp. (e.g., *P. canaliculata*), *Succinea* spp., *Tandonia* spp. (e.g., *T. budapestensis*, *T. sowerbyi*), *Theba* spp., *Vallonia* spp., and *Zonitoides* spp. (e.g., *Z. nitidus*).

### iii. Nematode hosts

The host of any of the compositions or methods described herein may also be any organism belonging to the phylum Nematoda. In some instances, the nematode is considered a pest, e.g., an agricultural pest. For example, the nematode may be parasitic or cause health problems to plant or to fungi (for example species of the orders Aphelenchida, Meloidogyne, Tylenchida and others) or to humans and animals (for example species of the orders Trichinellida, Tylenchida, Rhabditina, and Spirurida).

Plant nematodes encompass plant parasitic nematodes and nematodes living in the soil. Plant parasitic nematodes include, but are not limited to, ectoparasites such as *Xiphinema* spp., *Longidorus* spp., and *Trichodorus* spp.; semiparasites such as *Tylenchulus* spp.; migratory endoparasites such as *Pratylenchus* spp., *Radopholus* spp., and *Scutellonema* spp.; sedentary parasites such as *Heterodera* spp., *Globodera* spp., and *Meloidogyne* spp., and stem and leaf endoparasites such as *Ditylenchus* spp., *Aphelenchoides* spp., and *Hirschmaniella* spp. Especially harmful root parasitic soil nematodes are such as cystforming nematodes of the genera *Heterodera* or *Globodera*, and/or root knot nematodes of the genus *Meloidogyne*. Harmful species of these genera are for example *Meloidogyne incognita*, *Heterodera glycines* (soybean cyst nematode), *Globodera pallida* and *Globodera rostochiensis* (potato cyst nematode), which species are effectively controlled with the modulating agents described herein. However, the use of the modulating agents described herein is in no way restricted to these genera or species, but also extends in the same manner to other nematodes.

Plant nematodes include but are not limited to e.g. *Aglenchus agricola*, *Anguina tritici*, *Aphelenchoides arachidis*, *Aphelenchoides fragaria* and the stem and leaf endoparasites *Aphelenchoides* spp. in general, *Belonolaimus gracilis*, *Belonolaimus longicaudatus*, *Belonolaimus nortoni*, *Bursaphelenchus cocophilus*, *Bursaphelenchus eremus*, *Bursaphelenchus xylophilus*, *Bursaphelenchus mucronatus*, and *Bursaphelenchus* spp. in general, *Cacopaurus pestis*, *Criconemella curvata*, *Criconemella onoensis*, *Criconemella ornata*, *Criconemella rusium*, *Criconemella xenoplax* (= *Mesocriconema xenoplax*) and *Criconemella* spp. in general, *Criconemoides femiae*, *Criconemoides onoense*, *Criconemoides ornatum* and *Criconemoides* spp. in general, *Ditylenchus destructor*, *Ditylenchus dipsaci*, *Ditylenchus myceliophagus* and the stem and leaf endoparasites *Ditylenchus* spp. in general, *Dolichodorus heterocephalus*, *Globodera pallida* (= *Heterodera pallida*), *Globodera rostochiensis* (potato cyst nematode), *Globodera solanacearum*, *Globodera tabacum*, *Globodera virginia* and the sedentary, cyst forming parasites *Globodera* spp. in general, *Helicotylenchus digonicus*, *Helicotylenchus dihystra*, *Helicotylenchus erythrone*, *Helicotylenchus multicinctus*, *Helicotylenchus nannus*, *Helicotylenchus pseudorobustus* and *Helicotylenchus* spp. in general, *Hemicriconemoides*, *Hemicycliophora arenaria*, *Hemicycliophora nudata*, *Hemicycliophora parvana*, *Heterodera avenae*, *Heterodera cruciferae*, *Heterodera glycines* (soybean cyst nematode), *Heterodera oryzae*, *Heterodera schachtii*, *Heterodera zae* and the sedentary, cyst forming parasites *Heterodera* spp. in general, *Hirschmaniella gracilis*, *Hirschmaniella oryzae*, *Hirschmaniella spinicaudata* and the stem and leaf endoparasites *Hirschmaniella* spp. in general, *Hoplolaimus aegyptii*, *Hoplolaimus californicus*, *Hoplolaimus columbus*, *Hoplolaimus galeatus*, *Hoplolaimus indicus*, *Hoplolaimus magnistylus*, *Hoplolaimus pararobustus*, *Longidorus africanus*, *Longidorus breviannulatus*, *Longidorus elongatus*,

*Longidorus laevicapitatus*, *Longidorus vineacola* and the ectoparasites *Longidorus* spp. in general, *Meloidogyne acronea*, *Meloidogyne africana*, *Meloidogyne arenaria*, *Meloidogyne arenaria thamesi*, *Meloidogyne artiella*, *Meloidogyne chitwoodi*, *Meloidogyne coffeicola*, *Meloidogyne ethiopica*, *Meloidogyne exigua*, *Meloidogyne fallax*, *Meloidogyne graminicola*, *Meloidogyne graminis*, *Meloidogyne*  
5 *hapla*, *Meloidogyne incognita*, *Meloidogyne incognita acrita*, *Meloidogyne javanica*, *Meloidogyne kikuyensis*, *Meloidogyne minor*, *Meloidogyne naasi*, *Meloidogyne paranaensis*, *Meloidogyne thamesi* and the sedentary parasites *Meloidogyne* spp. in general, *Meloinema* spp., *Nacobbus aberrans*, *Neotylenchus vigissi*, *Paraphelenchus pseudoparietinus*, *Paratrichodorus allius*, *Paratrichodorus lobatus*, *Paratrichodorus minor*, *Paratrichodorus nanus*, *Paratrichodorus porosus*, *Paratrichodorus teres* and  
10 *Paratrichodorus* spp. in general, *Paratylenchus hamatus*, *Paratylenchus minutus*, *Paratylenchus projectus* and *Paratylenchus* spp. in general, *Pratylenchus agilis*, *Pratylenchus allenii*, *Pratylenchus andinus*, *Pratylenchus brachyurus*, *Pratylenchus cerealis*, *Pratylenchus coffeae*, *Pratylenchus crenatus*, *Pratylenchus delattrei*, *Pratylenchus giibbicaudatus*, *Pratylenchus goodeyi*, *Pratylenchus hamatus*, *Pratylenchus hexincisus*, *Pratylenchus loosi*, *Pratylenchus neglectus*, *Pratylenchus penetrans*,  
15 *Pratylenchus pratensis*, *Pratylenchus scribneri*, *Pratylenchus teres*, *Pratylenchus thornei*, *Pratylenchus vulnus*, *Pratylenchus zaeae* and the migratory endoparasites *Pratylenchus* spp. in general, *Pseudohalenchus minutus*, *Psilenchus magnidens*, *Psilenchus tumidus*, *Punctodera chalconensis*, *Quinisulcius acutus*, *Radopholus citrophilus*, *Radopholus similis*, the migratory endoparasites *Radopholus* spp. in general, *Rotylenchulus borealis*, *Rotylenchulus parvus*, *Rotylenchulus reniformis* and  
20 *Rotylenchulus* spp. in general, *Rotylenchus laurentinus*, *Rotylenchus macrodoratus*, *Rotylenchus robustus*, *Rotylenchus uniformis* and *Rotylenchus* spp. in general, *Scutellonema brachyurum*, *Scutellonema bradys*, *Scutellonema clathricaudatum* and the migratory endoparasites *Scutellonema* spp. in general, *Subanguina radiciola*, *Tetylenchus nicotianae*, *Trichodorus cylindricus*, *Trichodorus minor*, *Trichodorus primitivus*, *Trichodorus proximus*, *Trichodorus similis*, *Trichodorus sparsus* and the  
25 ectoparasites *Trichodorus* spp. in general, *Tylenchorhynchus agri*, *Tylenchorhynchus brassicae*, *Tylenchorhynchus clarus*, *Tylenchorhynchus claytoni*, *Tylenchorhynchus digitatus*, *Tylenchorhynchus ebriensis*, *Tylenchorhynchus maximus*, *Tylenchorhynchus nudus*, *Tylenchorhynchus vulgaris* and *Tylenchorhynchus* spp. in general, *Tylenchulus semipenetrans* and the semiparasites *Tylenchulus* spp. in general, *Xiphinema americanum*, *Xiphinema brevicolle*, *Xiphinema dimorphicaudatum*, *Xiphinema index*  
30 and the ectoparasites *Xiphinema* spp. in general.

Other examples of nematode hosts include species belonging to the family *Criconematidae*, *Belonolaimidae*, *Hoploaimidae*, *Heteroderidae*, *Longidoridae*, *Pratylenchidae*, *Trichodoridae*, or *Anguinidae*.

#### 35 iv. Beneficial hosts

In some instances, the host described herein is a beneficial insect, mollusk, or nematode (e.g., a pollinator, a natural competitor of a pest, or a producer of useful substances for humans). The term "beneficial insect," "beneficial mollusk," or "beneficial nematode," as used herein, refers to an insect, mollusk, or nematode that confers a benefit (e.g., economical and/or ecological) to humans, animals, an  
 40 ecosystem, and/or the environment. For example, the host may be an invertebrate (e.g., insect, mollusk,

or nematode) that is involved in the production of a commercial product, including, but not limited to, invertebrates cultivated to produce food (e.g., honey from honey bees, e.g., *Apis mellifera*), materials (such as silk from *Bombyx mori*), and/or substances (e.g., lac from *Laccifer lacca* or pigments from *Dactylopius coccus* and *Cynipidae*). Additionally, the host may include invertebrates (e.g., insects, mollusks, or nematodes) that are used in agricultural applications, including invertebrates (e.g., insects, mollusks, or nematodes) that aid in the pollination of crops, spreading seeds, or pest control. Further, in some instances, the host may be an invertebrate (e.g., insect, mollusk, or nematode) that is useful for waste disposal and/or organic recycling (e.g., earthworms, termites, or *Diptera* larvae).

In some instances, the host produces a useable product (e.g., honey, silk, beeswax, or shellac).

In some instances, the host is a bee. Exemplary bee genera include, but are not limited to *Apis*, *Bombus*, *Trigona*, and *Osmia*. In some instances, the bee is a honeybee (e.g., an insect belonging to the genus *Apis*). In some instances, the honeybee is the species *Apis mellifera* (the European or Western honey bee), *Apis cerana* (the Asiatic, Eastern, or Himalayan honey bee), *Apis dorsata* (the “giant” honey bee), *Apis florea* (the “red dwarf” honey bee), *Apis andreniformis* (the “black dwarf” honey bee), or *Apis nigrocincta*. In some instances, the host is a silkworm. The silkworm may be a species in the family *Bombycidae* or *Saturniidae*. In some instances, the silkworm is *Bombyx mori*. In some instances, the host is a lac bug. The lac bug may be a species in the family *Kerriidae*. In some instances, the lac bug is *Kerria lacca*.

In some instances, the host aids in pollination of a plant (e.g., bees, beetles, wasps, flies, butterflies, or moths). In some examples, the host aiding in pollination of a plant is beetle. In some instances, the beetle is a species in the family *Buprestidae*, *Cantharidae*, *Cerambycidae*, *Chrysomelidae*, *Cleridae*, *Coccinellidae*, *Elateridae*, *Melandryidae*, *Meloidae*, *Melyridae*, *Mordellidae*, *Nitidulidae*, *Oedemeridae*, *Scarabaeidae*, or *Staphyllinidae*. In some instances, the host aiding in pollination of a plant is a butterfly or moth (e.g., *Lepidoptera*). In some instances, the butterfly or moth is a species in the family *Geometridae*, *Hesperiidae*, *Lycaenidae*, *Noctuidae*, *Nymphalidae*, *Papilionidae*, *Pieridae*, or *Sphingidae*. In some instances, the host aiding in pollination of a plant is a fly (e.g., *Diptera*). In some instances, the fly is in the family *Anthomyiidae*, *Bibionidae*, *Bombyliidae*, *Calliphoridae*, *Cecidomyiidae*, *Certopogonidae*, *Chrionomidae*, *Conopidae*, *Culicidae*, *Dolichopodidae*, *Empididae*, *Ephydriidae*, *Lonchopteridae*, *Muscidae*, *Mycetophilidae*, *Phoridae*, *Simuliidae*, *Stratiomyidae*, or *Syrphidae*. In some instances, the host aiding in pollination is an ant (e.g., *Formicidae*), sawfly (e.g., *Tenthredinidae*), or wasp (e.g., *Sphecidae* or *Vespidae*). In some instances, the host aiding in pollination of a plant is a bee. In some instances, the bee is in the family *Andrenidae*, *Apidae*, *Colletidae*, *Halictidae*, or *Megachilidae*.

In some instances, the host aids in pest control. In some instances, the host aiding in pest control is a predatory nematode. In particular examples, the nematode is a species of *Heterorhabditis* or *Steinernema*. In some instances, the host aiding in pest control is an insect. For example, the host aiding in pest control may be a species belonging to the family *Braconidae* (e.g., parasitoid wasps), *Carabidae* (e.g., ground beetles), *Chrysopidae* (e.g., green lacewings), *Coccinellidae* (e.g., ladybugs), *Hemerobiidae* (e.g., brown lacewings), *Ichneumonidae* (e.g., ichneumon wasps), *Lampyridae* (e.g., fireflies), *Mantidae* (e.g., praying mantises), *Myrmeleontidae* (e.g., antlions), *Odonata* (e.g., dragonflies and damselflies), or *Syrphidae* (e.g., hoverfly). In other instances, the host aiding in pest control is an

insect that competes with an insect that is considered a pest (e.g., an agricultural pest). For example, the Mediterranean fruit fly, *Ceratitis capitata* is a common pest of fruits and vegetables worldwide. One way to control *C. capitata* is to release the sterilized male insect into the environment to compete with wild males to mate the females. In these instances, the host may be a sterilized male belonging to a species that is typically considered a pest.

In some instances, the host aids in degradation of waste or organic material. In some examples, the host aiding in degradation of waste or organic material belongs to *Coleoptera* or *Diptera*. In some instances, the host belonging to *Diptera* is in the family *Calliphoridae*, *Curtonotidae*, *Drosophilidae*, *Fanniidae*, *Heleomyzidae*, *Milichiidae*, *Muscidae*, *Phoridae*, *Psychodidae*, *Scatopsidae*, *Sepsidae*, *Sphaeroceridae*, *Stratiomyidae*, *Syrphidae*, *Tephritidae*, or *Ulidiidae*. In some instances, the host belonging to *Coleoptera* is in the family *Carabidae*, *Hydrophilidae*, *Phalacaridae*, *Ptiliidae*, or *Staphylinidae*.

In some instances, the host may be an insect or an arachnid that may be cultivated for a consumable product (e.g., food or feed). For example, the host may be a moth, butterfly, fly, cricket, spider, or beetle. In some instances, the host is in the order *Anoplura*, *Araneae*, *Blattodea*, *Coleoptera*, *Dermaptera*, *Dictyoptera*, *Diplura*, *Diptera*, *Embioptera*, *Ephemeroptera*, *Grylloblatodea*, *Hemiptera*, *Homoptera*, *Hymenoptera*, *Isoptera*, *Lepidoptera*, *Mantodea*, *Mecoptera*, *Neuroptera*, *Odonata*, *Orthoptera*, *Phasmida*, *Plecoptera*, *Protura*, *Psocoptera*, *Siphonaptera*, *Siphunculata*, *Thysanura*, *Strepsiptera*, *Thysanoptera*, *Trichoptera*, or *Zoraptera*.

In some examples, the host is a black soldier fly (*Hermetia illucens*), a common house fly, a lesser mealworm, a weaver ant, a silkworm (*Bombyx mori*), a grasshopper, a Chinese grasshopper (*Acrida cinerea*), a yellow mealworm (*Clarias gariepinns*), a moth (*Anaphe infracta* or *Bombyx mori*), *Spodoptera littoralis*, a house cricket, a termite, a palm weevil (*Rhynchophorus ferrugineus*), a giant water bug (*Lethocerus indicus*), a water beetle, a termite (*Macrotermes subhyalinus*), a drugstore beetle (*Stegobium paniceum*), *Imbrasia belina*, *Rhynchophorus phoenicis*, *Oryctes rhinoceros*, *Macrotermes bellicosus*, *Ruspolia differens*, *Oryctes Monoceros*, or *Oecophylla smaragdina*.

#### v. Decreasing host fitness

The methods and compositions provided herein may be used to decrease the fitness of any of the host invertebrates (e.g., insects, mollusks, or nematodes) described herein. The decrease in fitness arises from alterations in host pathways that mediate interactions between the host and microorganisms resident in the host, wherein the alterations are a consequence of administration of a modulating agent and have detrimental effects on the host.

In some instances, the decrease in host fitness may manifest as a deterioration or decline in the physiology of the host (e.g., reduced health or survival) as a consequence of administration of a modulating agent. In some instances, the fitness of an organism may be measured by one or more parameters, including, but not limited to, reproductive rate, fertility, lifespan, viability, mobility, fecundity, host development, body weight, metabolic rate or activity, or survival in comparison to a host organism to which the modulating agent has not been administered. For example, the methods or compositions provided herein may be effective to decrease the overall health of the host or to decrease the overall

survival of the host. In some instances, the decreased survival of the host is about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% greater relative to a reference level (e.g., a level found in a host that does not receive a modulating agent). In some instances, the methods and compositions are effective to decrease host reproduction (e.g., reproductive rate, fertility) in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods and compositions are effective to decrease other physiological parameters, such as mobility, body weight, life span, fecundity, or metabolic rate, by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent).

In some instances, the decrease in host fitness may manifest as a decrease in the production of one or more nutrients in the host (e.g., vitamins, carbohydrates, amino acids, or polypeptides) in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods or compositions provided herein may be effective to decrease the production of nutrients in the host (e.g., vitamins, carbohydrates, amino acids, or polypeptides) by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent). In some instances, the methods or compositions provided herein may decrease nutrients in the host by decreasing the production of nutrients by one or more microorganisms (e.g., endosymbiont) in the host in comparison to a host organism to which the modulating agent has not been administered.

In some instances, the decrease in host fitness may manifest as an increase in the host's sensitivity to a pesticidal agent (e.g., a pesticide listed in Table 11) and/or a decrease in the host's resistance to a pesticidal agent (e.g., a pesticide listed in Table 11) in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods or compositions provided herein may be effective to increase the host's sensitivity to a pesticidal agent (e.g., a pesticide listed in Table 11) by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent). The pesticidal agent may be any pesticidal agent known in the art, including insecticidal agents. In some instances, the methods or compositions provided herein may increase the host's sensitivity to a pesticidal agent (e.g., a pesticide listed in Table 11) by decreasing the host's ability to metabolize or degrade the pesticidal agent into usable substrates in comparison to a host organism to which the modulating agent has not been administered.

In some instances, the decrease in host fitness may manifest as an increase in the host's sensitivity to an allelochemical agent and/or a decrease in the host's resistance to an allelochemical agent in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods or compositions provided herein may be effective to decrease the host's resistance to an allelochemical agent by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent). In some instances, the allelochemical agent is caffeine, soyacystatin, fenitrothion, monoterpenes, diterpene acids, or phenolic compounds (e.g., tannins, flavonoids). In some instances, the methods or compositions provided herein may increase the host's sensitivity to an

allelochemical agent by decreasing the host's ability to metabolize or degrade the allelochemical agent into usable substrates in comparison to a host organism to which the modulating agent has not been administered.

In some instances, the methods or compositions provided herein may be effective to decrease the host's resistance to parasites or pathogens (e.g., fungal, bacterial, or viral pathogens or parasites) in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods or compositions provided herein may be effective to decrease the host's resistance to a pathogen or parasite (e.g., fungal, bacterial, or viral pathogens; or parasitic mites) by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent).

In some instances, the methods or compositions provided herein may be effective to decrease the host's ability to carry or transmit a plant pathogen (e.g., plant virus (e.g., TYLCV) or a plant bacterium (e.g., *Agrobacterium* spp)) in comparison to a host organism to which the modulating agent has not been administered. For example, the methods or compositions provided herein may be effective to decrease the host's ability to carry or transmit a plant pathogen (e.g., a plant virus (e.g., TYLCV) or plant bacterium (e.g., *Agrobacterium* spp)) by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent).

In some instances, the decrease in host fitness may manifest as other fitness disadvantages, such as a decreased tolerance to certain environmental factors (e.g., a high or low temperature tolerance), a decreased ability to survive in certain habitats, or a decreased ability to sustain a certain diet in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods or compositions provided herein may be effective to decrease host fitness in any plurality of ways described herein. Further, the modulating agent may decrease host fitness in any number of host classes, orders, families, genera, or species (e.g., 1 host species, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 500, or more host species). In some instances, the modulating agent acts on a single host class, order, family, genus, or species.

Host fitness may be evaluated using any standard methods in the art. In some instances, host fitness may be evaluated by assessing an individual host. Alternatively, host fitness may be evaluated by assessing a host population. For example, a decrease in host fitness may manifest as a decrease in successful competition against other insects, thereby leading to a decrease in the size of the host population.

#### *vi. Increasing host fitness*

The methods and compositions provided herein may be used to increase the fitness of any of the hosts described herein. The increase in fitness arises from alterations in host pathways that mediate interactions between the host and microorganisms resident in the host, wherein the alterations are a consequence of administration of a modulating agent and have beneficial effects on the host.

In some instances, the increase in host fitness may manifest as an improvement in the physiology of the host (e.g., improved health or survival) as a consequence of administration of a

modulating agent. In some instances, the fitness of an organism may be measured by one or more parameters, including, but not limited to, reproductive rate, lifespan, mobility, fecundity, body weight, metabolic rate or activity, or survival in comparison to a host organism to which the modulating agent has not been administered. For example, the methods or compositions provided herein may be effective to

5 improve the overall health of the host or to improve the overall survival of the host in comparison to a host organism to which the modulating agent has not been administered. In some instances, the improved survival of the host is about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% greater relative to a reference level (e.g., a level found in a host that does not receive a modulating agent). In some instances, the methods and compositions are effective to increase host

10 reproduction (e.g., reproductive rate) in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods and compositions are effective to increase other physiological parameters, such as mobility, body weight, life span, fecundity, or metabolic rate, by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent).

15 In some instances, the increase in host fitness may manifest as an increased production of a product generated by said host in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods or compositions provided herein may be effective to increase the production of a product generated by the host, as described herein (e.g., honey, beeswax, beebread, propolis, silk, or lac), by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%,

20 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent).

In some instances, the increase in host fitness may manifest as an increase in the frequency or efficacy of a desired activity carried out by the host (e.g., pollination, predation on pests, seed spreading, or breakdown of waste or organic material) in comparison to a host organism to which the modulating

25 agent has not been administered. In some instances, the methods or compositions provided herein may be effective to increase the frequency or efficacy of a desired activity carried out by the host (e.g., pollination, predation on pests, seed spreading, or breakdown of waste or organic material) by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent).

30 In some instances, the increase in host fitness may manifest as an increase in the production of one or more nutrients in the host (e.g., vitamins, carbohydrates, amino acids, or polypeptides) in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods or compositions provided herein may be effective to increase the production of nutrients in the host (e.g., vitamins, carbohydrates, amino acids, or polypeptides) by about 2%, 5%, 10%,

35 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent). In some instances, the methods or compositions provided herein may increase nutrients in the host by increasing the production of nutrients by one or more microorganisms (e.g., endosymbiont) in the host in comparison to a host organism to which the modulating agent has not been administered.

In some instances, the increase in host fitness may manifest as a decrease in the host's sensitivity to a pesticidal agent (e.g., a pesticide listed in Table 11) and/or an increase in the host's resistance to a pesticidal agent (e.g., a pesticide listed in Table 11) in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods or compositions provided herein may be effective to decrease the host's sensitivity to a pesticidal agent (e.g., a pesticide listed in Table 11) by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent). The pesticidal agent may be any pesticidal agent known in the art, including insecticidal agents. In some instances, the pesticidal agent is a neonicotinoid. In some instances, the methods or compositions provided herein may decrease the host's sensitivity to a pesticidal agent (e.g., a pesticide listed in Table 11) by increasing the host's ability to metabolize or degrade the pesticidal agent into usable substrates in comparison to a host organism to which the modulating agent has not been administered.

In some instances, the increase in host fitness may manifest as a decrease in the host's sensitivity to an allelochemical agent and/or an increase in the host's resistance to an allelochemical agent in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods or compositions provided herein may be effective to increase the host's resistance to an allelochemical agent by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent). In some instances, the allelochemical agent is caffeine, soyacystatin, fenitrothion, monoterpenes, diterpene acids, or phenolic compounds (e.g., tannins, flavonoids). In some instances, the methods or compositions provided herein may decrease the host's sensitivity to an allelochemical agent by increasing the host's ability to metabolize or degrade the allelochemical agent into usable substrates in comparison to a host organism to which the modulating agent has not been administered.

In some instances, the methods or compositions provided herein may be effective to increase the host's resistance to parasites or pathogens (e.g., fungal, bacterial, or viral pathogens; or parasitic mites (e.g., *Varroa destructor* mite in honeybees)) in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods or compositions provided herein may be effective to increase the host's resistance to a pathogen or parasite (e.g., fungal, bacterial, or viral pathogens; or parasitic mites (e.g., *Varroa destructor* mite in honeybees)) by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent).

In some instances, the increase in host fitness may manifest as other fitness advantages, such as improved tolerance to certain environmental factors (e.g., a high or low temperature tolerance), improved ability to survive in certain habitats, or an improved ability to sustain a certain diet (e.g., an improved ability to metabolize soy vs corn) in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods or compositions provided herein may be effective to increase host fitness in any plurality of ways described herein. Further, the modulating agent may increase host fitness in any number of host classes, orders, families, genera, or species (e.g.,

1 host species, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 200, 250, 500, or more host species). In some instances, the modulating agent acts on a single host class, order, family, genus, or species.

Host fitness may be evaluated using any standard methods in the art. In some instances, host fitness may be evaluated by assessing an individual host. Alternatively, host fitness may be evaluated by assessing a host population. For example, an increase in host fitness may manifest as an increase in successful competition against other insects, thereby leading to an increase in the size of the host population.

#### *vii. Hosts in agriculture*

The modulating agents described herein may be useful to promote the growth of plants. For example, by reducing the fitness of harmful invertebrates (e.g., insects, mollusks, or nematodes), the modulating agents provided herein may be effective to promote the growth of plants that are typically harmed by a host. Alternatively, by increasing the fitness of beneficial invertebrates (e.g., insects, mollusks, or nematodes), the modulating agents provided herein may be effective to promote the growth of plants that benefit from said hosts. The modulating agent may be delivered to the plant using any of the formulations and delivery methods described herein, in an amount and for a duration effective to modulate (e.g., increase or decrease) host fitness and thereby benefit the plant, e.g., increase crop growth, increase crop yield, decrease pest infestation, and/or decrease damage to plants. This may or may not involve direct application of the modulating agent to the plant. For example, in instances where the primary host habitat is different than the region of plant growth, the modulating agent may be applied to either the primary host habitat, the plants of interest, or a combination of both.

In some instances, the plant may be an agricultural food crop, such as a cereal, grain, legume, fruit, or vegetable crop, or a non-food crop, e.g., grasses, flowering plants, cotton, hay, hemp. The compositions described herein may be delivered to the crop any time prior to or after harvesting the cereal, grain, legume, fruit, vegetable, or other crop. Crop yield is a measurement often used for crop plants and is normally measured in metric tons per hectare (or kilograms per hectare). Crop yield can also refer to the actual seed generation from the plant. In some instances, the modulating agent may be effective to increase crop yield (e.g., increase metric tons of cereal, grain, legume, fruit, or vegetable per hectare and/or increase seed generation) by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or more in comparison to a reference level (e.g., a crop to which the modulating agent has not been administered).

In some instances, the plant (e.g., crop) may be at risk of developing a pest infestation (e.g., by an insect, mollusk, or nematode) or may have already developed a pest infestation. The methods and compositions described herein may be used to reduce or prevent pest infestation in such crops by reducing the fitness of invertebrates (e.g., insect, mollusk, or nematode) that infest the plants. In some instances, the modulating agent may be effective to reduce crop infestation (e.g., reduce the number of plants infested, reduce the pest population size, reduce damage to plants) by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or more in comparison to a reference level (e.g., a crop to which the modulating agent has not been administered). In other instances, the modulating agent may be effective to prevent or reduce the likelihood of crop infestation by about 2%, 5%, 10%, 20%, 30%, 40%,

50%, 60%, 70%, 80%, 90%, 100%, or more in comparison to a reference level (e.g., a crop to which the modulating agent has not been administered).

Any suitable plant tissues may benefit from the compositions and methods described herein, including, but not limited to, somatic embryos, pollen, leaves, stems, calli, stolons, microtubers, or shoots.

5 The methods and compositions described herein may include treatment of angiosperm or gymnosperm plants such as acacia, alfalfa, apple, apricot, artichoke, ash tree, asparagus, avocado, banana, barley, beans, beet, birch, beech, blackberry, blueberry, broccoli, brussels sprouts, cabbage, canola, cantaloupe, carrot, cassaya, cauliflower, cedar, a cereal, celery, chestnut, cherry, Chinese cabbage, citrus, clemintine, clover, coffee, corn, cotton, conifers, cowpea, cucumber, cypress, eggplant, elm, endive, eucalyptus, fava  
10 beans, fennel, forage crops, figs, fir, fruit and nut trees, geranium, grape, grapefruit, groundnuts, ground cherry, gum hemlock, hemp, hickory, kale, kiwifruit, kohlrabi, larch, lettuce, leek, lemon, lime, locust, pine, maidenhair, maize, mango, maple, melon, millet, mushroom, mustard, nuts, oak, oats, okra, onion, orange, an ornamental plant or flower or tree, papaya, palm, parsley, parsnip, pea, peach, peanut, pear, peat, pepper, persimmon, pigeon pea, pine, pineapple, plantain, plum, pomegranate, potato, pumpkin,  
15 radicchio, radish, rapeseed, raspberry, rice, rye, sorghum, sallow, soybean, spinach, spruce, squash, strawberry, sugarbeet, sugarcane, sunflower, sweet potato, sweet corn, tangerine, tea, tobacco, tomato, trees, triticale, turf grasses, turnips, a vine, walnut, watercress, watermelon, wheat, yams, yew, or zucchini.

#### 20 *viii. Host in feed/food production*

Upon reaching a desired life stage, the host may be harvested and, if desired, processed for use in the manufacture of a consumable product. In some instances, the harvested invertebrate host (e.g., insect, mollusk, or nematode) may be distributed in a whole form (e.g., as the whole, unprocessed host) as a consumable product. In some instances, the whole harvested host is processed (e.g., ground up)  
25 and distributed as a consumable product. Alternatively, one or more parts of the host (e.g., one or more body parts or one or more substances) may be extracted from the host for use in the manufacture of a consumable product.

The consumable product may be any product safe for human or animal consumption (e.g., ingestion). In some instances, the host may be used in the manufacture of a feed for an animal. In some  
30 instances, the animal is livestock or a farm animal (e.g., chicken, cow, horse, or pig). In some instances, the animal is a bird, reptile, amphibian, mammal, or fish. In some instances, the host may be used in the manufacture of a product that replaces the normal feed of an animal. Alternatively, the host may be used in the manufacture of a product that supplements the normal feed of an animal. The host may also be used in the manufacture of a food, food additive, or food ingredient for humans. In some instances, the  
35 host is used in the manufacture of a nutritional supplement (e.g., protein supplement) for humans.

The host may be a wild or domesticated host. Additionally, the host may be at any developmental stage at the time of delivering or applying the compositions described herein. Further, the host may be at any developmental stage at the time of harvesting the host for use in the manufacture of a consumable product. In some instances, the host is a larva, pupa, or adult insect at the time of

harvesting, using, processing, or manufacturing. The delivery of the modulating agent and the harvesting steps may occur at the same time or different times.

In some instances, a host species is selected based upon their natural nutritional profile. In some instances, the modulating agent is used to improve the nutritional profile of the insect, wherein the modulating agent leads to an increased production of a nutrient in comparison to a host organism to which the modulating agent has not been administered. Examples of nutrients include vitamins, carbohydrates, amino acids, polypeptides, or fatty acids. In some instances, the increased production may arise from increased production of a nutrient by a microorganism resident in the host. Alternatively, the increased production may arise from increased production of a nutrient by the host insect itself, wherein the host has increased fitness following delivery or administration of a modulating agent.

In some instances, in final processing, a first insect species is combined with a second insect species whose nutritional profile provides a complementary benefit to the overall nutritional value of the food or feed product. For example, a species containing a high protein profile could be combined with a species containing a high omega 3/6 fatty acid profile. In this manner, host protein meal may be custom blended to suit the needs of humans or different species of animals.

#### *ix. Host insects in disease transmission*

By decreasing the fitness of host insects that carry human and/or animal pathogens, the modulating agents provided herein may be effective to reduce the spread of vector-borne diseases. The modulating agent may be delivered to the hosts using any of the formulations and delivery methods described herein, in an amount and for a duration effective to reduce transmission of the disease, e.g., reduce vertical or horizontal transmission between vectors and/or reduce transmission to humans and/or animals. For example, the modulating agent described herein may reduce vertical or horizontal transmission of a vector-borne pathogen by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or more in comparison to a host organism to which the modulating agent has not been administered. As an another example, the modulating agent described herein may reduce vectorial competence of an host vector by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or more in comparison to a host organism to which the modulating agent has not been administered.

Non-limiting examples of diseases that may be controlled by the compositions and methods provided herein include diseases caused by Togaviridae viruses (e.g., Chikungunya, Ross River fever, Mayaro, Onyon-nyong fever, Sindbis fever, Eastern equine encephalomyelitis, Western equine encephalomyelitis, Venezuelan equine encephalomyelitis, or Barmah forest); diseases caused by Flaviviridae viruses (e.g., Dengue fever, Yellow fever, Kyasanur Forest disease, Omsk haemorrhagic fever, Japanese encephalitis, Murray Valley encephalitis, Rocio, St. Louis encephalitis, West Nile encephalitis, or Tick-borne encephalitis); diseases caused by Bunyaviridae viruses (e.g., Sandly fever, Rift Valley fever, La Crosse encephalitis, California encephalitis, Crimean-Congo haemorrhagic fever, or Oropouche fever); disease caused by Rhabdoviridae viruses (e.g., Vesicular stomatitis); disease caused by Orbiviridae (e.g., Bluetongue); diseases caused by bacteria (e.g., Plague, Tularemia, Q fever, Rocky Mountain spotted fever, Murine typhus, Boutonneuse fever, Queensland tick typhus, Siberian tick typhus,

Scrub typhus, Relapsing fever, or Lyme disease); or diseases caused by protozoa (e.g., Malaria, African trypanosomiasis, Nagana, Chagas disease, Leishmaniasis, Piroplasmosis, Bancroftian filariasis, or Brugian filariasis).

## 5 II. Target Microorganisms

The microorganisms targeted by the modulating agent described herein may include any microorganism resident in or on an invertebrate host (e.g., insect, mollusk, or nematode), including, but not limited to, any bacteria and/or fungi described herein. Microorganisms resident in the host may include, for example, symbiotic (e.g., endosymbiotic microorganisms that provide beneficial nutrients or enzymes to the host), commensal, pathogenic, or parasitic microorganisms. A symbiotic microorganism (e.g., bacteria or fungi) may be an obligate symbiont of the host or a facultative symbiont of the host. Microorganisms resident in the host may be acquired by any mode of transmission, including vertical, horizontal, or multiple origins of transmission.

### 15 i. Bacteria

Exemplary bacteria that may be targeted in accordance with the methods and compositions provided herein, include, but are not limited to, *Xenorhabdus spp*, *Photorhabdus spp*, *Candidatus spp*, *Buchnera spp*, *Blattabacterium spp*, *Baumania spp*, *Wigglesworthia spp*, *Wolbachia spp*, *Rickettsia spp*, *Orientia spp*, *Sodalis spp*, *Burkholderia spp*, *Cupriavidus spp*, *Frankia spp*, *Snirrhizobium spp*, *Streptococcus spp*, *Wolinella spp*, *Xylella spp* (e.g., *Xylella fastidiosa*), *Erwinia spp*, *Agrobacterium spp*, *Bacillus spp*, *Commensalibacter spp*. (e.g., *Commensalibacter intestini*), *Paenibacillus spp*, *Streptomyces spp*, *Micrococcus spp*, *Corynebacterium spp*, *Acetobacter spp* (e.g., *Acetobacter pomorum*), *Cyanobacteria spp*, *Salmonella spp*, *Rhodococcus spp*, *Pseudomonas spp* (e.g., *Psuedomonas fulva* or *Pseudomonas mandelii*, *Pseudomonas migulae*), *Pantoea spp*. (e.g., *Pantoea vagans*), *Lactobacillus spp* (e.g., *Lactobacillus plantarum*), *Lysobacter spp.*, *Herbaspirillum spp.*, *Enterococcus spp*, *Gluconobacter spp*. (e.g., *Gluconobacter morbifer*), *Alcaligenes spp*, *Hamiltonella spp.*, *Klebsiella spp*, *Paenibacillus spp*, *Serratia spp*. (e.g., *Serratia marcescens*), *Rahnella spp*. (e.g., *Rahnella aquatilis*), *Arthrobacter spp*, *Azotobacter spp.*, *Corynebacterium spp*, *Brevibacterium spp*, *Regiella spp*. (e.g., *Regiella insecticola*), *Thermus spp*, *Pseudomonas spp*, *Clostridium spp*, *Mortierella spp*. (e.g., *Mortierella elongata*) and *Escherichia spp*. In some instances, the targeted bacteria are species in the genera *Xenorhabdus spp.*, *Photorhabdus spp.*, or *Wolbachia spp*. In some instances, the targeted bacteria are species in the order *Streptomycetales*, *Rhizobiales*, *Pseudomonadales*, *Xanthomonadales*, *Sphingobacteriales*, *Chlorofelxales*, *Rhodospirillales*, *Enterobacteriales*, *Sphingomonadales*, *Gemmatimonadales*, *Micrococcales*, *Caulobacterales*, *Cytophagales*, *Firmicutes*, *Micromonosporales*, *Burkholderiales*, *Rickettsiales*, *Flavobacteriales*, *Acidimicroiales*, *Rhodocyclales*, or *Bdellovibrionales*. In some instances, the targeted bacteria are *Armatimonadetes*, *Firmicutes*, *TM7*, *Bacteroidetes*, *Proteobacteria*, or *Actinobacteria*. In some instances, the targeted bacteria are bacteria in the genera *Lactococcus spp.*, *Aeromonas spp.*, *Pseudomonas spp.*, *Enterobacter spp.*, *Citrobacter spp.*, *Sulfurospillium spp.*, *Phaeosphaeria spp.*, or *Mycosphaerella spp*. In some instances, the bacteria targeted by the modulating agent may be ones that can be transmitted from the host (e.g., insect, mollusk, or nematode) to a plant,

including, but not limited to, bacterial plant pathogens (e.g., *Agrobacterium* spp.). Non-limiting examples of bacteria that may be targeted by the methods and compositions provided herein are shown in Table 1. In some instances, the 16S rRNA sequence of the bacteria targeted by the modulating agent has at least 50%, 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 99.9%, or 100% identity with a sequence listed in

5 Table 1.

Table 1: Examples of Target Bacteria and Host Insects

| Primary endosymbiont     | Host                  | Location      | 16S rRNA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|--------------------------|-----------------------|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Gamma proteobacteria     |                       |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| <i>Carsonella ruddii</i> | Psyllids (Psylloidea) | bacteriocytes | TATCCAGCCACAGGTTCCC<br>CTACAGCTACCTTGTTACGA<br>CTTCACCCCAGTTACAAATC<br>ATACCGTTGTAATAGTAAAA<br>TTACTTATGATACAATTTAC<br>TTCCATGGTGTGACGGGCG<br>GTGTGTACAAGGCTCGAGA<br>ACGTATTCACCGTAACATTC<br>TGATTTACGATTACTAGCGA<br>TTCCAAC TTCATGAAATCGA<br>GTTACAGATTTCAATCCGAA<br>CTAAGAATATTTTTTAAGAT<br>TAGCATTATGTTGCCATATA<br>GCATATAACTTTTTGTAATA<br>CTCATTGTAGCACGTGTGT<br>AGCCCTACTTATAAGGGCC<br>ATGATGACTTGACGTCGTC<br>CTCACCTTCCTCCAATTTAT<br>CATTGGCAGTTTCTTATTAG<br>TTCTAATATATTTTTAGTAAA<br>ATAAGATAAGGGTTGCGCT<br>CGTTATAGGACTTAACCCAA<br>CATTTCACAACACGAGCTG<br>ACGACAGCCATGCAGCACC<br>TGTCTCAAAGCTAAAAAAGC<br>TTTATTATTTCTAATAAATTC<br>TTTGGATGTCAAAGTAGGT<br>AAGATTTTTCGTGTTGTATC<br>GAATTAAACCACATGCTCCA<br>CCGCTTGTGCGAGCCCCCG<br>TCAATTCATTTGAGTTTTAA |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|--|--|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | CCTTGCGGTCGTAATCCCC<br>AGGCGGTCAACTTAACGCG<br>TTAGCTTTTTTCACTAAAAAT<br>ATATAACTTTTTTTTCATAAAA<br>CAAAATTACAATTATAATATT<br>TAATAAATAGTTGACATCGT<br>TTRACTGCATGGACTACCAG<br>GGTATCTAATCCTGTTTGCT<br>CCCCATGCTTTCGTGTATTA<br>GTGTCAGTATTAAAATAGAA<br>ATACGCCTTCGCCACTAGT<br>ATTCTTTCAGATATCTAAGC<br>ATTTCACTGCTACTCCTGAA<br>ATTCTAATTTCTTCTTTTATA<br>CTCAAGTTTATAAGTATTAA<br>TTTCAATATTAAATTACTTTA<br>ATAAATTTAAAAATTAATTTT<br>TAAAAACAACCTGCACACC<br>CTTTACGCCCAATAATTCCG<br>ATTAACGCTTGCACCCCTC<br>GTATTACGCGGGCTGCTGG<br>CACGAAGTTAGCCGGTGCT<br>TCTTTTACAAATAACGTCAA<br>AGATAATATTTTTTTTATTATA<br>AAATCTCTTCTTACTTTGTT<br>GAAAGTGTTTTACAACCCTA<br>AGGCCTTCTTCACACACGC<br>GATATAGCTGGATCAAGCT<br>TTCGCTCATTGTCCAATATC<br>CCCCACTGCTGCCTTCCGT<br>AAAAGTTTGGGCCGTGTCT<br>CAGTCCCAATGTGGTTGTT<br>CATCCTCTAAGATCAACTAC<br>GAATCATAGTCTTGTTAAGC<br>TTTTACTTTAACAACATACT<br>AATTCGATATAAGCTCTTCT<br>ATTAGCGAACGACATTCTC<br>GTTCTTTATCCATTAGGATA<br>CATATTGAATTACTATACAT<br>TTCTATATACTTTTCTAATAC |
|--|--|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                              |                                  |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|----------------------------------------------|----------------------------------|---------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                              |                                  |               | TAATAGGTAGATTCTTATAT<br>ATTACTCACCCGTTTCGCTG<br>CTAATTATTTTTTTAATAATT<br>CGCACAACTTGCATGTGTT<br>AAGCTTATCGCTAGCGTTC<br>AATCTGAGCTATGATCAAAC<br>TCA<br>(SEQ ID NO: 1)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| <i>Portiera aleyrodidarum</i><br><i>BT-B</i> | whiteflies<br>(Aleyrodoide<br>a) | bacteriocytes | AAGAGTTTGATCATGGCTC<br>AGATTGAACGCTAGCGGCA<br>GACATAACACATGCAAGTC<br>GAGCGGCATCATAACAGGTT<br>GGCAAGCGGCGCACGGGT<br>GAGTAATACATGTAAATATA<br>CCTAAAAGTGGGGAATAAC<br>GTACGGAAACGTACGCTAA<br>TACCGCATAATTATTACGAG<br>ATAAAGCAGGGGCTTGATA<br>AAAAAAATCAACCTTGCGCT<br>TTTAGAAAATTACATGCCGG<br>ATTAGCTAGTTGGTAGAGTA<br>AAAGCCTACCAAGGTAACG<br>ATCCGTAGCTGGTCTGAGA<br>GGATGATCAGCCACACTGG<br>GACTGAGAAAAGGCCCAGA<br>CTCCTACGGGAGGCAGCAG<br>TGGGGAATATTGGACAATG<br>GGGGGAACCCTGATCCAGT<br>CATGCCGCGTGTGTGAAGA<br>AGGCCTTTGGGTTGTAAAG<br>CACTTTCAGCGAAGAAGAA<br>AAGTTAGAAAATAAAAAGTT<br>ATAACTATGACGGTACTCG<br>CAGAAGAAGCACCGGCTAA<br>CTCCGTGCCAGCAGCCGC<br>GGTAAGACGGAGGGTGCAA<br>GCGTTAATCAGAATTACTG<br>GGCGTAAAGGGCATGTAGG<br>TGGTTTGTTAAGCTTTATGT<br>GAAAGCCCTATGCTTAACAT |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | AGGAACGGAATAAAGAACT<br>GACAAACTAGAGTGCAGAA<br>GAGGAAGGTAGAATTCCCG<br>GTGTAGCGGTGAAATGCGT<br>AGATATCTGGAGGAATACC<br>AGTTGCGAAGGCGACCTTC<br>TGGGCTGACACTGACACTG<br>AGATGCGAAAGCGTGGGGA<br>GCAAACAGGATTAGATACC<br>CTGGTAGTCCACGCTGTAA<br>ACGATATCAACTAGCCGTT<br>GGATTCTTAAAGAATTTTGT<br>GGCGTAGCTAACGCGATAA<br>GTTGATCGCCTGGGGAGTA<br>CGGTCGCAAGGCTAAAACT<br>CAAATGAATTGACGGGGGC<br>CCGCACAAGCGGTGGAGC<br>ATGTGGTTTAATTCGATGCA<br>ACGCGCAAAACCTTACCTA<br>CTCTTGACATCCAAAGTACT<br>TTCCAGAGATGGAAGGGTG<br>CCTTAGGGAACTTTGAGAC<br>AGGTGCTGCATGGCTGTCG<br>TCAGCTCGTGTTGTGAAAT<br>GTTGGGTTAAGTCCCGTAA<br>CGAGCGCAACCCTTGTCCT<br>TAGTTGCCAACGCATAAGG<br>CGGGAACTTTAAGGAGACT<br>GCTGGTGATAAACCGGAGG<br>AAGGTGGGGACGACGTCAA<br>GTCATCATGGCCCTTAAGA<br>GTAGGGCAACACACGTGCT<br>ACAATGGCAAAAACAAAGG<br>GTCGCAAAATGGTAACATG<br>AAGCTAATCCCAAAAAAATT<br>GTCTTAGTTCGGATTGGAG<br>TCTGAAACTCGACTCCATAA<br>AGTCGGAATCGCTAGTAAT<br>CGTGAATCAGAATGTCACG<br>GTGAATACGTTCTCGGGCC |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                                                               |                        |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-------------------------------------------------------------------------------|------------------------|---------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                               |                        |               | TTGTACACACCGCCCGTCA<br>CACCATGGAAGTGAAATGC<br>ACCAGAAGTGGCAAGTTTA<br>ACCAAAAAACAGGAGAACA<br>GTCACTACGGTGTGGTTCA<br>TGA CTGGGGTGAAGTCGTA<br>ACAAGGTAGCTGTAGGGGA<br>ACCTGTGGCTGGATCACCT<br>CCTTAA<br>(SEQ ID NO: 2)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <i>Buchnera aphidicola</i><br><i>str. APS</i><br><i>(Acyrthosiphon pisum)</i> | Aphids<br>(Aphidoidea) | bacteriocytes | AGAGTTTGATCATGGCTCA<br>GATTGAACGCTGGCGGCAA<br>GCCTAACACATGCAAGTCG<br>AGCGGCAGCGAGAAGAGA<br>GCTTGCTCTCTTTGTCGGC<br>AAGCGGCAAACGGGTGAGT<br>AATATCTGGGGATCTACCC<br>AAAAGAGGGGGATAACTAC<br>TAGAAATGGTAGCTAATACC<br>GCATAATGTTGAAAAACCAA<br>AGTGGGGGACCTTTTGGCC<br>TCATGCTTTTGGATGAACCC<br>AGACGAGATTAGCTTGTTG<br>GTAGAGTAATAGCCTACCA<br>AGGCAACGATCTCTAGCTG<br>GTCTGAGAGGATAACCAGC<br>CACACTGGA ACTGAGACAC<br>GGTCCAGACTCCTACGGGA<br>GGCAGCAGTGGGGAATATT<br>GCACAATGGGCGAAAGCCT<br>GATGCAGCTATGCCGCGTG<br>TATGAAGAAGGCCTTAGGG<br>TTGTAAAGTACTTTCAGCGG<br>GGAGGAAAAAAATAAACT<br>AATAATTTTATTTCGTGACG<br>TTACCCGCAGAAGAAGCAC<br>CGGCTAACTCCGTGCCAGC<br>AGCCGCGGTAATACGGAGG<br>GTGCAAGCGTTAATCAGAA<br>T TACTGGGCGTAAAGAGCG |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|--|--|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | CGTAGGTGGTTTTTTAAGTC<br>AGGTGTGAAATCCCTAGGC<br>TCAACCTAGGAACTGCATTT<br>GAAACTGGAAAAGTAGAGT<br>TTCGTAGAGGGAGGTAGAA<br>TTCTAGGTGTAGCGGTGAA<br>ATGCGTAGATATCTGGAGG<br>AATACCCGTGGCGAAAGCG<br>GCCTCCTAAACGAAAAGT<br>ACACTGAGGCGCGAAAGCG<br>TGGGGAGCAAACAGGATTA<br>GATACCCTGGTAGTCCATG<br>CCGTAAACGATGTGCGACTT<br>GGAGGTTGTTTCCAAGAGA<br>AGTGACTTCCGAAGCTAAC<br>GCATTAAGTCGACCGCCTG<br>GGGAGTACGGCCGCAAGG<br>CTAAAGTCAAATGAATTGA<br>CGGGGGCCCGCACAAGCG<br>GTGGAGCATGTGGTTTAAT<br>TCGATGCAACGCGAAAAAC<br>CTTACCTGGTCTTGACATCC<br>ACAGAATTCTTTAGAAATAA<br>AGAAGTGCCTTCGGGAGCT<br>GTGAGACAGGTGCTGCATG<br>GCTGTCGTCAGCTCGTGTT<br>GTGAAATGTTGGGTAAAGT<br>CCCGCAACGAGCGCAACCC<br>TTATCCCCTGTTGCCAGCG<br>GTTCCGGCCGGGAAGTCAGA<br>GGAGACTGCCGGTTATAAA<br>CCGGAGGAAGGTGGGGAC<br>GACGTCAAGTCATCATGGC<br>CCTTACGACCAGGGCTACA<br>CACGTGCTACAATGGTTTAT<br>ACAAAGAGAAGCAAATCTG<br>CAAAGACAAGCAAACCTCA<br>TAAAGTAAATCGTAGTCCG<br>GACTGGAGTCTGCAACTCG<br>ACTCCACGAAGTCGGAATC |
|--|--|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                                                              |                        |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|------------------------------------------------------------------------------|------------------------|---------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                              |                        |               | GCTAGTAATCGTGGATCAG<br>AATGCCACGGTGAATACGT<br>TCCCGGGCCTTGTACACAC<br>CGCCCGTCACACCATGGGA<br>GTGGGTTGCAAAGAAGCA<br>GGTATCCTAACCCTTTAAAA<br>GGAAGGCGCTTACCACTTT<br>GTGATTCATGACTGGGGTG<br>AAGTCGTAACAAGGTAACC<br>GTAGGGGAACCTGCGGTTG<br>GATCACCTCCTT<br>(SEQ ID NO: 3)                                                                                                                                                                                                                                                                                                                                                                                                |
| <i>Buchnera aphidicola</i><br><i>str. Sg (Schizaphis</i><br><i>graminum)</i> | Aphids<br>(Aphidoidea) | bacteriocytes | AAACTGAAGAGTTTGATCAT<br>GGCTCAGATTGAACGCTGG<br>CGGCAAGCCTAACACATGC<br>AAGTCGAGCGGCAGCGAAA<br>AGAAAGCTTGCTTTCTTGTC<br>GGCGAGCGGCAAACGGGT<br>GAGTAATATCTGGGGATCT<br>GCCCAAAGAGGGGGGATAA<br>CTACTAGAAATGGTAGCTAA<br>TACCGCATAAAGTTGAAAAA<br>CCAAAGTGGGGGACCTTTT<br>TTAAAGGCCTCATGCTTTTG<br>GATGAACCCAGACGAGATT<br>AGCTTGTTGGTAAGGTAAA<br>AGCTTACCAAGGCAACGAT<br>CTCTAGCTGGTCTGAGAGG<br>ATAACCAGCCACACTGGAA<br>CTGAGACACGGTCCAGACT<br>CCTACGGGAGGCAGCAGT<br>GGGGAATATTGCACAATGG<br>GCGAAAGCCTGATGCAGCT<br>ATGCCGCGTGTATGAAGAA<br>GGCCTTAGGGTTGTAAAGT<br>ACTTTCAGCGGGGAGGAAA<br>AAATTAAACTAATAATTTTA<br>TTTTGTGACGTTACCCGCA<br>GAAGAAGCACCGGCTAACT<br>CCGTGCCAGCAGCCGCGG |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|--|--|--|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | TAATACGGAGGGTGCGAGC<br>GTTAATCAGAATTACTGGG<br>CGTAAAGAGCACGTAGGTG<br>GTTTTTTAAGTCAGATGTGA<br>AATCCCTAGGCTTAACCTA<br>GGAAGTGCATTTGAAACTG<br>AAATGCTAGAGTATCGTAG<br>AGGGAGGTAGAATTCTAGG<br>TGTAGCGGTGAAATGCGTA<br>GATATCTGGAGGAATACCC<br>GTGGCGAAAGCGGCCTCCT<br>AAACGAATACTGACACTGA<br>GGTGCGAAAGCGTGGGGA<br>GCAAACAGGATTAGATACC<br>CTGGTAGTCCATGCCGTAA<br>ACGATGTCGACTTGGAGGT<br>TGTTTCCAAGAGAAGTGAC<br>TTCCGAAGCTAACGCGTTA<br>AGTCGACCGCCTGGGGAGT<br>ACGGCCGCAAGGCTAAAAC<br>TCAAATGAATTGACGGGGG<br>CCCGCACAAAGCGGTGGAG<br>CATGTGGTTTAATTGATGC<br>AACGCGAAAAACCTTACCT<br>GGTCTTGACATCCACAGAA<br>TTTTTTAGAAATAAAAAAGT<br>GCCTTCGGGAACTGTGAGA<br>CAGGTGCTGCATGGCTGTC<br>GTCAGCTCGTGTTGTGAAA<br>TGTTGGGTTAAGTCCCGCA<br>ACGAGCGCAACCCTTATCC<br>CCTGTTGCCAGCGGTTCCG<br>CCGGGAACTCAGAGGAGAC<br>TGCCGGTTATAAACCGGAG<br>GAAGGTGGGGACGACGTC<br>AAGTCATCATGGCCCTTAC<br>GACCAGGGCTACACACGTG<br>CTACAATGGTTTATACAAAG<br>AGAAGCAAATCTGTAAAGA<br>CAAGCAAACCTCATAAAGTA |
|--|--|--|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                                                              |                        |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|------------------------------------------------------------------------------|------------------------|---------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                              |                        |               | AATCGTAGTCCGGACTGGA<br>GTCTGCAACTCGACTCCAC<br>GAAGTCGGAATCGCTAGTA<br>ATCGTGGATCAGAATGCCA<br>CGGTGAATACGTTCCCGGG<br>CCTTGTACACACCGCCCGT<br>CACACCATGGGAGTGGGTT<br>GCAAAAGAAGCAGATTTCC<br>TAACCACGAAAGTGGAAGG<br>CGTCTACCACTTTGTGATTC<br>ATGACTGGGGTGAAGTCGT<br>AACAAGGTAACCGTAGGGG<br>AACCTGCGGTTGGATCACC<br>TCCTTA<br>(SEQ ID NO: 4)                                                                                                                                                                                                                                                              |
| <i>Buchnera aphidicola</i><br><i>str. Bp (Baizongia</i><br><i>pistaciae)</i> | Aphids<br>(Aphidoidea) | bacteriocytes | ACTTAAAATTGAAGAGTTTG<br>ATCATGGCTCAGATTGAAC<br>GCTGGCGGCAAGCTTAACA<br>CATGCAAGTCGAGCGGCAT<br>CGAAGAAAAGTTTACTTTTC<br>TGGCGGCGAGCGGCAAAC<br>GGGTGAGTAACATCTGGGG<br>ATCTACCTAAAAGAGGGGG<br>ACAACCATTGGAAACGATG<br>GCTAATACCGCATAATGTTT<br>TTAAATAAACCAAAGTAGGG<br>GACTAAAATTTTTAGCCTTA<br>TGCTTTTAGATGAACCCAGA<br>CGAGATTAGCTTGATGGTA<br>AGGTAATGGCTTACCAAGG<br>CGACGATCTCTAGCTGGTC<br>TGAGAGGATAACCAGCCAC<br>ACTGGAAGTGAAGTACGGT<br>CCAGACTCCTACGGGAGGC<br>AGCAGTGGGGAATATTGCA<br>CAATGGGCTAAAGCCTGAT<br>GCAGCTATGCCGCGTGTAT<br>GAAGAAGGCCTTAGGGTTG<br>TAAAGTACTTTCAGCGGGG<br>AGGAAAGAATTATGTCTAAT |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|--|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | ATACATATTTTGTGACGTTA<br>CCCGAAGAAGAAGCACCGG<br>CTAACTCCGTGCCAGCAGC<br>CGCGGTAATACGGAGGGTG<br>CGAGCGTTAATCAGAATTA<br>CTGGGCGTAAAGAGCACGT<br>AGGCGGTTTATTAAGTCAG<br>ATGTGAAATCCCTAGGCTTA<br>ACTTAGGAACTGCATTTGAA<br>ACTAATAGACTAGAGTCTCA<br>TAGAGGGAGGTAGAATTCT<br>AGGTGTAGCGGTGAAATGC<br>GTAGATATCTAGAGGAATA<br>CCCGTGGCGAAAGCGACCT<br>CCTAAATGAAAAGTACGCG<br>TGAGGTGCGAAAGCGTGG<br>GGAGCAAACAGGATTAGAT<br>ACCCTGGTAGTCCATGCTG<br>TAAACGATGTCGACTTGGA<br>GGTTGTTTCCTAGAGAAGT<br>GGCTTCCGAAGCTAACGCA<br>TTAAGTCGACCGCCTGGGG<br>AGTACGGTCGCAAGGCTAA<br>AACTCAAATGAATTGACGG<br>GGGCCCCGCACAAGCGGTG<br>GAGCATGTGGTTTAATTCTG<br>ATGCAACGCGAAGAACCTT<br>ACCTGGTCTTGACATCCATA<br>GAATTTTTTAGAGATAAAAG<br>AGTGCCTTAGGGAACTATG<br>AGACAGGTGCTGCATGGCT<br>GTCGTCAGCTCGTGTTGTG<br>AAATGTTGGGTTAAGTCCC<br>GCAACGAGCGCAACCCCTA<br>TCCTTTGTTGCCATCAGGTT<br>ATGCTGGGAACTCAGAGGA<br>GACTGCCGGTTATAAACCG<br>GAGGAAGGTGGGGATGAC<br>GTCAAGTCATCATGGCCCT<br>TACGACCAGGGCTACACAC |
|--|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                          |                        |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------------------------------|------------------------|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                          |                        |               | GTGCTACAATGGCATATAC<br>AAAGAGATGCAACTCTGCG<br>AAGATAAGCAAACCTCATAA<br>AGTATGTCGTAGTCCGGAC<br>TGGAGTCTGCAACTCGACT<br>CCACGAAGTAGGAATCGCT<br>AGTAATCGTGGATCAGAAT<br>GCCACGGTGAATACGTTCC<br>CGGGCCTTGTACACACCGC<br>CCGTCACACCATGGGAGTG<br>GGTTGCAAAAGAAGCAGGT<br>AGCTTAACCAGATTATTTTA<br>TTGGAGGGCGCTTACCACT<br>TTGTGATTCATGACTGGGG<br>TGAAGTCGTAACAAGGTAA<br>CCGTAGGGGAACCTGCGGT<br>TGGATCACCTCCTTA<br>(SEQ ID NO: 5)                                                                                                                                |
| <i>Buchnera aphidicola</i><br><i>BCc</i> | Aphids<br>(Aphidoidea) | bacteriocytes | ATGAGATCATTAAATATATAA<br>AAATCATGTTCCAATTAAAA<br>AATTAGGACAAAATTTTTTA<br>CAGAATAAAGAAATTATTAA<br>TCAGATAATTAATTTAATAA<br>ATATTAATAAAAATGATAAT<br>ATTATTGAAATAGGATCAGG<br>ATTAGGAGCGTTAACTTTTC<br>CTATTTGTAGAATCATTAAA<br>AAAATGATAGTATTAGAAAT<br>TGATGAAGATCTTGTGTTTT<br>TTTTAACTCAAAGTTTATTTA<br>TTAAAAAATTACAAATTATAA<br>TTGCTGATATTATAAAATTT<br>GATTTTTGTTGTTTTTTTTCT<br>TTACAGAAATATAAAAAATA<br>TAGGTTTATTGGTAATTTAC<br>CATATAATATTGCTACTATA<br>TTTTTTTTTAAAAACAATTAAA<br>TTTCTTTATAATATAATTGAT<br>ATGCATTTTATGTTTCAAAA<br>AGAAGTAGCAAAGAGATTA |

|                                                             |                        |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-------------------------------------------------------------|------------------------|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                             |                        |               | TTAGCTACTCCTGGTACTAA<br>AGAATATGGTAGATTAAGTA<br>TTATTGCACAATATTTTTATA<br>AGATAGAAACTGTTATTAAT<br>GTTAATAAATTTAATTTTTTT<br>CCTACTCCTAAAGTAGATTC<br>TACTTTTTTACGATTTACTC<br>CTAAATATTTTAATAGTAAAT<br>ATAAAATAGATAAACATTTT<br>TCTGTTTTAGAATTAATTAC<br>TAGATTTTCTTTTCAACATA<br>GAAGAAAATTTTTAAATAAT<br>AATTTAATATCTTTATTTTCT<br>ACAAAAGAATTAATTTCTTT<br>AGATATTGATCCATATTCAA<br>GAGCAGAAAATGTTTCTTTA<br>ATTCAATATTGTAAATTAAT<br>GAAATATTATTTGAAAAGAA<br>AAATTTTATGTTTAGATTAA<br>(SEQ ID NO: 6)   |
| <i>Buchnera aphidicola</i><br>( <i>Cinara tujaefilina</i> ) | Aphids<br>(Aphidoidea) | bacteriocytes | TTATCTTATTTACATATAC<br>GTAATATTGCGCTGCGTGC<br>ACGAGGATTTTTTTGAATTT<br>CAGATATATTTGGTTTAATA<br>CGTTTAATAAAACGTATTTT<br>TTTTTTTATTTTTCTTATTTG<br>CAATTCAGTAATAGGAAGTT<br>TTTAGGTATATTTGGATAA<br>TACTGTAATTCTTAATAAA<br>GTTTTTTACAATCCTATCTT<br>CAATAGAATGAAAACATAA<br>ATAGCAATTTTTGATCCGGA<br>ATGTAATATGTTAATAATAA<br>TTTTTAATATTTTATGTAATT<br>CATTTATTTCTTGGTTAATAT<br>ATATTCGAAAAGCTTGAAAT<br>GTTCTCGTAGCTGGATGTTT<br>AAATTTGTCATATTTTGGGA<br>TTGATTTTTTTTATGATTTGAA<br>CTAACTCTAACGTGCTTGTT |

|                                                                           |                        |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|---------------------------------------------------------------------------|------------------------|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                           |                        |               | ATGGTTTTTTTTTTTATTGT<br>AATATGATGGCTCGGGATA<br>TTTTTTTTGCGTATTTTCTT<br>CGCCAAAATTTTTTATTACC<br>TGTTCTATTGTTTTTTGGTTT<br>GTTTTTTTTTAACCATTGACT<br>AACTGATATTCCAGATTTAG<br>GGTTCATACGCATATCTAAA<br>GGTCCATCATTCATAAATGA<br>AAATCCTCGGATACTAGAAT<br>TTAACTGTATTGAAGAAATA<br>CCTAAATCTAATAATATTCC<br>ATCTATTTTATCTCTATTTTT<br>TTCTTTTTTTAATATTTTTTC<br>AATATTAGAAAATTTACCTA<br>AAAATATTTTAAATCGCGAA<br>TCTTTTATTTTTTTTCCGATT<br>TTTATAGATTGTGGGTCTTG<br>ATCAATACTATATAACTTTC<br>CATTAACCCCTAATTCTTGA<br>AGAATTGCTTTTGAATGACC<br>ACCACCTCCAAATGTACAAT<br>CAACATATGTACCGTCTTTT<br>TTTATTTTTTAAGTATTGTATG<br>ATTTCTTTTGTTAAAACAGG<br>TTTATGAATCAT<br>(SEQ ID NO: 7) |
| <i>Buchnera aphidicola</i><br><i>str. G002 (Myzus</i><br><i>persicae)</i> | Aphids<br>(Aphidoidea) | bacteriocytes | ATGAAAAGTATAAAAACTTT<br>TAAAAAACACTTTCCTGTGA<br>AAAAATATGGACAAAATTTT<br>CTTATTAATAAAGAGATCAT<br>AAAAAATATTGTTAAAAAAA<br>TTAATCCAAATATAGAACAA<br>ACATTAGTAGAAATCGGAC<br>CAGGATTAGCTGCATTAACT<br>GAGCCCATATCTCAGTTATT<br>AAAAGAGTTAATAGTTATTG<br>AAATAGACTGTAATCTATTA<br>TATTTTTTAAAAAACAACC<br>ATTTTATTCAAATTAATAGT                                                                                                                                                                                                                                                                                                                                           |

|                                                                               |                        |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-------------------------------------------------------------------------------|------------------------|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                               |                        |               | TTTTGTCAAGATGCTTTAA<br>ACTTTAATTATACAAATTTAT<br>TTTATAAAAAAATAAATTAA<br>TTCGTATTTTTTGGTAATTTA<br>CCATATAATATCTCTACATC<br>TTTAATTATTTTTTTATTTCA<br>ACACATTAGAGTAATTCAAG<br>ATATGAATTTTATGCTTCAA<br>AAAGAAGTTGCTGCAAGAT<br>TAATTGCATTACCTGGAAAT<br>AAATATTACGGTCGTTTGAG<br>CATTATATCTCAATATTATT<br>GTGATATCAAATTTTATTA<br>AATGTTGCTCCTGAAGATTT<br>TTGGCCTATTCCGAGAGTT<br>CATTCTATATTTGTAAATTTA<br>ACACCTCATCATAATTCTCC<br>TTATTTTGTTTATGATATTAA<br>TATTTTAAGCCTTATTACAA<br>ATAAGGCTTTCCAAATAGA<br>AGAAAAATATTACGTCATAG<br>TTAAAAAATTTATTTTCTGA<br>AACAACTTTATTAAATTTAG<br>ATATTAATCCCAGATTAAGA<br>GCTGAAAATATTTCTGTTTT<br>TCAGTATTGTCAATTAGCTA<br>ATTATTTGTATAAAAAAAATT<br>ATACTAAAAAAATTAA<br>(SEQ ID NO: 8) |
| <i>Buchnera aphidicola</i><br><i>str. Ak (Acyrthosiphon</i><br><i>kondoi)</i> | Aphids<br>(Aphidoidea) | bacteriocytes | ATTATAAAAAATTTTAAAAAA<br>CATTTTCCTTTAAAAAGGTA<br>TGGACAAAATTTTCTTGTCA<br>ATACAAAACTATTCAAAG<br>ATAATTAATATAATTAATCCA<br>AACACCAAACAAACATTAGT<br>GGAAATTGGACCTGGATTA<br>GCTGCATTAACAAAACCAAT<br>TTGTCAATTATTAGAAGAAT<br>TAATTGTTATTGAAATAGAT<br>CCTAATTTATTGTTTTTATTA                                                                                                                                                                                                                                                                                                                                                                                                                                        |

|                                                                              |                        |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|------------------------------------------------------------------------------|------------------------|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                              |                        |               | AAAAAACGTTTCATTTTATTC<br>AAAATTAACAGTTTTTTATC<br>AAGACGCTTTAAATTTCAAT<br>TATACAGATTTGTTTTATAA<br>GAAAAATCAATTAATTCGTG<br>TTTTTGGAACCTTGCCATAT<br>AATATTTCTACATCTTTAATT<br>ATTTCTTTATTCAATCATATT<br>AAAGTTATTCAAGATATGAA<br>TTTTATGTTACAGAAAGAGG<br>TTGCTGAAAGATTAATTTCT<br>ATTCCTGGAAATAAATCTTA<br>TGGCCGTTTAAGCATTATTT<br>CTCAGTATTATTGTAAAATT<br>AAAATATTATTAAATGTTGT<br>ACCTGAAGATTTTCGACCTA<br>TACCGAAAGTGCATTCTGTT<br>TTTATCAATTTAACTCCTCA<br>TACCAATTCTCCATATTTTG<br>TTTATGATACAAATATCCTC<br>AGTTCTATCACAAGAAATGC<br>TTTTCAAATAGAAGGAAAA<br>TTTTGCGTCATAGTTTAAAA<br>AATTTATTTTCTGAAAAAGA<br>ACTAATTCAATTAGAAATTA<br>ATCCAAATTTACGAGCTGAA<br>AATATTTCTATCTTTCAGTAT<br>TGTC AATTAGCTGATTATTT<br>ATATAAAAAATTAAATAATCT<br>TGTA AAAATCAATTAA<br>(SEQ ID NO: 9) |
| <i>Buchnera aphidicola</i><br><i>str. Ua (Uroleucon</i><br><i>ambrosiae)</i> | Aphids<br>(Aphidoidea) | bacteriocytes | ATGATACTAAATAAATATAA<br>AAAATTTATTCTTTAAAAA<br>GATACGGACAAAATTTTCTT<br>GTAAATAGAGAAATAATCAA<br>AAATATTATCAAAATAATTAA<br>TCCTAAAAAAACGCAAACAT<br>TATTAGAAATTGGACCGGG<br>TTTAGGTGCGTTAACAAAAC<br>CTATTTGTGAATTTTAAAT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |

|                                                         |                        |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|---------------------------------------------------------|------------------------|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                         |                        |               | <div>GAAC TTATCGTCATTGAAAT<br/>AGATCCTAATATATTATCTT<br/>TTTTAAAGAAATGTATATTTT<br/>TTGATAAATTAAAAATATATT<br/>GTCATAATGCTTTAGATTTT<br/>AATTATAAAAATATATTCTAT<br/>AAAAAAAGTCAATTAATTCTG<br/>TATTTTTTGGAAATTTACCAT<br/>ATAATATTTCTACATCTTTAA<br/>TAATATATTTATTTTCGGAAT<br/>ATTGATATTATTCAAGATAT<br/>GAATTTTATGTTACAACAAG<br/>AAGTGGCTAAAAGATTAGTT<br/>GCTATTCCTGGTGAAAAACT<br/>TTATGGTCGTTTAAGTATTA<br/>TATCTCAATATTATTGTAATA<br/>TTAAAATATTATTACATATTC<br/>GACCTGAAAATTTTCAACCT<br/>ATTCCTAAAGTTAATTCAAT<br/>GTTTGTAAATTTAACTCCGC<br/>ATATTCATTCTCCTTATTTTG<br/>TTTATGATATTAATTTATTAA<br/>CTAGTATTACAAAACATGCT<br/>TTTCAACATAGAAGAAAAAT<br/>ATTGCGTCATAGTTTAAGAA<br/>ATTTTTTTTCTGAGCAAGAT<br/>TTAATTCATTTAGAAATTAAT<br/>CCAAATTTAAGAGCTGAAAA<br/>TGTTTCTATTATTCAATATTG<br/>TCAATTGGCTAATAATTTAT<br/>ATAAAAAACATAAACAGTTT<br/>ATTAATAATTAA<br/>(SEQ ID NO: 10)</div> |
| <i>Buchnera aphidicola</i><br>( <i>Aphis glycines</i> ) | Aphids<br>(Aphidoidea) | bacteriocytes | <div>ATGAAAAAGCATATTCCTAT<br/>AAAAAAATTTAGTCAAAATT<br/>TTCTTG TAGATTTGAGTGTG<br/>ATTAAAAAAATAATTAAATTT<br/>ATTAATCCGCAGTTAAATGA<br/>AATATTGGTTGAAATTGGAC<br/>CGGGATTAGCTGCTATCAC</div>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |

|                          |                      |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|--------------------------|----------------------|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                          |                      |               | TCGACCTATTTGTGATTGA<br>TAGATCATTTAATTGTGATT<br>GAAATTGATAAAAATTTTATT<br>AGATAGATTAAAACAGTTCT<br>CATTTTATTCAAATTAACA<br>GTATATCATCAAGATGCTTT<br>AGCATTTGATTACATAAAGT<br>TATTTAATAAAAAAATAAAT<br>TAGTTCGAATTTTGGTAAT<br>TTACCATATCATGTTTCTAC<br>GTCTTTAATATTGCATTTATT<br>TAAAAGAATTAATATTATTAA<br>AGATATGAATTTTATGCTAC<br>AAAAAGAAGTTGCTGAACG<br>TTTAATTGCAACTCCAGGTA<br>GTAAATTATATGGTCGTTTA<br>AGTATTATTTCTCAATATTAT<br>TGTAATATAAAAGTTTTATT<br>GCATGTGTCTTCAAATGTT<br>TTAAACCAGTTCCTAAAGTA<br>GAATCAATTTTCTTAATTT<br>GACACCTTATACTGATTATT<br>TCCCTTATTTTACTTATAAT<br>GTAAACGTTCTTAGTTATAT<br>TACAAATTTAGCTTTTCAA<br>AAAGAAGAAAAATATTACGT<br>CATAGTTTAGGTAAAATATT<br>TTCTGAAAAAGTTTTTATAA<br>AATTAAATATTAATCCCAA<br>TTAAGACCTGAGAATATTT<br>TATATTACAATATTGTCAGT<br>TATCTAATTATATGATAGAA<br>AATAATATTCATCAGGAACA<br>TGTTTGTATTTAA<br>(SEQ ID NO: 11) |
| <i>Annandia pinicola</i> | (Phylloxeroid<br>ea) | bacteriocytes | AGATTGAACGCTGGCGGCA<br>TGCCTTACACATGCAAGTC<br>GAACGGTAACAGGTCTTCG<br>GACGCTGACGAGTGGCGAA<br>CGGGTGAGTAATACATCGG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|--|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | AACGTGCCCAGTCGTGGGG<br>GATAACTACTCGAAAGAGT<br>AGCTAATACCGCATACGAT<br>CTGAGGATGAAAGCGGGG<br>GACCTTCGGGCCTCGCGC<br>GATTGGAGCGGCCGATGG<br>CAGATTAGGTAGTTGGTGG<br>GATAAAAGCTTACCAAGCC<br>GACGATCTGTAGCTGGTCT<br>GAGAGGACGACCAGCCACA<br>CTGGAACTGAGATACGGTC<br>CAGACTCTTACGGGAGGCA<br>GCAGTGGGGAATATTGCAC<br>AATGGGCGCAAGCCTGATG<br>CAGCTATGTCGCGTGTATG<br>AAGAAGACCTTAGGGTTGT<br>AAAGTACTTTCGATAGCATA<br>AGAAGATAATGAGACTAATA<br>ATTTTATTGTCTGACGTTAG<br>CTATAGAAGAAGCACCGGC<br>TAACTCCGTGCCAGCAGCC<br>GCGGTAATACGGGGGGTG<br>CTAGCGTTAATCGGAATTAC<br>TGGGCGTAAAGAGCATGTA<br>GGTGGTTTATTAAGTCAGAT<br>GTGAAATCCCTGGACTTAAT<br>CTAGGAACTGCATTTGAAA<br>CTAATAGGCTAGAGTTTCGT<br>AGAGGGAGGTAGAATTCTA<br>GGTGTAGCGGTGAAATGCA<br>TAGATATCTAGAGGAATATC<br>AGTGGCGAAGGCGACCTTC<br>TGGACGATAACTGACGCTA<br>AAATGCGAAAGCATGGGTA<br>GCAAACAGGATTAGATACC<br>CTGGTAGTCCATGCTGTAA<br>ACGATGTCGACTAAGAGGT<br>TGGAGGTATAACTTTTAATC<br>TCTGTAGCTAACGCGTTAA<br>GTCGACCGCCTGGGGAGTA |
|--|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                          |             |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|--------------------------|-------------|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                          |             |               | CGGTCGCAAGGCTAAACT<br>CAAATGAATTGACGGGGGC<br>CTGCACAAGCGGTGGAGCA<br>TGTGGTTTAATTGATGCAA<br>CGCGTAAAACCTTACCTGG<br>TCTTGACATCCACAGAATTT<br>TACAGAAATGTAGAAGTGC<br>AATTTGAACTGTGAGACAG<br>GTGCTGCATGGCTGTCGTC<br>AGCTCGTGTTGTGAAATGTT<br>GGGTAAAGTCCCGCAACGA<br>GCGCAACCCTTGTCCTTTG<br>TTACCATAAGATTTAAGGAA<br>CTCAAAGGAGACTGCCGGT<br>GATAAACTGGAGGAAGGCG<br>GGGACGACGTCAAGTCATC<br>ATGGCCCTTATGACCAGGG<br>CTACACACGTGCTACAATG<br>GCATATACAAAGAGATGCA<br>ATATTGCGAAATAAAGCCAA<br>TCTTATAAAATATGTCCTAG<br>TTCGGACTGGAGTCTGCAA<br>CTCGACTCCACGAAGTCGG<br>AATCGCTAGTAATCGTGGA<br>TCAGCATGCCACGGTGAAT<br>ATGTTTCCAGGCCTTGTACA<br>CACCGCCCGTCACACCATG<br>GAAGTGGATTGCAAAGAA<br>GTAAGAAAATTAACCTTCTT<br>AACAAGGAAATAACTTACCA<br>CTTTGTGACTCATAACTGG<br>GGTGA<br>(SEQ ID NO: 12) |
| <i>Moranella endobia</i> | (Coccoidea) | bacteriocytes | TCTTTTGGTAAGGAGGTG<br>ATCCAACCGCAGGTTCCCC<br>TACGGTTACCTTGTTACGAC<br>TTCACCCCAGTCATGAATCA<br>CAAAGTGGTAAGCGCCCTC<br>CTAAAAGGTTAGGCTACCT<br>ACTTCTTTTGCAACCCACTT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | CCATGGTGTGACGGGCGGT<br>GTGTACAAGGCCCGGGAAC<br>GTATTCACCGTGGCATTCT<br>GATCCACGATTACTAGCGA<br>TTCCTACTTCATGGAGTCGA<br>GTTGCAGACTCCAATCCGG<br>ACTACGACGCACTTTATGA<br>GGTCCGCTAACTCTCGCGA<br>GCTTGCTTCTCTTTGTATGC<br>GCCATTGTAGCACGTGTGT<br>AGCCCTACTCGTAAGGGCC<br>ATGATGACTTGACGTCATC<br>CCCACCTTCCTCCGGTTTAT<br>CACCGGCAGTCTCCTTTGA<br>GTTCCCGACCGAATCGCTG<br>GCAAAAAGGATAAGGGTT<br>GCGCTCGTTGCGGGACTTA<br>ACCCAACATTTACAACAC<br>GAGCTGACGACAGCCATGC<br>AGCACCTGTCTCAGAGTTC<br>CCGAAGGTACCAAAACATC<br>TCTGCTAAGTTCTCTGGATG<br>TCAAGAGTAGGTAAGGTTC<br>TTCGCGTTGCATCGAATTAA<br>ACCACATGCTCCACCGCTT<br>GTGCGGGCCCCCGTCAATT<br>CATTTGAGTTTTAACCTTGC<br>GGCCGTA CTCCCCAGGCG<br>GTCGATTTAACGCGTTAACT<br>ACGAAAGCCACAGTTCAAG<br>ACCACAGCTTTCAAATCGA<br>CATAGTTTACGGCGTGGAC<br>TACCAGGGTATCTAATCCT<br>GTTTGCTCCCCACGCTTTC<br>GTACCTGAGCGTCAGTATT<br>CGTCCAGGGGGCCGCCTT<br>CGCCACTGGTATTCCTCCA<br>GATATCTACACATTTACCG<br>CTACACCTGGAATTCTACC<br>CCCCTCTACGAGACTCTAG |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                              |               |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|----------------------------------------------|---------------|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                              |               |               | <div>CCTATCAGTTTCAAATGCAG<br/>TTCCTAGGTTAAGCCCAGG<br/>GATTTACATCTGACTTAAT<br/>AAACCGCCTACGTACTCTTT<br/>ACGCCCAGTAATTCCGATT<br/>AACGCTTGCACCCTCCGTA<br/>TTACCGCGGCTGCTGGCAC<br/>GGAGTTAGCCGGTGCTTCT<br/>TCTGTAGGTAACGTCAATCA<br/>ATAACCGTATTAAGGATATT<br/>GCCTTCCTCCCTACTGAAA<br/>GTGCTTTACAACCCGAAGG<br/>CCTTCTTCACACACGCGGC<br/>ATGGCTGCATCAGGGTTTC<br/>CCCCATTGTGCAATATTCCC<br/>CACTGCTGCCTCCCGTAGG<br/>AGTCTGGACCGTGTCTCAG<br/>TTCCAGTGTGGCTGGTCAT<br/>CCTCTCAGACCAGCTAGGG<br/>ATCGTCGCCTAGGTAAGCT<br/>ATTACCTCACCTACTAGCTA<br/>ATCCCATCTGGGTTCATCT<br/>GAAGGTGTGAGGCCAAAAG<br/>GTCCCCCACTTTGGTCTTA<br/>CGACATTATGCGGTATTAG<br/>CTACCGTTTCCAGCAGTTAT<br/>CCCCCTCCATCAGGCAGAT<br/>CCCCAGACTTTACTCACCC<br/>GTTGCTGCTCGCCGGCAA<br/>AAAAGTAACTTTTTTCCGT<br/>TGCCGCTCAACTTGCAATGT<br/>GTTAGGCCTGCCGCCAGCG<br/>TTCAATCTGAGCCATGATCA<br/>AACTCTTCAATTAAA<br/>(SEQ ID NO: 13)</div> |
| <div>Ishikawaella capsulata<br/>Mpkobe</div> | (Heteroptera) | bacteriocytes | <div>AAATTGAAGAGTTTGATCAT<br/>GGCTCAGATTGAACGCTAG<br/>CGGCAAGCTTAACACATGC<br/>AAGTCGAACGGTAACAGAA<br/>AAAAGCTTGCTTTTTTGCTG</div>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | ACGAGTGGCGGACGGGTG<br>AGTAATGTCTGGGGATCTA<br>CCTAATGGCGGGGGGATAAC<br>TACTGGAAACGGTAGCTAA<br>TACCGCATAATGTTGTAAAA<br>CCAAAGTGGGGGACCTTAT<br>GGCCTCACACCATTAGATG<br>AACCTAGATGGGATTAGCT<br>TGTAGGTGGGGTAAAGGCT<br>CACCTAGGCAACGATCCCT<br>AGCTGGTCTGAGAGGATGA<br>CCAGCCACACTGGAAGTGA<br>GATACGGTCCAGACTCCTA<br>CGGGAGGCAGCAGTGGGG<br>AATCTTGCACAATGGGCGC<br>AAGCCTGATGCAGCTATGT<br>CGCGTGTATGAAGAAGGCC<br>TTAGGGTTGTAAAGTACTTT<br>CATCGGGGAAGAAGGATAT<br>GAGCCTAATATTCTCATATA<br>TTGACGTTACCTGCAGAAG<br>AAGCACCGGCTAACTCCGT<br>GCCAGCAGCCGCGGTAAAC<br>CGGAGGGTGCGAGCGTTAA<br>TCGGAATTACTGGGCGTAA<br>AGAGCACGTAGGTGGTTTA<br>TTAAGTCATATGTGAAATCC<br>CTGGGCTTAACCTAGGAAC<br>TGCATGTGAAACTGATAAAC<br>TAGAGTTTCGTAGAGGGAG<br>GTGGAATTCCAGGTGTAGC<br>GGTGAAATGCGTAGATATC<br>TGGAGGAATATCAGAGGCG<br>AAGGCGACCTTCTGGACGA<br>AACTGACACTCAGGTGCG<br>AAAGCGTGGGGAGCAAACA<br>GGATTAGATACCCTGGTAG<br>TCCACGCTGTAAACAATGT<br>CGACTAAAAAACTGTGAGC<br>TTGACTTGTGGTTTTTGTAG |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                              |                             |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|------------------------------|-----------------------------|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                              |                             |               | CTAACGCATTAAGTCGACC<br>GCCTGGGGAGTACGGCCG<br>CAAGGTTAAAACTCAAATGA<br>ATTGACGGGGGTCCGCACA<br>AGCGGTGGAGCATGTGGTT<br>TAATTCGATGCAACGCGAA<br>AAACCTTACCTGGTCTTGAC<br>ATCCAGCGAATTATATAGAA<br>ATATATAAGTGCCTTTCGGG<br>GAACTCTGAGACGCTGCAT<br>GGCTGTCGTCAGCTCGTGT<br>TGTGAAATGTTGGGTAAAGT<br>CCCGCAACGAGCGCCCTTA<br>TCCTCTGTTGCCAGCGGCA<br>TGGCCGGGAACTCAGAGGA<br>GACTGCCAGTATTAACTG<br>GAGGAAGGTGGGGATGAC<br>GTCAAGTCATCATGGCCCT<br>TATGACCAGGGCTACACAC<br>GTGCTACAATGGTGTATAC<br>AAAGAGAAGCAATCTCGCA<br>AGAGTAAGCAAACTCAA<br>AAGTACATCGTAGTTCGGA<br>TTAGAGTCTGCAACTCGAC<br>TCTATGAAGTAGGAATCGC<br>TAGTAATCGTGGATCAGAAT<br>GCCACGGTGAATACGTTCT<br>CTGGCCTTGACACACCGC<br>CCGTCACACCATGGGAGTA<br>AGTTGCAAAAGAAGTAGGT<br>AGCTTAACCTTTATAGGAG<br>GGCGCTTACCACTTTGTGA<br>TTTATGACTGGGGTGAAGT<br>CGTAACAAGGTAAGTGTAG<br>GGGAACCTGTGGTTGGATT<br>ACCTCCTTA<br>(SEQ ID NO: 14) |
| Baumannia<br>cicadellinicola | sharpshooter<br>leafhoppers | bacteriocytes | TTCAATTGAAGAGTTTGATC<br>ATGGCTCAGATTGAACGCT<br>GGCGGTAAGCTTAACACAT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |

|  |                    |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|--|--------------------|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | (Cicadellinae<br>) |  | GCAAGTCGAGCGGCATCG<br>GAAAGTAAATTAATTACTTT<br>GCCGGCAAGCGGCGAACG<br>GGTGAGTAATATCTGGGGA<br>TCTACCTTATGGAGAGGGA<br>TAACTATTGGAAACGATAGC<br>TAACACCGCATAATGTCGT<br>CAGACCAAAATGGGGGACC<br>TAATTTAGGCCTCATGCCAT<br>AAGATGAACCCAGATGAGA<br>TTAGCTAGTAGGTGAGATA<br>ATAGCTCACCTAGGCAACG<br>ATCTCTAGTTGGTCTGAGA<br>GGATGACCAGCCACACTGG<br>AACTGAGACACGGTCCAGA<br>CTCCTACGGGAGGCAGCAG<br>TGGGGAATCTTGCACAATG<br>GGGGAAACCCTGATGCAGC<br>TATACCGCGTGTGTGAAGA<br>AGGCCTTCGGGTGTGTAAG<br>CACTTTCAGCGGGGAAGAA<br>AATGAAGTTACTAATAATA<br>TTGTCAATTGACGTTACCCG<br>CAAAAGAAGCACCGGCTAA<br>CTCCGTGCCAGCAGCCGC<br>GGTAAGACGGAGGGTGCAA<br>GCGTTAATCGGAATTACTG<br>GGCGTAAAGCGTATGTAGG<br>CGGTTTATTTAGTCAGGTGT<br>GAAAGCCCTAGGCTTAACC<br>TAGGAATTGCATTTGAAACT<br>GGTAAGCTAGAGTCTCGTA<br>GAGGGGGGGAGAATTCCA<br>GGTGTAGCGGTGAAATGCG<br>TAGAGATCTGGAAGAATAC<br>CAGTGGCGAAGGCGCCCC<br>CCTGGACGAAAACGACGC<br>TCAAGTACGAAAGCGTGGG<br>GAGCAAACAGGATTAGATA<br>CCCTGGTAGTCCACGCTGT |
|--|--------------------|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | <div>AAACGATGTCGATTTGAAG<br/>GTTGTAGCCTTGAGCTATA<br/>GCTTTCGAAGCTAACGCAT<br/>TAAATCGACCGCCTGGGGA<br/>GTACGACCGCAAGGTTAAA<br/>ACTCAAATGAATTGACGGG<br/>GGCCCGCACAAAGCGGTGG<br/>AGCATGTGGTTTAATTCGAT<br/>ACAACGCGAAAAACCTTAC<br/>CTACTCTTGACATCCAGAGT<br/>ATAAAGCAGAAAAGCTTTAG<br/>TGCCTTCGGGAACTCTGAG<br/>ACAGGTGCTGCATGGCTGT<br/>CGTCAGCTCGTGTTGTGAA<br/>ATGTTGGGTTAAGTCCCGC<br/>AACGAGCGCAACCCTTATC<br/>CTTTGTTGCCAACGATTAAG<br/>TCGGGAACTCAAAGGAGAC<br/>TGCCGGTGATAAACCGGAG<br/>GAAGGTGAGGATAACGTCA<br/>AGTCATCATGGCCCTTACG<br/>AGTAGGGCTACACACGTGC<br/>TACAATGGTGCATACAAAG<br/>AGAAGCAATCTCGTAAGAG<br/>TTAGCAAACCTCATAAAGTG<br/>CATCGTAGTCCGGATTAGA<br/>GTCTGCAACTCGACTCTAT<br/>GAAGTCGGAATCGCTAGTA<br/>ATCGTGGATCAGAATGCCA<br/>CGGTGAATACGTTCCCGGG<br/>CCTTGTACACACCGCCCGT<br/>CACACCATGGGAGTGTATT<br/>GCAAAGAAGTTAGTAGCT<br/>TAACTCATAATACGAGAGG<br/>GCGCTTACCACTTTGTGATT<br/>CATAACTGGGGTGAAGTCG<br/>TAACAAGGTAACCGTAGGG<br/>GAACCTGCGGTTGGATCAC<br/>CTCCTTACACTAAA<br/>(SEQ ID NO: 15)</div> |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                     |                         |                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|---------------------|-------------------------|-------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Sodalis like</i> | Rhopalus<br>sapporensis | wider tissue<br>tropism | ATTGAACGCTGGCGGCAGG<br>CCTAACACATGCAAGTCGA<br>GCGGCAGCGGGAAGAAGC<br>TTGCTTCTTTGCCGGCGAG<br>CGGCGGACGGGTGAGTAAT<br>GTCTGGGGATCTGCCCGAT<br>GGAGGGGGATAACTACTGG<br>AAACGGTAGCTAATACCGC<br>ATAACGTCGCAAGACCAAA<br>GTGGGGGACCTTCGGGCC<br>TCACACCATCGGATGAACC<br>CAGGTGGGATTAGCTAGTA<br>GGTGGGGTAATGGCTCACC<br>TAGGCGACGATCCCTAGCT<br>GGTCTGAGAGGATGACCAG<br>TCACACTGGAAGTGAACA<br>CGGTCCAGACTCCTACGGG<br>AGGCAGCAGTGGGGAATAT<br>TGCACAATGGGGGAACCC<br>TGATGCAGCCATGCCGCGT<br>GTGTGAAGAAGGCCTTCGG<br>GTTGTAAAGCACTTTCAGC<br>GGGGAGGAAGGCGATGGC<br>GTTAATAGCGCTATCGATTG<br>ACGTTACCCGCAGAAGAAG<br>CACCGGCTAACTCCGTGCC<br>AGCAGCCGCGGTAATACGG<br>AGGGTGCGAGCGTTAATCG<br>GAATTACTGGGCGTAAAGC<br>GTACGCAGGCGGTCTGTTA<br>AGTCAGATGTGAAATCCCC<br>GGGCTCAACCTGGGAACTG<br>CATTTGAAACTGGCAGGCT<br>AGAGTCTCGTAGAGGGGG<br>GTAGAATTCCAGGTGTAGC<br>GGTGAAATGCGTAGAGATC<br>TGGAGGAATACCGGTGGCG<br>AAGGCGGCCCCCTGGACG<br>AAGACTGACGCTCAGGTAC<br>GAAAGCGTGGGGAGCAAAC |
|---------------------|-------------------------|-------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | <div>AGGATTAGATACCCTGGTA<br/>GTCCACGCTGTAAACGATG<br/>TCGATTTGAAGGTTGTGGC<br/>CTTGAGCCGTGGCTTTCGG<br/>AGCTAACGTGTTAAATCGA<br/>CCGCCTGGGGAGTACGGC<br/>CGCAAGGTTAAAACTCAAAT<br/>GAATTGACGGGGGCCCCGC<br/>ACAAGCGGTGGAGCATGTG<br/>GTTTAATTGATGCAACGC<br/>GAAGAACCTTACCTACTCTT<br/>GACATCCAGAGAACTTGGC<br/>AGAGATGCTTTGGTGCCTT<br/>CGGGAACCTCTGAGACAGGT<br/>GCTGCATGGCTGTCGTCAG<br/>CTCGTGTTGTGAAATGTTG<br/>GGTTAAGTCCCGCAACGAG<br/>CGCAACCCTTATCCTTTATT<br/>GCCAGCGATTCCGGTCGGGA<br/>ACTCAAAGGAGACTGCCGG<br/>TGATAAACCGGAGGAAGGT<br/>GGGGATGACGTCAAGTCAT<br/>CATGGCCCTTACGAGTAGG<br/>GCTACACACGTGCTACAAT<br/>GGCGCATACAAAGAGAAGC<br/>GATCTCGCGAGAGTCAGCG<br/>GACCTCATAAAGTGCGTCG<br/>TAGTCCGGATTGGAGTCTG<br/>CAACTCGACTCCATGAAGT<br/>CGGAATCGCTAGTAATCGT<br/>GGATCAGAATGCCACGGTG<br/>AATACGTTCCCGGGCCTTG<br/>TACACACCGCCCGTCACAC<br/>CATGGGAGTGGGTTGCAAA<br/>AGAAGTAGGTAGCTTAACC<br/>TTCGGGAGGGCGCTTACCA<br/>CTTTGTGATTCATGACTGG<br/>GGTG<br/>(SEQ ID NO: 16)</div> |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                          |                          |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|--------------------------|--------------------------|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Hartigia pinicola</i> | The pine<br>bark adelgid | bacteriocytes | AGATTTAACGCTGGCGGCA<br>GGCCTAACACATGCAAGTC<br>GAGCGGTACCAGAAGAAGC<br>TTGCTTCTTGCTGACGAGC<br>GGCGGACGGGTGAGTAAT<br>GTATGGGGATCTGCCCGAC<br>AGAGGGGGATAACTATTGG<br>AAACGGTAGCTAATACCGC<br>ATAATCTCTGAGGAGCAAA<br>GCAGGGGAACCTTCGGTCCT<br>TGCGCTATCGGATGAACCC<br>ATATGGGATTAGCTAGTAG<br>GTGAGGTAATGGCTCCCCT<br>AGGCAACGATCCCTAGCTG<br>GTCTGAGAGGATGATCAGC<br>CACACTGGGACTGAGACAC<br>GGCCCAGACTCCTACGGGA<br>GGCAGCAGTGGGGAATATT<br>GCACAATGGGCGAAAGCCT<br>GATGCAGCCATGCCGCGTG<br>TATGAAGAAGGCTTTAGGG<br>TTGTAAAGTACTTTCAGTCG<br>AGAGGAAAACATTGATGCT<br>AATATCATCAATTATTGACG<br>TTTCCGACAGAAGAAGCAC<br>CGGCTAACTCCGTGCCAGC<br>AGCCGCGGTAATACGGAGG<br>GTGCAAGCGTTAATCGGAA<br>TACTGGGCGTAAAGCGCA<br>CGCAGGCGGTTAATTAAGT<br>TAGATGTGAAAGCCCCGGG<br>CTTAACCCAGGAATAGCAT<br>ATAAACTGGTCAACTAGA<br>GTATTGTAGAGGGGGGTAG<br>AATTCCATGTGTAGCGGTG<br>AAATGCGTAGAGATGTGGA<br>GGAATACCAGTGGCGAAGG<br>CGGCCCCCTGGACAAAAC<br>TGACGCTCAAATGCGAAAG<br>CGTGGGGAGCAAACAGGAT |
|--------------------------|--------------------------|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|--|--|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | <div> <div> TAGATACCCTGGTAGTCCA TGCTGTAAACGATGTCGATT TGGAGGTTGTTCCCTTGAG GAGTAGCTTCCGTAGCTAA CGCGTTAAATCGACCGCCT GGGGGAGTACGACTGCAA GGTTAAAACTCAAATGAATT GACGGGGGCCCCGCACAAG CGGTGGAGCATGTGGTTTA ATTCGATGCAACGCGAAAA ACCTTACCTACTCTTGACAT CCAGATAATTTAGCAGAAAT GCTTTAGTACCTTCGGGAA ATCTGAGACAGGTGCTGCA TGGCTGTCGTCAGCTCGTG TTGTGAAATGTTGGGTAAAG TCCCGCAACGAGCGCAACC CTTATCCTTTGTTGCCAGCG ATTAGGTCGGGAACTCAAA GGAGACTGCCGGTGATAAA CCGGAGGAAGGTGGGGAT GACGTCAAGTCATCATGGC CCTTACGAGTAGGGCTACA CACGTGCTACAATGGCATA TACAAAGGGAAGCAACCTC GCGAGAGCAAGCGAAACTC ATAAATTATGTCGTAGTTCA GATTGGAGTCTGCAACTCG ACTCCATGAAGTCGGAATC GCTAGTAATCGTAGATCAG AATGCTACGGTGAATACGT TCCCGGGCCTTGTACACAC CGCCCGTCACACCATGGGA GTGGGTTGCAAAGAAGTA GGTAACTTAACCTTATGGAA AGCGCTTACCACTTTGTGAT TCATAACTGGGGTG </div> <div>(SEQ ID NO: 17)</div> </div> |
|--|--|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                   |                                         |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-----------------------------------|-----------------------------------------|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Wigglesworthia glossinidia</i> | tsetse fly<br>(Diptera:<br>Glossinidae) | bacteriocytes |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|                                   |                                         |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Beta proteobacteria               |                                         |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| <i>Tremblaya phenacola</i>        | Phenacoccus<br>avenae<br>(TPPAVE).      | bacteriomes   | AGGTAATCCAGCCACACCT<br>TCCAGTACGGCTACCTTGT<br>TACGACTTCACCCCAGTCA<br>CAACCCTTACCTTCGGAAC<br>TGCCCTCCTCACAACCTCAA<br>ACCACCAAACACTTTTAAAT<br>CAGGTTGAGAGAGGTTAGG<br>CCTGTTACTTCTGGCAAGA<br>ATTATTTCCATGGTGTGACG<br>GGCGGTGTGTACAAGACCC<br>GAGAACATATTCACCGTGG<br>CATGCTGATCCACGATTACT<br>AGCAATTCCAACCTTCATGCA<br>CTCGAGTTTCAGAGTACAAT<br>CCGAACTGAGGCCGGCTTT<br>GTGAGATTAGCTCCCTTTTG<br>CAAGTTGGCAACTCTTTGG<br>TCCGGCCATTGTATGATGT<br>GTGAAGCCCCACCCATAAA<br>GGCCATGAGGACTTGACGT<br>CATCCCCACCTTCCTCCAA<br>CTTATCGCTGGCAGTCTCTT<br>TAAGGTAAGTACTAATCCA<br>GTAGCAATTAAAGACAGGG<br>GTTGCGCTCGTTACAGGAC<br>TTAACCCAACATCTCACGAC<br>ACGAGCTGACGACAGCCAT<br>GCAGCACCTGTGCACTAAT<br>TCTCTTTCAAGCACTCCCG<br>CTTCTCAACAGGATCTTAGC<br>CATATCAAAGGTAGGTAAG<br>GTTTTTCGCGTTGCATCGAA<br>TTAATCCACATCATCCACTG<br>CTTGTGCGGGTCCCCGTCA<br>ATTCCTTTGAGTTTTAACCT |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|--|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | TGCGGCCGTACTCCCCAGG<br>CGGTCGACTTGTGCGTTAG<br>CTGCACCACTGAAAAGGAA<br>AACTGCCCAATGGTTAGTC<br>AACATCGTTTAGGGCATGG<br>ACTACCAGGGTATCTAATC<br>CTGTTTGCTCCCCATGCTTT<br>AGTGTCTGAGCGTCAGTAA<br>CGAACCAGGAGGCTGCCTA<br>CGCTTTCGGTATTCCTCCA<br>CATCTCTACACATTTCACTG<br>CTACATGCGGAATTCTACCT<br>CCCCCTCTCGTACTCCAGC<br>CTGCCAGTAACTGCCGCAT<br>TCTGAGGTTAAGCCTCAGC<br>CTTTCACAGCAATCTTAACA<br>GGCAGCCTGCACACCCTTT<br>ACGCCCAATAAATCTGATTA<br>ACGCTCGCACCCCTACGTAT<br>TACCGCGGCTGCTGGCACG<br>TAGTTTGCCGGTGCTTATTC<br>TTTCGGTACAGTCACACCA<br>CCAAATTGTTAGTTGGGTG<br>GCTTTCTTTCCGAACAAAAG<br>TGCTTTACAACCCAAAGGC<br>CTTCTTCACACACGCGGCA<br>TTGCTGGATCAGGCTTCCG<br>CCCATTGTCCAAGATTCCTC<br>ACTGCTGCCTTCCTCAGAA<br>GTCTGGGCCGTGTCTCAGT<br>CCCAGTGTGGCTGGCCGTC<br>CTCTCAGACCAGCTACCGA<br>TCATTGCCTTGGGAAGCCA<br>TTACCTTTCCAACAAGCTAA<br>TCAGACATCAGCCAATCTC<br>AGAGCGCAAGGCAATTGGT<br>CCCCTGCTTTCATTCTGCTT<br>GGTAGAGAACTTTATGCGG<br>TATTAATTAGGCTTTCACCT<br>AGCTGTCCCCCACTCTGAG |
|--|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                           |                                            |             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|---------------------------|--------------------------------------------|-------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                           |                                            |             | GCATGTTCTGATGCATTACT<br>CACCCGTTTGCCACTTGCC<br>ACCAAGCCTAAGCCCGTGT<br>TGCCGTTGCGACTTGCAATGT<br>GTAAGGCATGCCGCTAGCG<br>TTCAATCTGAGCCAGGATC<br>AAACTCT<br>(SEQ ID NO: 18)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <i>Tremblaya princeps</i> | citrus<br>mealybug<br>Planococcus<br>citri | bacteriomes | AGAGTTTGATCCTGGCTCA<br>GATTGAACGCTAGCGGCAT<br>GCATTACACATGCAAGTCG<br>TACGGCAGCACGGGCTTAG<br>GCCTGGTGGCGAGTGGCG<br>AACGGGTGAGTAACGCCTC<br>GGAACGTGCCTTGTAGTGG<br>GGGATAGCCTGGCGAAAGC<br>CAGATTAATACCGCATGAA<br>GCCGCACAGCATGCGCGG<br>TGAAAGTGGGGGATTCTAG<br>CCTCACGCTACTGGATCGG<br>CCGGGGTCTGATTAGCTAG<br>TTGGCGGGGTAAATGGCCCA<br>CCAAGGCTTAGATCAGTAG<br>CTGGTCTGAGAGGACGATC<br>AGCCACACTGGGACTGAGA<br>CACGGCCCAGACTCCTACG<br>GGAGGCAGCAGTGGGGAA<br>TCTTGGACAATGGGCGCAA<br>GCCTGATCCAGCAATGCCG<br>CGTGTGTGAAGAAGGCCTT<br>CGGGTCGTAAAGCACTTTT<br>GTTCGGGATGAAGGGGGG<br>CGTGCAAACACCATGCCCT<br>CTTGACGATACCGAAAGAA<br>TAAGCACCGGCTAACTACG<br>TGCCAGCAGCCGCGGTAAT<br>ACGTAGGGTGCGAGCGTTA<br>ATCGGAATCACTGGGCGTA<br>AAGGGTGCGCGGGTGGTTT<br>GCCAAGACCCCTGTAAAAT |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|--|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | CCTACGGCCCAACCGTAGT<br>GCTGCGGAGGTTACTGGTA<br>AGCTTGAGTATGGCAGAGG<br>GGGGTAGAATTCCAGGTGT<br>AGCGGTGAAATGCGTAGAT<br>ATCTGGAGGAATACCGAAG<br>GCGAAGGCAACCCCCTGG<br>GCCATCACTGACACTGAGG<br>CACGAAAGCGTGGGGAGC<br>AAACAGGATTAGATACCCT<br>GGTAGTCCACGCCCTAAAC<br>CATGTCGACTAGTTGTCGG<br>GGGGAGCCCTTTTTCCTCG<br>GTGACGAAGCTAACGCATG<br>AAGTCGACCGCCTGGGGA<br>GTACGACCGCAAGGTTAAA<br>ACTCAAAGGAATTGACGGG<br>GACCCGCACAAGCGGTGG<br>ATGATGTGGATTAATTCGAT<br>GCAACGCGAAAAACCTTAC<br>CTACCCTTGACATGGCGGA<br>GATTCTGCCGAGAGGCGGA<br>AGTGCTCGAAAGAGAATCC<br>GTGCACAGGTGCTGCATGG<br>CTGTCGTCAGCTCGTGTCTG<br>TGAGATGTTGGGTTAAGTC<br>CCATAACGAGCGCAACCCC<br>CGTCTTTAGTTGCTACCACT<br>GGGGCACTCTATAGAGACT<br>GCCGGTGATAAACCGGAGG<br>AAGGTGGGGACGACGTCAA<br>GTCATCATGGCCTTTATGG<br>GTAGGGCTTCACACGTCAT<br>ACAATGGCTGGAGCAAAGG<br>GTCGCCAACTCGAGAGAGG<br>GAGCTAATCCCACAAACCC<br>AGCCCCAGTTCGGATTGCA<br>CTCTGCAACTCGAGTGCAT<br>GAAGTCGGAATCGCTAGTA<br>ATCGTGGATCAGCATGCCA |
|--|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                  |                                                                                                     |             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|----------------------------------|-----------------------------------------------------------------------------------------------------|-------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                  |                                                                                                     |             | CGGTGAATACGTTCTCGGG<br>TCTTGTACACACCGCCCGT<br>CACACCATGGGAGTAAGCC<br>GCATCAGAAGCAGCCTCCC<br>TAACCCTATGCTGGGAAGG<br>AGGCTGCGAAGGTGGGGT<br>CTATGACTGGGGTGAAGTC<br>GTAACAAGGTAGCCGTACC<br>GGAAGGTGCGGCTGGATTA<br>CCT<br>(SEQ ID NO: 19)                                                                                                                                                                                                                                                                                                                                                                                                                |
|                                  |                                                                                                     |             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Vidania                          |                                                                                                     | bacteriomes |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <i>Nasuia deltocephalinicola</i> | pestiferous<br>insect host,<br>Macrosteles<br>quadripunctul<br>atus<br>(Hemiptera:<br>Cicadellidae) | bacteriomes | AGTTTAATCCTGGCTCAGAT<br>TTAACGCTTGCGACATGCC<br>TAACACATGCAAGTTGAAC<br>GTTGAAAATATTTCAAAGTA<br>GCGTATAGGTGAGTATAAC<br>ATTTAAACATACCTTAAAGT<br>TCGGAATACCCCGATGAAA<br>ATCGGTATAATACCGTATAA<br>AAGTATTTAAGAATTAAAGC<br>GGGGAAAACCTCGTGCTAT<br>AAGATTGTTAAATGCCTGAT<br>TAGTTTGTTGGTTTTTAAGG<br>TAAAAGCTTACCAAGACTTT<br>GATCAGTAGCTATTCTGTGA<br>GGATGTATAGCCACATTGG<br>GATTGAAATAATGCCCAAAC<br>CTCTACGGAGGGCAGCAGT<br>GGGGAATATTGGACAATGA<br>GCGAAAGCTTGATCCAGCA<br>ATGTCGCGTGTGCGATTAA<br>GGGAAACTGTAAAGCACTT<br>TTTTTTAAGAATAAGAAATTT<br>TAATTAATAATTAATTTT<br>GAATGTATTAAGAATAAG<br>TACCGACTAATCACGTGCC<br>AGCAGTCGCGGTAATACGT<br>GGGGTGCGAGCGTTAATCG |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|--|--|--|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | GATTTATTGGGCGTAAAGT<br>GTATTCAGGCTGCTTAAAAA<br>GATTTATATTAAATATTTAAA<br>TTAAATTTAAAAAATGTATAA<br>ATTACTATTAAGCTAGAGTT<br>TAGTATAAGAAAAAAGAATT<br>TTATGTGTAGCAGTGAAATG<br>CGTTGATATATAAAGGAAC<br>GCCGAAAGCGAAAGCATT<br>TTCTGTAATAGAACTGACGC<br>TTATATACGAAAGCGTG<br>TAGCAAACAGGATTAGATA<br>CCCTGGTAGTCCACGCCCT<br>AACTATGTCAATTA<br>TAGAATTTTTTTTAGTGGTG<br>TAGCTAACGCGTTAAATTGA<br>CCGCCTGGGTATTACGATC<br>GCAAGATTAAACTCAAAG<br>GAATTGACGGGGACCAGCA<br>CAAGCGGTGGATGATGTGG<br>ATTAATTCGATGATACGCGA<br>AAAACCTTACCTGCCCTTGA<br>CATGGTTAGAATTTTATTGA<br>AAAATAAAAGTGCTTGGAAA<br>AGAGCTAACACACAGGTGC<br>TGCATGGCTGTCGTCAGCT<br>CGTGTCGTGAGATGTTGGG<br>TTAAGTCCCGCAACGAGCG<br>CAACCCCTACTCTTAGTTGC<br>TAATTAAAGAACTTTAAGAG<br>AACAGCTAACAATAAGTTTA<br>GAGGAAGGAGGGGATGAC<br>TTCAAGTCCTCATGGCCCTT<br>ATGGGCAGGGCTTCACACG<br>TCATACAATGGTTAATACAA<br>AAAGTTGCAATATCGTAAGA<br>TTGAGCTAATCTTTAAAATT<br>AATCTTAGTTCGGATTGTAC<br>TCTGCAACTCGAGTACATG<br>AAGTTGGAATCGCTAGTAAT |
|--|--|--|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                     |                                                      |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|-------------------------------------|------------------------------------------------------|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                     |                                                      |               | CGCGGATCAGCATGCCGC<br>GGTGAATAGTTTAACTGGT<br>CTTGTACACACCGCCCGTC<br>ACACCATGGAAATAAATCTT<br>GTTTTAAATGAAGTAATATA<br>TTTTATCAAACAGGTTTTG<br>TAACCGGGGTGAAGTCGTA<br>ACA<br>(SEQ ID NO: 20)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| <i>Zinderia insecticola</i><br>CARI | spittlebug<br><i>Clastoptera</i><br><i>arizonana</i> | bacteriocytes | ATATAAATAAGAGTTTGATC<br>CTGGCTCAGATTGAACGCT<br>AGCGGTATGCTTTACACAT<br>GCAAGTCGAACGACAATAT<br>TAAAGCTTGCTTTAATATAA<br>AGTGGCGAACGGGTGAGTA<br>ATATATCAAACGTACCTTA<br>AAGTGGGGGATAACTAATT<br>GAAAAATTAGATAATACCGC<br>ATATTAATCTTAGGATGAAA<br>ATAGGAATAATATCTTATGC<br>TTTTAGATCGGTTGATATCT<br>GATTAGCTAGTTGGTAGGG<br>TAAATGCTTACCAAGGCAAT<br>GATCAGTAGCTGGTTTTAG<br>CGAATGATCAGCCCACTG<br>GAACTGAGACACGGTCCAG<br>ACTTCTACGGAAGGCAGCA<br>GTGGGGAATATTGGACAAT<br>GGGAGAAATCCTGATCCAG<br>CAATACCGCGTGAGTGATG<br>AAGGCCTTAGGGTCGTAAA<br>ACTCTTTTGTTAGGAAAGAA<br>ATAATTTTAAATAATATTAA<br>AATTGATGACGGTACCTAAA<br>GAATAAGCACCGGCTAACT<br>ACGTGCCAGCAGCCGCGG<br>TAATACGTAGGGTGCAAGC<br>GTTAATCGGAATTATTGGG<br>CGTAAAGAGTGCGTAGGCT<br>GTTATATAAGATAGATGTGA |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | <div>AATACTTAAGCTTAACTTAA<br/>GAACTGCATTTATTACTGTT<br/>TAACTAGAGTTTATTAGAGA<br/>GAAGTGGAATTTTATGTGTA<br/>GCAGTGAAATGCGTAGATA<br/>TATAAAGGAATATCGATGG<br/>CGAAGGCAGCTTCTTGGAA<br/>TAATACTGACGCTGAGGCA<br/>CGAAAGCGTGGGGAGCAAA<br/>CAGGATTAGATACCCTGGT<br/>AGTCCACGCCCTAAACTAT<br/>GTCTACTAGTTATTAAATTA<br/>AAAATAAAATTTAGTAACGT<br/>AGCTAACGCATTAAGTAGA<br/>CCGCCTGGGGAGTACGATC<br/>GCAAGATTAAAACTCAAAG<br/>GAATTGACGGGGACCCGCA<br/>CAAGCGGTGGATGATGTGG<br/>ATTAATTCGATGCAACACGA<br/>AAAACCTTACCTACTCTTGA<br/>CATGTTTGAATTTTAAAGA<br/>AATTTAAAAGTGCTTGAAAA<br/>AGAACC AAAACACAGGTGC<br/>TGCATGGCTGTCGTCAGCT<br/>CGTGTCGTGAGATGTTGGG<br/>TTAAGTCCCGCAACGAGCG<br/>CAACCCTTGTTATTATTTGC<br/>TAATAAAAAGAACTTTAATA<br/>AGACTGCCAATGACAAATT<br/>GGAGGAAGGTGGGGATGA<br/>CGTCAAGTCCTCATGGCCC<br/>TTATGAGTAGGGCTTCACA<br/>CGTCATACAATGATATATAC<br/>AATGGGTAGCAAATTTGTG<br/>AAAATGAGCCAATCCTTAAA<br/>GTATATCTTAGTTCGGATTG<br/>TAGTCTGCAACTCGACTAC<br/>ATGAAGTTGGAATCGCTAG<br/>TAATCGCGGATCAGCATGC<br/>CGCGGTGAATACGTTCTCG</div> |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                           |                                                    |             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|---------------------------|----------------------------------------------------|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                           |                                                    |             | GGTCTTGTACACACCGCCC<br>GTCACACCATGGAAGTGAT<br>TTTTACCAGAAATTATTTGT<br>TTAACCTTTATTGGAAAAAA<br>ATAATTAAGGTAGAATTCAT<br>GACTGGGGTGAAGTCGTAA<br>CAAGGTAGCAGTATCGGAA<br>GGTGCGGCTGGATTACATT<br>TTAAAT<br>(SEQ ID NO: 21)                                                                                                                                                                                                                                                                                                                                                    |
| <i>Proffella armatura</i> | <i>Diaphorina citri</i> , the Asian citrus psyllid | bacteriomes |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|                           |                                                    |             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Alpha proteobacteria      |                                                    |             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Hodgkinia                 | Cicada<br><i>Diceroprocta semicincta</i>           | bacteriome  | AATGCTGGCGGCAGGCCTA<br>ACACATGCAAGTCGAGCGG<br>ACAACGTTCAAACGTTGTTA<br>GCGGCGAACGGGTGAGTA<br>ATACGTGAGAATCTACCCAT<br>CCCAACGTGATAACATAGT<br>CAACACCATGTCAATAACGT<br>ATGATTCCTGCAACAGGTA<br>AAGATTTTATCGGGGATGG<br>ATGAGCTCACGCTAGATTA<br>GCTAGTTGGTGAGATAAAA<br>GCCCACCAAGGCCAAGATC<br>TATAGCTGGTCTGGAAGGA<br>TGGACAGCCACATTGGGAC<br>TGAGACAAGGCCCAACCCT<br>CTAAGGAGGGCAGCAGTGA<br>GGAATATTGGACAATGGGC<br>GTAAGCCTGATCCAGCCAT<br>GCCGCATGAGTGATTGAAG<br>GTCCAACGGACTGTAAAC<br>TCTTTTCTCCAGAGATCATA<br>AATGATAGTATCTGGTGATA<br>TAAGCTCCGGCCAACTTCG<br>TGCCAGCAGCCGCGGTAAT |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | ACGAGGGGAGCGAGTATTG<br>TTCGGTTTTATTGGGCGTAA<br>AGGGTGTCCAGGTTGCTAA<br>GTAAGTTAACAACAAAATCT<br>TGAGATTCAACCTCATAACG<br>TTCGGTTAATACTACTAAGC<br>TCGAGCTTGGATAGAGACA<br>AACGGAATTCCGAGTGTAG<br>AGGTGAAATTCGTTGATACT<br>TGGAGGAACACCAGAGGC<br>GAAGGCGGTTTGTCATACC<br>AAGCTGACACTGAAGACAC<br>GAAAGCATGGGGAGCAAAC<br>AGGATTAGATACCCTGGTA<br>GTCCATGCCCTAAACGTTG<br>AGTGCTAACAGTTCGATCA<br>AGCCACATGCTATGATCCA<br>GGATTGTACAGCTAACGCG<br>TTAAGCACTCCGCCTGGGT<br>ATTACGACCGCAAGGTAA<br>AACTCAAAGGAATTGACGG<br>AGACCCGCACAAGCGGTG<br>GAGCATGTGGTTTAATTG<br>AAGCTACACGAAGAACCTT<br>ACCAGCCCTTGACATACCA<br>TGGCCAACCATCCTGGAAA<br>CAGGATGTTGTTCAAGTTAA<br>ACCCTTGAAATGCCAGGAA<br>CAGGTGCTGCATGGCTGTT<br>GTCAGTTCGTGTCGTGAGA<br>TGTATGGTTAAGTCCCAAAA<br>CGAACACAACCCTCACCCA<br>TAGTTGCCATAAACACAATT<br>GGGTTCTCTATGGGTACTG<br>CTAACGTAAGTTAGAGGAA<br>GGTGAGGACCACAACAAGT<br>CATCATGGCCCTTATGGGC<br>TGGGCCACACACATGCTAC<br>AATGGTGGTTACAAAGAGC<br>CGCAACGTTGTGAGACCGA |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                           |                                                                |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|---------------------------|----------------------------------------------------------------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                           |                                                                |            | GCAAATCTCCAAAGACCAT<br>CTCAGTCCGGATTGTACTC<br>TGCAACCCGAGTACATGAA<br>GTAGGAATCGCTAGTAATC<br>GTGGATCAGCATGCCACGG<br>TGAATACGTTCTCGGGTCTT<br>GTACACGCCGCCCGTCACA<br>CCATGGGAGCTTCGCTCCG<br>ATCGAAGTCAAGTTACCCTT<br>GACCACATCTTGGCAAGTG<br>ACCGA<br>(SEQ ID NO: 22)                                                                                                                                                                                                                                                                                                                                                                                                          |
| <i>Wolbachia sp. wPip</i> | Mosquito<br><i>Culex</i><br><i>quinquefasci</i><br><i>atus</i> | bacteriome | AAATTTGAGAGTTTGATCCT<br>GGCTCAGAATGAACGCTGG<br>CGGCAGGCCTAACACATGC<br>AAGTCGAACGGAGTTATATT<br>GTAGCTTGCTATGGTATAAC<br>TTAGTGGCAGACGGGTGAG<br>TAATGTATAGGAATCTACCT<br>AGTAGTACGGAATAATTGTT<br>GGAAACGACAACTAATACC<br>GTATACGCCCTACGGGGGA<br>AAAATTTATTGCTATTAGAT<br>GAGCCTATATTAGATTAGCT<br>AGTTGGTGGGGTAATAGCC<br>TACCAAGGTAATGATCTATA<br>GCTGATCTGAGAGGATGAT<br>CAGCCACACTGGAAGTGAAG<br>ATACGGTCCAGACTCCTAC<br>GGGAGGCAGCAGTGGGGA<br>ATATTGGACAATGGGCGAA<br>AGCCTGATCCAGCCATGCC<br>GCATGAGTGAAGAAGGCCT<br>TTGGGTTGTAAAGCTCTTTT<br>AGTGAGGAAGATAATGACG<br>GTAATCACAGAAGAAGTCC<br>TGGCTAACTCCGTGCCAGC<br>AGCCGCGGTAATACGGAGA<br>GGGCTAGCGTTATTCGGAA<br>TTATTGGGCGTAAAGGGCG |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|--|--|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | CGTAGGCTGGTTAATAAGT<br>TAAAAGTGAAATCCCGAGG<br>CTTAACCTTGGAATTGCTTT<br>TAAAACTATTAATCTAGAGA<br>TTGAAAGAGGATAGAGGAA<br>TTCCTGATGTAGAGGTAAAA<br>TTCGTAAATATTAGGAGGAA<br>CACCAGTGGCGAAGGCGTC<br>TATCTGGTTCAAATCTGACG<br>CTGAAGCGCGAAGGCGTG<br>GGGAGCAAACAGGATTAGA<br>TACCCTGGTAGTCCACGCT<br>GTAAACGATGAATGTTAAAT<br>ATGGGGAGTTTACTTTCTGT<br>ATTACAGCTAACGCGTTAAA<br>CATTCCGCCTGGGGACTAC<br>GGTCGCAAGATTAAAACTC<br>AAAGGAATTGACGGGGACC<br>CGCACAAGCGGTGGAGCAT<br>GTGGTTTAATTCGATGCAAC<br>GCGAAAAACCTTACCACTT<br>CTTGACATGAAAATCATACC<br>TATTCGAAGGGATAGGGTC<br>GGTTCGGCCGGATTTTACA<br>CAAGTGTTGCATGGCTGTC<br>GTCAGCTCGTGTCGTGAGA<br>TGTTGGGTTAAGTCCCGCA<br>ACGAGCGCAACCCTCATCC<br>TTAGTTGCCATCAGGTAATG<br>CTGAGTACTTTAAGGAAACT<br>GCCAGTGATAAGCTGGAGG<br>AAGGTGGGGATGATGTCAA<br>GTCATCATGGCCTTTATGG<br>AGTGGGCTACACACGTGCT<br>ACAATGGTGTCTACAATGG<br>GCTGCAAGGTGCGCAAGCC<br>TAAGCTAATCCCTAAAAGAC<br>ATCTCAGTTCGGATTGTACT<br>CTGCAACTCGAGTACATGA<br>AGTTGGAATCGCTAGTAAT |
|--|--|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                             |                           |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-----------------------------|---------------------------|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                             |                           |               | CGTGGATCAGCATGCCACG<br>GTGAATACGTTCTCGGGTC<br>TTGTACACACTGCCCCGTCA<br>CGCCATGGGAATTGGTTTC<br>ACTCGAAGCTAATGGCCTA<br>ACCGCAAGGAAGGAGTTAT<br>TTAAAGTGGGATCAGTGAC<br>TGGGGTGAAGTCGTAACAA<br>GGTAGCAGTAGGGGAATCT<br>GCAGCTGGATTACCTCCTT<br>A<br>(SEQ ID NO: 23)                                                                                                                                                                                                                                                                                                                                                              |
|                             |                           |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Bacteroidetes               |                           |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <i>Uzinura diaspidicola</i> | armoured<br>scale insects | bacteriocytes | AAAGGAGATATTCCAACCA<br>CACCTTCCGGTACGGTTAC<br>CTTGTTACGACTTAGCCCTA<br>GTCATCAAGTTTACCTTAGG<br>CAGACCACTGAAGGATTAC<br>TGACTTCAGGTACCCCCGA<br>CTCCCATGGCTTGACGGGC<br>GGTGTGTACAAGGTTCGAG<br>AACATATTCACCGCGCCATT<br>GCTGATGCGCGATTACTAG<br>CGATTCCTGCTTCATAGAGT<br>CGAATTGCAGACTCCAATC<br>CGAACTGAGACTGGTTTTA<br>GAGATTAGCTCCTGATCAC<br>CCAGTGGCTGCCCTTTGTA<br>ACCAGCCATTGTAGCACGT<br>GTGTAGCCCAAGGCATAGA<br>GGCCATGATGATTTGACAT<br>CATCCCCACCTTCCTCACA<br>GTTTACACCGGCAGTTTTGT<br>TAGAGTCCCCGGCTTTACC<br>CGATGGCAACTAACAATAG<br>GGGTTGCGCTCGTTATAGG<br>ACTTAACCAAACACTTCACA<br>GCACGAACTGAAGACAACC<br>ATGCAGCACCTTGTAATAC |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|--|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | GTCGTATAGACTAAGCTGTT<br>TCCAGCTTATTTCGTAATACA<br>TTTAAGCCTTGGTAAGGTTC<br>CTCGCGTATCATCGAATTAA<br>ACCACATGCTCCACCGCTT<br>GTGCGAACCCCCGTCAATT<br>CCTTTGAGTTTCAATCTTGC<br>GACTGTACTTCCCAGGTGG<br>ATCACTTATCGCTTTCGCTA<br>AGCCACTGAATATCGTTTTT<br>CCAATAGCTAGTGATCATC<br>GTTTAGGGCGTGGACTACC<br>AGGGTATCTAATCCTGTTTG<br>CTCCCCACGCTTTCGTGCA<br>CTGAGCGTCAGTAAAGATT<br>TAGCAACCTGCCTTCGCTA<br>TCGGTGTTCTGTATGATATC<br>TATGCATTTACCGCTACAC<br>CATACATTCCAGATGCTCCA<br>ATCTTACTCAAGTTTACCAG<br>TATCAATAGCAATTTTACAG<br>TTAAGCTGTAAGCTTTCACT<br>ACTGACTTAATAAACAGCCT<br>ACACACCCTTTAAACCCAAT<br>AAATCCGAATAACGCTTGT<br>GTCATCCGTATTGCCGCGG<br>CTGCTGGCACGGAATTAGC<br>CGACACTTATTCGTATAGTA<br>CCTTCAATCTCCTATCACGT<br>AAGATATTTTATTTCTATACA<br>AAAGCAGTTTACAACCTAAA<br>AGACCTTCATCCTGCACGC<br>GACGTAGCTGGTTCAGAGT<br>TTCCTCCATTGACCAATATT<br>CCTCACTGCTGCCTCCCGT<br>AGGAGTCTGGTCCGTGTCT<br>CAGTACCAGTGTGGAGGTA<br>CACCTCTTAGGCCCCCTA<br>CTGATCATAGTCTTGGTAGA<br>GCCATTACCTCACCAACTAA |
|--|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                        |                                                                             |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|------------------------|-----------------------------------------------------------------------------|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                        |                                                                             |               | CTAATCAAACGCAGGCTCA<br>TCTTTTGCCACCTAAGTTTT<br>AATAAAGGCTCCATGCAGA<br>AACTTTATATTATGGGGGAT<br>TAATCAGAATTTCTTCTGGC<br>TATACCCCAGCAAAAGGTA<br>GATTGCATACGTGTTACTCA<br>CCCATTGCGCCGGTCGCCGA<br>CAAATTAAAAATTTTTTCGAT<br>GCCCCTCGACTTGCATGTG<br>TTAAGCTCGCCGCTAGCGT<br>TAATTCTGAGCCAGGATCA<br>AACTCTTCGTTGTAG<br>(SEQ ID NO: 24)                                                                                                                                                                                                                                                                                                 |
| <i>Sulcia muelleri</i> | Blue-Green<br>Sharpshooter<br>and several<br>other<br>leafhopper<br>species | bacteriocytes | CTCAGGATAAACGCTAGCG<br>GAGGGCTTAACACATGCAA<br>GTCGAGGGGGCAGCAAAAAT<br>AATTATTTTTTGGCGACCGG<br>CAAACGGGTGAGTAATACA<br>TACGTAACCTTTCCTTATGCT<br>GAGGAATAGCCTGAGGAAA<br>CTTGGATTAATACCTCATAA<br>TACAATTTTTTTAGAAAGAAA<br>AATTGTTAAAGTTTTATTAT<br>GGCATAAGATAGGCGTATG<br>TCCAATTAGTTAGTTGGTAA<br>GGTAATGGCTTACCAAGAC<br>GATGATTGGTAGGGGGCCT<br>GAGAGGGGGCGTTCCCCCA<br>CATTGGTACTGAGACACGG<br>ACCAACTTCTACGGAAGG<br>CTGCAGTGAGGAATATTGG<br>TCAATGGAGGAACTCTGA<br>ACCAGCCACTCCGCGTGCA<br>GGATGAAAGAAAGCCTTAT<br>TGGTTGTAACTGCTTTTGT<br>ATATGAATAAAAAATTCTAA<br>TTATAGAAATAATTGAAGGT<br>AATATACGAATAAGTATCGA<br>CTAACCTCTGTGCCAGCAGT |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|--|--|--|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | CGCGGTAAGACAGAGGATA<br>CAAGCGTTATCCGGATTTAT<br>TGGGTTTAAAGGGTGCGTA<br>GGCGGTTTTTAAAGTCAGT<br>AGTGAAATCTTAAAGCTTAA<br>CTTTAAAAGTGCTATTGATA<br>CTGAAAACTAGAGTAAGG<br>TTGGAGTAACTGGAATGTG<br>TGGTGTAGCGGTGAAATGC<br>ATAGATATCACACAGAACAC<br>CGATAGCGAAAGCAAGTTA<br>CTAACCCTATACTGACGCT<br>GAGTCACGAAAGCATGGGG<br>AGCAAACAGGATTAGATAC<br>CCTGGTAGTCCATGCCGTA<br>AACGATGATCACTAACTATT<br>GGGTTTTATACGTTGTAATT<br>CAGTGGTGAAGCGAAAGTG<br>TTAAGTGATCCACCTGAGG<br>AGTACGACCGCAAGGTTGA<br>AACTCAAAGGAATTGACGG<br>GGGCCCCGCACAATCGGTG<br>GAGCATGTGGTTTAATTCTG<br>ATGATACACGAGGAACCTT<br>ACCAAGACTTAAATGTACTA<br>CGAATAAATTGGAAACAATT<br>TAGTCAAGCGACGGAGTAC<br>AAGGTGCTGCATGGTTGTC<br>GTCAGCTCGTGCCGTGAGG<br>TGTAAGGTTAAGTCCTTTAA<br>ACGAGCGCAACCCTTATTA<br>TTAGTTGCCATCGAGTAATG<br>TCAGGGGACTCTAATAAGA<br>CTGCCGGCGCAAGCCGAG<br>AGGAAGGTGGGGATGACGT<br>CAAATCATCACGGCCCTTA<br>CGTCTTGGGCCACACACGT<br>GCTACAATGATCGGTACAA<br>AAGGGAGCGACTGGGTGA<br>CCAGGAGCAAATCCAGAAA |
|--|--|--|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                                             |                                                           |                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|-------------------------------------------------------------|-----------------------------------------------------------|--------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                             |                                                           |                                                  | GCCGATCTAAGTTCGGATT<br>GGAGTCTGAAACTCGACTC<br>CATGAAGCTGGAATCGCTA<br>GTAATCGTGCATCAGCCAT<br>GGCACGGTGAATATGTTCC<br>CGGGCCTTGTACACACCGC<br>CCGTCAAGCCATGGAAGTT<br>GGAAGTACCTAAAGTTGGT<br>TCGCTACCTAAGGTAAGTC<br>TAATAACTGGGGCTAAGTC<br>GTAACAAGGTA<br>(SEQ ID NO: 25)                                                                                                                                                                                                                                                                                                                                                           |
|                                                             |                                                           |                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Yeast like                                                  |                                                           |                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| <i>Symbiotaphrina</i><br><i>buchneri</i> voucher<br>JCM9740 | Anobiid<br>beetles<br><i>Stegobium</i><br><i>paniceum</i> | mycetome<br>between the<br>foregut and<br>midgut | AGATTAAGCCATGCAAGTC<br>TAAGTATAAGNAATCTATAC<br>NGTGAAACTGCGAATGGCT<br>CATTAAATCAGTTATCGTTT<br>ATTTGATAGTACCTTACTAC<br>ATGGATAACCGTGGTAATT<br>CTAGAGCTAATACATGCTAA<br>AAACCCCGACTTCGGAAGG<br>GGTGTATTTATTAGATAAAA<br>AACCAATGCCCTTCGGGGC<br>TCCTTGGTGATTCATGATAA<br>CTTAACGAATCGCATGGCC<br>TTGCGCCGGCGATGGTTCA<br>TTCAAATTTCTGCCCTATCA<br>ACTTTCGATGGTAGGATAG<br>TGGCCTACCATGGTTTTAAC<br>GGGTAACGGGGAATTAGGG<br>TTCGATTCCGGAGAGGGAG<br>CCTGAGAAACGGCTACCAC<br>ATCCAAGGAAGGCAGCAGG<br>CGCGCAAATTACCCAATCC<br>CGACACGGGGAGGTAAGTG<br>ACAATAAATACTGATACAGG<br>GCTCTTTTGGGTCTTGTAAT<br>TGGAATGAGTACAATTTAAA<br>TCCCTTAACGAGGAACAATT |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|--|--|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | GGAGGGCAAGTCTGGTGC<br>CAGCAGCCGCGGTAATTCC<br>AGCTCCAATAGCGTATATTA<br>AAGTTGTTGCAGTTAAAAAG<br>CTCGTAGTTGAACCTTGGG<br>CCTGGCTGGCCGGTCCGC<br>CTAACCGCGTGTACTGGTC<br>CGGCCGGGCCTTTCCTTCT<br>GGGGAGCCGCATGCCCTTC<br>ACTGGGTGTGTCTGGGGAAC<br>CAGGACTTTTACTTTGAAAA<br>AATTAGAGTGTTCAAAGCA<br>GGCCTATGCTCGAATACAT<br>TAGCATGGAATAATAGAATA<br>GGACGTGCGGTTCTATTTT<br>GTTGGTTTCTAGGACCGCC<br>GTAATGATTAATAGGGATAG<br>TCGGGGGCATCAGTATTCA<br>ATTGTCAGAGGTGAAATTCT<br>TGGATTTATTGAAGACTAAC<br>TACTGCGAAAGCATTTGCC<br>AAGGATGTTTTCATTAATCA<br>GTGAACGAAAGTTAGGGGA<br>TCGAAGACGATCAGATACC<br>GTCGTAGTCTTAACCATAAA<br>CTATGCCGACTAGGGATCG<br>GGCGATGTTATTATTTTGAC<br>TCGCTCGGCACCTTACGAG<br>AAATCAAAGTCTTTGGGTTC<br>TGGGGGGAGTATGGTCGCA<br>AGGCTGAAACTTAAAGAAAT<br>TGACGGAAGGGCACCACCA<br>GGAGTGGAGCCTGCGGCTT<br>AATTTGACTCAACACGGGG<br>AAACTCACCAGGTCCAGAC<br>ACATTAAGGATTGACAGATT<br>GAGAGCTCTTTCTTGATTAT<br>GTGGGTGGTGGTGCATGG<br>CCGTTCTTAGTTGGTGGAG<br>TGATTTGTCTGCTTAATTGC |
|--|--|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                                    |                                                              |          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|----------------------------------------------------|--------------------------------------------------------------|----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                    |                                                              |          | GATAACGAACGAGACCTTA<br>ACCTGCTAAATAGCCCGGT<br>CCGCTTTGGCGGGCCGCT<br>GGCTTCTTAGAGGGACTAT<br>CGGCTCAAGCCGATGGAAG<br>TTTGAGGCAATAACAGGTC<br>TGTGATGCCCTTAGATGTTC<br>TGGGCCGCACGCGCGCTA<br>CACTGACAGAGCCAACGAG<br>TAAATCACCTTGGCCGGAA<br>GGTCTGGGTAATCTTGTTAA<br>ACTCTGTCGTGCTGGGGAT<br>AGAGCATTGCAATTATTGCT<br>CTTCAACGAGGAATTCCTA<br>GTAAGCGCAAGTCATCAGC<br>TTGCGCTGATTACGTCCCT<br>GCCCTTTGTACACACCGCC<br>CGTCGCTACTACCGATTGA<br>ATGGCTCAGTGAGGCCTTC<br>GGACTGGCACAGGGACGTT<br>GGCAACGACGACCCAGTGC<br>CGGAAAGTTGGTCAAACCTT<br>GGTCATTTAGAGGAAGTAA<br>AAGTCGTAACAAGGTTTCC<br>GTAGGTGAACCTGCGGAAG<br>GATCATTA<br>(SEQ ID NO: 26) |
| <i>Symbiotaphrina kochii</i><br>voucher CBS 589.63 | Anobiid<br>beetles<br><i>Lasioderma</i><br><i>serricorne</i> | mycetome | TACCTGGTTGATTCTGCCA<br>GTAGTCATATGCTTGTCTCA<br>AAGATTAAGCCATGCAAGT<br>CTAAGTATAAGCAATCTATA<br>CGGTGAAACTGCGAATGGC<br>TCATTAAATCAGTTATCGTT<br>TATTTGATAGTACCTTACTA<br>CATGGATAACCGTGGTAAT<br>TCTAGAGCTAATACATGCTA<br>AAAACCTCGACTTCGGAAG<br>GGGTGTATTTATTAGATAAA<br>AAACCAATGCCCTTCGGGG<br>CTCCTTGGTGATTCATGATA                                                                                                                                                                                                                                                                                                               |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|--|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | ACTTAACGAATCGCATGGC<br>CTTGCGCCGGCGATGGTTC<br>ATTCAAATTTCTGCCCTATC<br>AACTTTCGATGGTAGGATA<br>GTGGCCTACCATGGTTTCA<br>ACGGGTAACGGGGAATTAG<br>GGTTCGATTCCGGAGAGGG<br>AGCCTGAGAAACGGCTACC<br>ACATCCAAGGAAGGCAGCA<br>GGCGCGCAAATTACCCAAT<br>CCCGACACGGGGAGGTAG<br>TGACAATAAATACTGATACA<br>GGGCTCTTTTGGGTCTTGT<br>AATTGGAATGAGTACAATTT<br>AAATCCCTTAACGAGGAAC<br>AATTGGAGGGCAAGTCTGG<br>TGCCAGCAGCCGCGGTAAT<br>TCCAGCTCCAATAGCGTAT<br>ATTAAAGTTGTTGCAGTTAA<br>AAAGCTCGTAGTTGAACCTT<br>GGGCCTGGCTGGCCGGTC<br>CGCCTAACCGCGTGTACTG<br>GTCCGGCCGGGCCTTTCCT<br>TCTGGGGAGCCGCATGCCC<br>TTCACTGGGTGTGTCGGGG<br>AACCAGGACTTTTACTTTGA<br>AAAAATTAGAGTGTTCAAAG<br>CAGGCCTATGCTCGAATAC<br>ATTAGCATGGAATAATAGAA<br>TAGGACGTGTGGTTCTATTT<br>TGTTGGTTTCTAGGACCGC<br>CGTAATGATTAATAGGGATA<br>GTCGGGGGCATCAGTATTC<br>AATTGTCAGAGGTGAAATTC<br>TTGGATTTATTGAAGACTAA<br>CTACTGCGAAAGCATTTGC<br>CAAGGATGTTTTCATTAATC<br>AGTGAACGAAAGTTAGGGG<br>ATCGAAGACGATCAGATAC<br>CGTCGTAGTCTTAACCATAA |
|--|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|--|--|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | ACTATGCCGACTAGGGATC<br>GGGCGATGTTATTATTTTGA<br>CTCGCTCGGCACCTTACGA<br>GAAATCAAAGTCTTTGGGT<br>CTGGGGGGAGTATGGTCG<br>CAAGGCTGAAACTTAAAGA<br>AATTGACGGAAGGGCACCA<br>CCAGGAGTGGAGCCTGCG<br>GCTTAATTTGACTCAACACG<br>GGGAAACTCACCAGGTCCA<br>GACACATTAAGGATTGACA<br>GATTGAGAGCTCTTTCTTGA<br>TTATGTGGGTGGTGGTGCA<br>TGGCCGTTCTTAGTTGGTG<br>GAGTGATTTGTCTGCTTAAT<br>TGCGATAACGAACGAGACC<br>TTAACCTGCTAAATAGCCC<br>GGTCCGCTTTGGCGGGCC<br>GCTGGCTTCTTAGAGGGAC<br>TATCGGCTCAAGCCGATGG<br>AAGTTTGAGGCAATAACAG<br>GTCTGTGATGCCCTTAGAT<br>GTTCTGGGCCGCACGCGC<br>GCTACACTGACAGAGCCAA<br>CGAGTACATCACCTTGGCC<br>GGAAGGTCTGGGTAATCTT<br>GTTAAACTCTGTCGTGCTG<br>GGGATAGAGCATTGCAATT<br>ATTGCTCTTCAACGAGGAAT<br>TCCTAGTAAGCGCAAGTCA<br>TCAGCTTGCGCTGATTACG<br>TCCCTGCCCTTTGTACACA<br>CCGCCCGTCGCTACTACCG<br>ATTGAATGGCTCAGTGAGG<br>CCTTCGGACTGGCACAGGG<br>ACGTTGGCAACGACGACCC<br>AGTGCCGGAAAGTTCGTCA<br>AACTTGGTCATTTAGAGGAA<br>GNNNAAGTCGTAACAAGGT |
|--|--|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                           |                                |                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-------------------------------------------|--------------------------------|-----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                           |                                |                 | TTCCGTAGGTGAACCTGCG<br>GAAGGATCATTA<br>(SEQ ID NO: 27)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|                                           |                                |                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| <b>Primary extracellular<br/>symbiont</b> | <b>Host</b>                    | <b>location</b> | <b>16 rRNA</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| fenitrothion-degrading<br>bacteria        |                                |                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| <i>Burkholderia sp. SFA1</i>              | <i>Riptortus<br/>pedestris</i> | Gut             | AGTTTGATCCTGGCTCAGA<br>TTGAACGCTGGCGGCATGC<br>CTTACACATGCAAGTCGAA<br>CGGCAGCACGGGGGCAAC<br>CCTGGTGGCGAGTGGCGA<br>ACGGGTGAGTAATACATCG<br>GAACGTGTCCTGTAGTGGG<br>GGATAGCCCGGCGAAAGC<br>CGGATTAATACCGCATACG<br>ACCTAAGGGAGAAAGCGGG<br>GGATCTTCGGACCTCGCGC<br>TATAGGGGCGGCCGATGG<br>CAGATTAGCTAGTTGGTGG<br>GGTAAAGGCCTACCAAGGC<br>GACGATCTGTAGCTGGTCT<br>GAGAGGACGACCAGCCACA<br>CTGGGACTGAGACACGGCC<br>CAGACTCCTACGGGAGGCA<br>GCAGTGGGGAATTTTGGAC<br>AATGGGGGCAACCCTGATC<br>CAGCAATGCCGCGTGTGTG<br>AAGAAGGCTTCGGGTTGTA<br>AAGCACTTTTGTCCGAAA<br>GAAAACCTTCGTCCCTAATAT<br>GGATGGAGGATGACGGTAC<br>CGGAAGAATAAGCACCGGC<br>TAACTACGTGCCAGCAGCC<br>GCGGTAATACGTAGGGTGC<br>GAGCGTTAATCGGAATTAC<br>TGGGCGTAAAGCGTGCGCA<br>GGCGGTCTGTTAAGACCGA<br>TGTGAAATCCCCGGGCTTA |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|--|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | ACCTGGGAACTGCATTGGT<br>GACTGGCAGGCTTTGAGTG<br>TGGCAGAGGGGGGTAGAAT<br>TCCACGTGTAGCAGTGAAA<br>TGCGTAGAGATGTGGAGGA<br>ATACCGATGGCGAAGGCAG<br>CCCCCTGGGCCAACTACTG<br>ACGCTCATGCACGAAAGCG<br>TGGGGAGCAAACAGGATTA<br>GATACCCTGGTAGTCCACG<br>CCCTAAACGATGTCAACTA<br>GTTGTTGGGGATTTCATTTCC<br>TTAGTAACGTAGCTAACGC<br>GTGAAGTTGACCGCCTGGG<br>GAGTACGGTCGCAAGATTA<br>AAACTCAAAGGAATTGACG<br>GGGACCCGCACAAGCGGT<br>GGATGATGTGGATTAATTC<br>GATGCAACGCGAAAAACCT<br>TACCTACCCTTGACATGGT<br>CGGAACCCTGCTGAAAGGT<br>GGGGGTGCTCGAAAGAGAA<br>CCGGCGCACAGGTGCTGC<br>ATGGCTGTCTCAGCTCGT<br>GTCGTGAGATGTTGGGTTA<br>AGTCCCGCAACGAGCGCAA<br>CCCTTGTCCTTAGTTGCTAC<br>GCAAGAGCACTCTAAGGAG<br>ACTGCCGGTGACAAACCGG<br>AGGAAGGTGGGGATGACGT<br>CAAGTCCTCATGGCCCTTA<br>TGGGTAGGGCTTCACACGT<br>CATACAATGGTCGGAACAG<br>AGGGTTGCCAAGCCGCGA<br>GGTGGAGCCAATCCCAGAA<br>AACCGATCGTAGTCCGGAT<br>CGCAGTCTGCAACTCGACT<br>GCGTGAAGCTGGAATCGCT<br>AGTAATCGCGGATCAGCAT<br>GCCGCGGTGAATACGTTCC |
|--|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                              |                            |     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|------------------------------|----------------------------|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                              |                            |     | CGGGTCTTGTACACACCGC<br>CCGTCACACCATGGGAGTG<br>GGTTTCACCAGAAGTAGGT<br>AGCCTAACCGCAAGGAGGG<br>CGCTTACCACGGTGGGATT<br>CATGACTGGGGTGAAGTCG<br>TAACAAGGTAGC<br>(SEQ ID NO: 28)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| <i>Burkholderia sp. KM-A</i> | <i>Riptortus pedestris</i> | Gut | GCAACCCTGGTGGCGAGTG<br>GCGAACGGGTGAGTAATAC<br>ATCGGAACGTGTCCTGTAG<br>TGGGGGATAGCCCGGCGA<br>AAGCCGGATTAATACCGCA<br>TACGATCTACGGAAGAAAG<br>CGGGGGATCCTTCGGGAC<br>CTCGCGCTATAGGGGCGG<br>CCGATGGCAGATTAGCTAG<br>TTGGTGGGGTAAAGGCCTA<br>CCAAGGCGACGATCTGTAG<br>CTGGTCTGAGAGGACGACC<br>AGCCACACTGGGACTGAGA<br>CACGGCCCAGACTCCTACG<br>GGAGGCAGCAGTGGGGAA<br>TTTTGGACAATGGGGGCAA<br>CCCTGATCCAGCAATGCCG<br>CGTGTGTGAAGAAGGCCTT<br>CGGGTTGTAAAGCACTTTT<br>GTCCGGAAAGAAAACGTCT<br>TGGTTAATACCTGAGGCGG<br>ATGACGGTACCGGAAGAAT<br>AAGCACCGGCTAACTACGT<br>GCCAGCAGCCGCGGTAATA<br>CGTAGGGTGCGAGCGTTAA<br>TCGGAATTACTGGGCGTAA<br>AGCGTGCGCAGGCGGTCT<br>GTTAAGACCGATGTGAAAT<br>CCCCGGGCTTAACCTGGGA<br>ACTGCATTGGTGACTGGCA<br>GGCTTTGAGTGTGGCAGAG<br>GGGGGTAGAATTCCACGTG |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|--|--|--|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | <div>TAGCAGTGAAATGCGTAGA<br/>GATGTGGAGGAATACCGAT<br/>GGCGAAGGCAGCCCCCTG<br/>GGCCAACACTGACGCTCAT<br/>GCACGAAAGCGTGGGGAG<br/>CAAACAGGATTAGATACCC<br/>TGGTAGTCCACGCCCTAAA<br/>CGATGTCAACTAGTTGTTG<br/>GGGATTCATTTCTTAGTAA<br/>CGTAGCTAACGCGTGAAGT<br/>TGACCGCCTGGGGAGTACG<br/>GTCGCAAGATTAAAACTCAA<br/>AGGAATTGACGGGGACCCG<br/>CACAAGCGGTGGATGATGT<br/>GGATTAATTCGATGCAACG<br/>CGAAAAACCTTACCTACCCT<br/>TGACATGGTCGGAAGTCTG<br/>CTGAGAGGTGGACGTGCTC<br/>GAAAGAGAACCGGCGCACA<br/>GGTGCTGCATGGCTGTCGT<br/>CAGCTCGTGTCGTGAGATG<br/>TTGGGTAAAGTCCCGCAAC<br/>GAGCGCAACCCTTGTCCTT<br/>AGTTGCTACGCAAGAGCAC<br/>TCTAAGGAGACTGCCGGTG<br/>ACAAACCGGAGGAAGGTGG<br/>GGATGACGTCAAGTCCTCA<br/>TGGCCCTTATGGGTAGGGC<br/>TTCACACGTCATACAATGGT<br/>CGGAACAGAGGGTTGCCAA<br/>GCCGCGAGGTGGAGCCAA<br/>TCCCAGAAAACCGATCGTA<br/>GTCCGGATCGCAGTCTGCA<br/>ACTCGACTGCGTGAAGCTG<br/>GAATCGCTAGTAATCGCGG<br/>ATCAGCATGCCGCGGTGAA<br/>TACGTTCCCGGGTCTTGTA<br/>CACACCGCCCGTCACACCA<br/>TGGGAGTGGGTTTCACCAG<br/>AAGTAGGTAGCCTAACCGC</div> |
|--|--|--|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                       |                     |     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-----------------------|---------------------|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                       |                     |     | AAGGAGGGCGCTTACCACG<br>GTGGGATTCATGACTGGGG<br>TGAAGT<br>(SEQ ID NO: 29)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Burkholderia sp. KM-G | Riptortus pedestris | Gut | GCAACCCTGGTGGCGAGTG<br>GCGAACGGGTGAGTAATAC<br>ATCGGAACGTGTCCTGTAG<br>TGGGGGATAGCCCGGCGA<br>AAGCCGGATTAATACCGCA<br>TACGACCTAAGGGAGAAAG<br>CGGGGGATCTTCGGACCTC<br>GCGCTATAGGGGCGGCCG<br>ATGGCAGATTAGCTAGTTG<br>GTGGGGTAAAGGCCTACCA<br>AGGCGACGATCTGTAGCTG<br>GTCTGAGAGGACGACCAGC<br>CACACTGGGACTGAGACAC<br>GGCCCAGACTCCTACGGGA<br>GGCAGCAGTGGGGAATTTT<br>GGACAATGGGGGCAACCCT<br>GATCCAGCAATGCCGCGTG<br>TGTGAAGAAGGCCTTCGGG<br>TTGTAAAGCACTTTTGTCCG<br>GAAAGAAAACCTTCGAGGTT<br>AATACCCTTGGAGGATGAC<br>GGTACCGGAAGAATAAGCA<br>CCGGCTAACTACGTGCCAG<br>CAGCCGCGGTAATACGTAG<br>GGTGCGAGCGTTAATCGGA<br>ATTACTGGGCGTAAAGCGT<br>GCGCAGGCGGTCTGTTAAG<br>ACCGATGTGAAATCCCCGG<br>GCTTAACCTGGGAACTGCA<br>TTGGTGA CTGGCAGGCTTT<br>GAGTGTGGCAGAGGGGGG<br>TAGAATTCCACGTGTAGCA<br>GTGAAATGCGTAGAGATGT<br>GGAGGAATACCGATGGCGA<br>AGGCAGCCCCCTGGGCCA<br>ACACTGACGCTCATGCACG |

|           |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-----------|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|           |  |  | AAAGCGTGGGGAGCAAACA<br>GGATTAGATACCCTGGTAG<br>TCCACGCCCTAAACGATGT<br>CAACTAGTTGTTGGGGATT<br>CATTTCCCTTAGTAACGTAGC<br>TAACGCGTGAAGTTGACCG<br>CCTGGGGAGTACGGTTCGCA<br>AGATTAAAACTCAAAGGAAT<br>TGACGGGGACCCGCACAA<br>GCGGTGGATGATGTGGATT<br>AATTCGATGCAACGCGAAA<br>AACCTTACCTACCCTTGACA<br>TGGTCGGAAGTCTGCTGAG<br>AGGTGGACGTGCTCGAAAG<br>AGAACCGGCGCACAGGTG<br>CTGCATGGCTGTCGTCAGC<br>TCGTGTCGTGAGATGTTGG<br>GTTAAGTCCCGCAACGAGC<br>GCAACCCTTGTCCTTAGTT<br>GCTACGCAAGAGCACTCTA<br>AGGAGACTGCCGGTGACAA<br>ACCGGAGGAAGGTGGGGA<br>TGACGTCAAGTCCTCATGG<br>CCCTTATGGGTAGGGCTTC<br>ACACGTCATACAATGGTTCG<br>GAACAGAGGGTTGCCAAGC<br>CGCGAGGTGGAGCCAATCC<br>CAGAAAACCGATCGTAGTC<br>CGGATCGCAGTCTGCAACT<br>CGACTGCGTGAAGCTGGAA<br>TCGCTAGTAATCGCGGATC<br>AGCATGCCGCGGTGAATAC<br>GTTCCCGGGTCTTGTACAC<br>ACCGCCCGTCCACCATGG<br>GAGTGGGTTTCACCAGAAG<br>TAGGTAGCCTAACCTGCAA<br>AGGAGGGCGCTTACCACG<br>(SEQ ID NO: 30) |
|           |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Nematodes |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

|                              |                             |                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|------------------------------|-----------------------------|-------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Xiphinematobacter sp.</i> | <i>Xiphinema americanum</i> | ovaries,<br>developing<br>eggs, and gut<br>lining                                                     | GCAAGTCGAACGGAGTGGA<br>ACCTGCAGTAATGCAGATT<br>CGATTCAGTGGCGTACGGG<br>TGCGTAACACGTGAGTGAT<br>CTACCGGTAAGTGGGGGAT<br>AACCCGCCGAAAGGCGAAT<br>TAATACCGCATGTGGCTAG<br>GGATGCCTTCATCCTGTAG<br>CTAAAGTCGATTTTGACGCT<br>TTCTGATGAGCTCGCGGCC<br>TATCAGCTTGTTGGTGGAG<br>GTAATGGCCCACCAAGGCA<br>ATGACGGGTAGCTGGTCTG<br>AGAGGACGATCAGCCACAC<br>TGGAAGTGAAGACACGGTCC<br>AGACACCTACGGGTGGCAG<br>CAGTCGAGAATTTTTCACAA<br>TGGGGGAAACCCTGATGAA<br>GCAACGCCGCGTGGAGGA<br>TGAAGGGCTTCGCGCTCGT<br>AACTCCTGTCAAGCGGGA<br>ACAAGAAAGTGATAGTACC<br>GCTAGAGGAAGAGACGGCT<br>AACTCTGTGCCAGCAGCCG<br>CGGTAATACAGAGGTCTCG<br>AGCGTTGTTCCGATTTATTG<br>GGCGTAAAGGGTGCGTAG<br>GCGGTGTGGCAAGTCAAGT<br>GTGAA<br>(SEQ ID NO: 31) |
|                              |                             |                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| <i>Radopholus similis</i>    |                             |                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| <i>Wolbachia sp. wOo</i>     | <i>Onchocerca ochengi</i>   | somatic<br>hypodermal<br>cords that run<br>along the<br>length of the<br>worms and in<br>the germinal | ATGACACACATACCGGTTTT<br>ACTAAAAGAAATGCTATCGC<br>AACTTTCACCACAAAATGGT<br>AGTGTATATGTGGATGCCA<br>CATTTGGAGCTGGAGGATA<br>TAGTAAAGCAATATTGGAGT<br>CAGCTGATTGCAGAGTGTA<br>TGCAATCGACAGAGATGAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

|  |  |                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|--|--|---------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  | zones of the female gonad | ACGGTTATTAAATTTTATAA<br>TAGTTTGAATACCAAGTACC<br>ACGGTAAAATAAAACTATTT<br>ATTGAAAAGTTTAGCAATAT<br>TCAAAC TATACTAAACAGTA<br>GTAATCTCAAACACTTTACA<br>GAACCTTCCGTCATTGTTTC<br>AGCTGGAATTCAGAAAAAA<br>AATGCAAGGTCAAGCACCG<br>AGATGATACAAAGTAATACC<br>GTAGATGGAGTTGTGTTTCG<br>ATATAGGAGTATCGTCTATG<br>CAGCTTGATGAAGAAAATA<br>GAGGATTTTCATTTTACAT<br>AACAGTCCGCTTGATATGC<br>GCATGGATACCTCTTCTCA<br>CATTAA CGCTTCAATATTTG<br>TTAATGCCTTACGCGAAGA<br>AGAAATTGCAAACACTATAT<br>ATAGCTATGGAGGTGAACG<br>TTATTCTCGCAA AATTGCAA<br>GAGCAATAGTGAACGTACG<br>TAAGAAAAAACTATCGACA<br>CTACATTTGAGCTTGCAGA<br>CATTGTACGTTCCGTGGTAT<br>CTCGCGGAAAAAGCAAGAT<br>TGATCCTGCAACTAGGACA<br>TTTCAAGCAATCAGAATATG<br>GGTAAACGATGAGCTTAGA<br>GAGCTTGAAAAGGGTATTA<br>AAGCTGCATCCAAAATCTTA<br>AATAGGAATGGCAAGCTGA<br>TTGTCATTACTTTTCATTCC<br>TTGGAAGATCGTATAGTCAA<br>GACCTTTTTTTAAAGGCTTAT<br>GTGAGCCAAAATTCACCAA<br>CTGTAGAACGTTTTCTCTTC<br>TGAATAAAAAAGTAATCAAG<br>GCAAGCGCAGAAGAAATAA<br>GTGCAAATCCACGTGCGCG |
|--|--|---------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                           |                                                                    |       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|---------------------------|--------------------------------------------------------------------|-------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                           |                                                                    |       | TTCAGCAAACTAAGAGCTA<br>TACAAAGGTTATTATGA<br>(SEQ ID NO: 32)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| <i>Snodgrassella alvi</i> | Honeybee<br>( <i>Apis mellifera</i> )<br>and <i>Bombus</i><br>spp. | Ileum | GAGAGTTTGATCCTGGCTC<br>AGATTGAACGCTGGCGGCA<br>TGCCTTACACATGCAAGTC<br>GAACGGCAGCACGGAGAG<br>CTTGCTCTCTGGTGGCGAG<br>TGGCGAACGGGTGAGTAAT<br>GCATCGGAACGTACCGAGT<br>AATGGGGGATAACTGTCCG<br>AAAGGATGGCTAATACCGC<br>ATACGCCCTGAGGGGGAAA<br>GCGGGGGATCGAAAGACCT<br>CGCGTTATTTGAGCGGCCG<br>ATGTTGGATTAGCTAGTTG<br>GTGGGGTAAAGGCCTACCA<br>AGGCGACGATCCATAGCGG<br>GTCTGAGAGGATGATCCGC<br>CACATTGGGACTGAGACAC<br>GGCCCAAACCTCCTACGGGA<br>GGCAGCAGTGGGGAATTTT<br>GGACAATGGGGGGAACCCCT<br>GATCCAGCCATGCCGCGTG<br>TCTGAAGAAGGCCTTCGGG<br>TTGTAAAGGACTTTTGTTAG<br>GGAAGAAAAGCCGGGTGTT<br>AATACCATCTGGTGCTGAC<br>GGTACCTAAAGAATAAGCA<br>CCGGCTAACTACGTGCCAG<br>CAGCCGCGGTAATACGTAG<br>GGTGCGAGCGTTAATCGGA<br>ATTACTGGGCGTAAAGCGA<br>GCGCAGACGGTTAATTAAG<br>TCAGATGTGAAATCCCCGA<br>GCTCAACTTGGGACGTGCA<br>TTTGAAACTGGTTAACTAGA<br>GTGTGTCAGAGGGAGGTAG<br>AATTCCACGTGTAGCAGTG<br>AAATGCGTAGAGATGTGGA |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|--|--|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | GGAATACCGATGGCGAAGG<br>CAGCCTCCTGGGATAACAC<br>TGACGTTTCATGCTCGAAAG<br>CGTGGGTAGCAAACAGGAT<br>TAGATACCCTGGTAGTCCA<br>CGCCCTAAACGATGACAAT<br>TAGCTGTTGGGACACTAGA<br>TGTCTTAGTAGCGAAGCTA<br>ACGCGTGAAATTGTCCGCC<br>TGGGGAGTACGGTCGCAAG<br>ATTAAAACTCAAAGGAATTG<br>ACGGGGACCCGCACAAGC<br>GGTGGATGATGTGGATTAA<br>TTCGATGCAACGCGAAGAA<br>CCTTACCTGGTCTTGACAT<br>GTACGGAATCTCTTAGAGA<br>TAGGAGAGTGCCTTCGGGA<br>ACCGTAACACAGGTGCTGC<br>ATGGCTGTCGTCAGCTCGT<br>GTCGTGAGATGTTGGGTTA<br>AGTCCCGCAACGAGCGCAA<br>CCCTTGTCATTAGTTGCCAT<br>CATTAAGTTGGGCACTCTAA<br>TGAGACTGCCGGTGACAAA<br>CCGGAGGAAGGTGGGGAT<br>GACGTCAAGTCCTCATGGC<br>CCTTATGACCAGGGCTTCA<br>CACGTCATACAATGGTCGG<br>TACAGAGGGTAGCGAAGCC<br>GCGAGGTGAAGCCAATCTC<br>AGAAAGCCGATCGTAGTCC<br>GGATTGCACTCTGCAACTC<br>GAGTGCATGAAGTCGGAAT<br>CGCTAGTAATCGCAGGTCA<br>GCATACTGCGGTGAATACG<br>TTCCCGGGTCTTGTACACA<br>CCGCCCCGTCACACCATGGG<br>AGTGGGGGATACCAGAATT<br>GGGTAGACTAACCGCAAGG<br>AGGTCGCTTAACACGGTAT |
|--|--|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                            |                                                                    |       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|----------------------------|--------------------------------------------------------------------|-------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                            |                                                                    |       | GCTTCATGACTGGGGTGAA<br>GTCGTAACAAGGTAGCCGT<br>AG<br>(SEQ ID NO: 33)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| <i>Gilliamella apicola</i> | honeybee<br>( <i>Apis mellifera</i> )<br>and <i>Bombus</i><br>spp. | Ileum | TTAAATTGAAGAGTTTGATC<br>ATGGCTCAGATTGAACGCT<br>GGCGGCAGGCTTAACACAT<br>GCAAGTCGAACGGTAACAT<br>GAGTGCTTGCACTTGATGA<br>CGAGTGGCGGACGGGTGA<br>GTAAAGTATGGGGATCTGC<br>CGAATGGAGGGGGACAACA<br>GTTGGAAACGACTGCTAAT<br>ACCGCATAAAGTTGAGAGA<br>CCAAAGCATGGGACCTTCG<br>GGCCATGCGCCATTTGATG<br>AACCCATATGGGATTAGCT<br>AGTTGGTAGGGTAATGGCT<br>TACCAAGGCGACGATCTCT<br>AGCTGGTCTGAGAGGATGA<br>CCAGCCACACTGGAAGTGA<br>GACACGGTCCAGACTCCTA<br>CGGGAGGCAGCAGTGGGG<br>AATATTGCACAATGGGGGA<br>AACCCTGATGCAGCCATGC<br>CGCGTGTATGAAGAAGGCC<br>TTCGGGTTGTAAAGTACTTT<br>CGGTGATGAGGAAGGTGGT<br>GTATCTAATAGGTGCATCAA<br>TTGACGTTAATTACAGAAGA<br>AGCACCGGCTAACTCCGTG<br>CCAGCAGCCGCGGTAATAC<br>GGAGGGTGCGAGCGTTAAT<br>CGGAATGACTGGGCGTAAA<br>GGGCATGTAGGCGGATAAT<br>TAAGTTAGGTGTGAAAGCC<br>CTGGGCTCAACCTAGGAAT<br>TGCACTTAAACTGGTTAAC<br>TAGAGTATTGTAGAGGAAG<br>GTAGAATTCCACGTGTAGC |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|--|--|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | GGTGAAATGCGTAGAGATG<br>TGGAGGAATACCGGTGGCG<br>AAGGCGGCCTTCTGGACAG<br>ATACTGACGCTGAGATGCG<br>AAAGCGTGGGGAGCAAACA<br>GGATTAGATACCCTGGTAG<br>TCCACGCTGTAAACGATGT<br>CGATTTGGAGTTTGTTCCT<br>AGAGTGATGGGCTCCGAAG<br>CTAACGCGATAAATCGACC<br>GCCTGGGGAGTACGGCCG<br>CAAGGTTAAAACTCAAATGA<br>ATTGACGGGGGGCCCGCACA<br>AGCGGTGGAGCATGTGGTT<br>TAATTCGATGCAACGCGAA<br>GAACCTTACCTGGTCTTGA<br>CATCCACAGAATCTTGCAG<br>AGATGCGGGAGTGCCTTCG<br>GGAAGTGTGAGACAGGTGC<br>TGCATGGCTGTCGTCAGCT<br>CGTGTTGTGAAATGTTGGG<br>TTAAGTCCCGCAACGAGCG<br>CAACCCTTATCCTTTGTTGC<br>CATCGGTTAGGCCGGGAAC<br>TCAAAGGAGACTGCCGTTG<br>ATAAAGCGGAGGAAGGTGG<br>GGACGACGTCAAGTCATCA<br>TGGCCCTTACGACCAGGGC<br>TACACACGTGCTACAATGG<br>CGTATACAAAGGGAGGCGA<br>CCTCGCGAGAGCAAGCGG<br>ACCTCATAAAGTACGTCTAA<br>GTCCGGATTGGAGTCTGCA<br>ACTCGACTCCATGAAGTCG<br>GAATCGCTAGTAATCGTGA<br>ATCAGAATGTCACGGTGAA<br>TACGTTCCCGGGCCTTGTA<br>CACACCGCCCGTCACACCA<br>TGGGAGTGGGTTGCACCAG<br>AAGTAGATAGCTTAACCTTC |
|--|--|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                        |                                       |     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|------------------------|---------------------------------------|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                        |                                       |     | GGGAGGGCGTTTACCACG<br>GTGTGGTCCATGACTGGGG<br>TGAAGTCGTAACAAGGTAA<br>CCGTAGGGGAACCTGCGGT<br>TGGATCACCTCCTTAC<br>(SEQ ID NO: 34)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| <i>Bartonella apis</i> | honeybee<br>( <i>Apis mellifera</i> ) | Gut | AAGCCAAAATCAAATTTTCA<br>ACTTGAGAGTTTGATCCTG<br>GCTCAGAACGAACGCTGGC<br>GGCAGGCTTAACACATGCA<br>AGTCGAACGCACTTTTCGG<br>AGTGAGTGGCAGACGGGT<br>GAGTAACGCGTGGGAATCT<br>ACCTATTTCTACGGAATAAC<br>GCAGAGAAATTTGTGCTAAT<br>ACCGTATACGTCCTTCGGG<br>AGAAAGATTTATCGGAGATA<br>GATGAGCCCGCGTTGGATT<br>AGCTAGTTGGTGAGGTAAT<br>GGCCCACCAAGGCGACGAT<br>CCATAGCTGGTCTGAGAGG<br>ATGACCAGCCACATTGGGA<br>CTGAGACACGGCCCAGACT<br>CCTACGGGAGGCAGCAGT<br>GGGGAATATTGGACAATGG<br>GCGCAAGCCTGATCCAGCC<br>ATGCCGCGTGAGTGATGAA<br>GGCCCTAGGGTTGTAAAGC<br>TCTTTCACCGGTGAAGATAA<br>TGACGGTAACCGGAGAAGA<br>AGCCCCGGCTAACTTCGTG<br>CCAGCAGCCGCGGTAATAC<br>GAAGGGGGCTAGCGTTGTT<br>CGGATTTACTGGGCGTAAA<br>GCGCACGTAGGCGGATATT<br>TAAGTCAGGGGTGAAATCC<br>CGGGGCTCAACCCCGGAA<br>CTGCCTTTGATACTGGATAT<br>CTTGAGTATGGAAGAGGTA<br>AGTGGAATTCCGAGTGTAG |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | AGGTGAAATTCGTAGATATT<br>CGGAGGAACACCAGTGGC<br>GAAGGCGGCTTACTGGTCC<br>ATTACTGACGCTGAGGTGC<br>GAAAGCGTGGGGAGCAAAC<br>AGGATTAGATACCCTGGTA<br>GTCCACGCTGTAAACGATG<br>AATGTTAGCCGTTGGACAG<br>TTTACTGTTCCGGTGGCGCA<br>GCTAACGCATTAAACATTCC<br>GCCTGGGGAGTACGGTCG<br>CAAGATTAAAACTCAAAGGA<br>ATTGACGGGGGGCCCGCACA<br>AGCGGTGGAGCATGTGGTT<br>TAATTCGAAGCAACGCGCA<br>GAACCTTACCAGCCCTTGA<br>CATCCCGATCGCGGATGGT<br>GGAGACACCGTCTTTCAGT<br>TCGGCTGGATCGGTGACAG<br>GTGCTGCATGGCTGTCGTC<br>AGCTCGTGTCTGTGAGATGT<br>TGGGTTAAGTCCCGCAACG<br>AGCGCAACCCTCGCCCTTA<br>GTTGCCATCATTTAGTTGG<br>GCACTCTAAGGGGACTGCC<br>GGTGATAAGCCGAGAGGAA<br>GGTGGGGATGACGTCAAGT<br>CCTCATGGCCCTTACGGGC<br>TGGGCTACACACGTGCTAC<br>AATGGTGGTGACAGTGGGC<br>AGCGAGACCGCGAGGTCCG<br>AGCTAATCTCCAAAAGCCAT<br>CTCAGTTCGGATTGCACTC<br>TGCAACTCGAGTGCATGAA<br>GTTGGAATCGCTAGTAATC<br>GTGGATCAGCATGCCACGG<br>TGAATACGTTCCCGGGCCT<br>TGTAACACCGCCCGTCAC<br>ACCATGGGAGTTGGTTTTA<br>CCCGAAGGTGCTGTGCTAA |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                 |                                       |     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|---------------------------------|---------------------------------------|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                 |                                       |     | CCGCAAGGAGGCAGGCAA<br>CCACGGTAGGGTCAGCGAC<br>TGGGGTGAAGTCGTAACAA<br>GGTAGCCGTAGGGGAACCT<br>GCGGCTGGATCACCTCCTT<br>TCTAAGGAAGATGAAGAATT<br>GGAA<br>(SEQ ID NO: 35)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <i>Parasaccharibacter apium</i> | honeybee<br>( <i>Apis mellifera</i> ) | Gut | CTACCATGCAAGTCGCACG<br>AAACCTTTCGGGGTTAGTG<br>GCGGACGGGTGAGTAACG<br>CGTTAGGAACCTATCTGGA<br>GGTGGGGGATAACATCGG<br>GAAACTGGTGCTAATACCG<br>CATGATGCCTGAGGGCCAA<br>AGGAGAGATCCGCCATTGG<br>AGGGGCCTGCGTTCGATTA<br>GCTAGTTGGTTGGGTAAAG<br>GCTGACCAAGGCGATGATC<br>GATAGCTGGTTTGAGAGGA<br>TGATCAGCCACACTGGGAC<br>TGAGACACGGCCCAGACTC<br>CTACGGGAGGCAGCAGTG<br>GGGAATATTGGACAATGGG<br>GGCAACCCTGATCCAGCAA<br>TGCCGCGTGTGTGAAGAAG<br>GTCTTCGGATTGTAAAGCA<br>CTTTCACTAGGGAAGATGA<br>TGACGGTACCTAGAGAAGA<br>AGCCCCGGCTAACTTCGTG<br>CCAGCAGCCGCGGTAATAC<br>GAAGGGGGCTAGCGTTGCT<br>CGGAATGACTGGGCGTAAA<br>GGGCGCGTAGGCTGTTTGT<br>ACAGTCAGATGTGAAATCC<br>CCGGGCTTAACCTGGGAAC<br>TGCATTTGATACGTGCAGA<br>CTAGAGTCCGAGAGAGGGT<br>TGTGGAATTCCCAGTGTAG<br>AGGTGAAATTCGTAGATATT |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|--|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | <div>GGGAAGAACACCGGTTGCG<br/>AAGGCGGCAACCTGGCTNN<br/>NNNNNNNNNNNNNNNNNNNN<br/>NNNNNNNNNNNNNNNNNNNN<br/>NNNNNNNNNNNNNNNNNNNN<br/>NNNNNNNNNNNNNNNNNNNN<br/>NNNNNNNNNNNNNNNNNNNN<br/>NNNNNNNNNNNNNNNNNNNN<br/>NNNNNNNNNNNNNNNNNNNN<br/>NNNGAGCTAACGCGTTAA<br/>GCACACCGCCTGGGGAGTA<br/>CGGCCGCAAGGTTGAACT<br/>CAAAGGAATTGACGGGGGC<br/>CCGCACAAGCGGTGGAGC<br/>ATGTGGTTTAATTCGAAGCA<br/>ACGCGCAGAACCTTACCAG<br/>GGCTTGCATGGGGAGGCT<br/>GTATTCAGAGATGGATATTT<br/>CTTCGGACCTCCCGCACAG<br/>GTGCTGCATGGCTGTCGTC<br/>AGCTCGTGTCGTGAGATGT<br/>TGGGTTAAGTCCCGCAACG<br/>AGCGCAACCCTTGTCTTTA<br/>GTTGCCATCACGTCTGGGT<br/>GGGCACTCTAGAGAGACTG<br/>CCGGTGACAAGCCGGAGG<br/>AAGGTGGGGATGACGTCAA<br/>GTCCTCATGGCCCTTATGT<br/>CCTGGGCTACACACGTGCT<br/>ACAATGGCGGTGACAGAGG<br/>GATGCTACATGGTGACATG<br/>GTGCTGATCTCAAAAACC<br/>GTCTCAGTTCGGATTGTACT<br/>CTGCAACTCGAGTGCATGA<br/>AGGTGGAATCGCTAGTAAT<br/>CGCGGATCAGCATGCCGC<br/>GGTGAATACGTTCCCGGGC<br/>CTTGTAACACACCGCCCGTC<br/>ACACCATGGGAGTTGGTTT<br/>GACCTTAAGCCGGTGAGCG<br/>AACCGCAAGGAACGCAGCC</div> |
|--|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                              |                                       |     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------------------|---------------------------------------|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                              |                                       |     | GACCACCGGTTCTGGGTTCA<br>GCGACTGGGGGA<br>(SEQ ID NO: 36)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| <i>Lactobacillus kunkeei</i> | honeybee<br>( <i>Apis mellifera</i> ) | Gut | TTCCTTAGAAAGGAGGTGA<br>TCCAGCCGCAGGTTCTCCT<br>ACGGCTACCTTGTTACGAC<br>TTCACCCTAATCATCTGTCC<br>CACCTTAGACGACTAGCTC<br>CTAAAAGGTTACCCCATCG<br>TCTTTGGGTGTTACAACTC<br>TCATGGTGTGACGGGCGGT<br>GTGTACAAGGCCCGGGAAC<br>GTATTCACCGTGGCATGCT<br>GATCCACGATTACTAGTGAT<br>TCCAACCTTCATGCAGGCGA<br>GTTGCAGCCTGCAATCCGA<br>ACTGAGAATGGCTTTAAGA<br>GATTAGCTTGACCTCGCGG<br>TTTCGCGACTCGTTGTACC<br>ATCCATTGTAGCACGTGTG<br>TAGCCCAGCTCATAAGGGG<br>CATGATGATTTGACGTCGT<br>CCCCACCTTCCTCCGGTTT<br>ATCACCGGCAGTCTCACTA<br>GAGTGCCCAACTAAATGCT<br>GGCAACTAATAATAAGGGT<br>TGCGCTCGTTGCGGGACTT<br>AACCCAACATCTCACGACA<br>CGAGCTGACGACAACCATG<br>CACCACCTGTCATTCTGTC<br>CCCGAAGGGAACGCCCAAT<br>CTCTTGGGTTGGCAGAAGA<br>TGTCAGAGCTGGTAAGGT<br>TCTTCGCGTAGCATCGAATT<br>AAACCACATGCTCCACCAC<br>TTGTGCGGGCCCCCGTCAA<br>TTCCTTTGAGTTTCAACCTT<br>GCGGTCTGACTCCCCAGGC<br>GGAATACTTAATGCGTTAG<br>CTGCGGCACTGAAGGGCG |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|--|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | GAAACCCTCCAACACCTAG<br>TATTCATCGTTTACGGCATG<br>GACTACCAGGGTATCTAAT<br>CCTGTTCTGCTACCCATGCT<br>TTCGAGCCTCAGCGTCAGT<br>AACAGACCAGAAAGCCGCC<br>TTCGCCACTGGTGTTCCTC<br>CATATATCTACGCATTTTAC<br>CGCTACACATGGAGTTCCA<br>CTTTCCTCTTCTGTACTION<br>GTTTTGTAGTTTCCACTGCA<br>CTTCCTCAGTTGAGCTGAG<br>GGCTTTCACAGCAGACTTA<br>CAAAACCGCCTGCGCTCGC<br>TTACGCCCAATAAATCCG<br>GACAACGCTTGCCACCTAC<br>GTATTACCGCGGCTGCTGG<br>CACGTAGTTAGCCGTGGCT<br>TTCTGGTTAAATACCGTCAA<br>AGTGTTAACAGTTACTCTAA<br>CACTTGTTCTTCTTTAACAA<br>CAGAGTTTTACGATCCGAA<br>AACCTTCATCACTCACGCG<br>GCGTTGCTCCATCAGACTT<br>TCGTCCATTGTGGAAGATT<br>CCCTACTGCTGCCTCCCGT<br>AGGAGTCTGGGCCGTGTCT<br>CAGTCCCAATGTGGCCGAT<br>TACCCTCTCAGGTCGGCTA<br>CGTATCATCGTCTTGGTGG<br>GCTTTTATCTCACCAACTAA<br>CTAATACGGCGCGGGTCCA<br>TCCCAAAGTGATAGCAAAG<br>CCATCTTTCAAGTTGGAACC<br>ATGCGGTTCCAACTAATTAT<br>GCGGTATTAGCACTTGTTTC<br>CAAATGTTATCCCCGCTTC<br>GGGGCAGGTTACCCACGTG<br>TACTCACCAAGTTCGCCACT<br>CGCTCCGAATCCAAAAATC |
|--|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                             |                                       |     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-----------------------------|---------------------------------------|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                             |                                       |     | ATTTATGCAAGCATAAAATC<br>AATTTGGGAGAACTCGTTC<br>GACTTGCATGTATTAGGCA<br>CGCCGCCAGCGTTCGTCCT<br>GAGCCAGGATCAAACCTCTC<br>ATCTTAA<br>(SEQ ID NO: 37)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| <i>Lactobacillus Firm-4</i> | honeybee<br>( <i>Apis mellifera</i> ) | Gut | ACGAACGCTGGCGGGCGTG<br>CCTAATACATGCAAGTCGA<br>GCGCGGGAAGTCAGGGAA<br>GCCTTCGGGTGGAACCTGGT<br>GGAACGAGCGGCGGATGG<br>GTGAGTAACACGTAGGTAA<br>CCTGCCCTAAAGCGGGGGA<br>TACCATCTGGAAACAGGTG<br>CTAATACCGCATAAACCCA<br>GCAGTCACATGAGTGCTGG<br>TTGAAAGACGGCTTCGGCT<br>GTCACTTTAGGATGGACCT<br>GCGGCGTATTAGCTAGTTG<br>GTGGAGTAACGGTTCACCA<br>AGGCAATGATACGTAGCCG<br>ACCTGAGAGGGTAATCGGC<br>CACATTGGGACTGAGACAC<br>GGCCCAAACCTCCTACGGGA<br>GGCAGCAGTAGGGAATCTT<br>CCACAATGGACGCAAGTCT<br>GATGGAGCAACGCCGCGT<br>GGATGAAGAAGGTCTTCGG<br>ATCGTAAAATCCTGTTGTTG<br>AAGAAGAACGGTTGTGAGA<br>GTAAC TGCTCATAACGTGA<br>CGGTAATCAACCAGAAAGT<br>CACGGCTAACTACGTGCCA<br>GCAGCCGCGGTAATACGTA<br>GGTGGCAAGCGTTGTCCGG<br>ATTTATTGGGCGTAAAGGG<br>AGCGCAGGCGGTCTTTTAA<br>GTCTGAATGTGAAAGCCCT<br>CAGCTTAACTGAGGAAGAG |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|--|--|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | CATCGGAAACTGAGAGACT<br>TGAGTGCAGAAGAGGAGAG<br>TGGAACTCCATGTGTAGCG<br>GTGAAATGCGTAGATATAT<br>GGAAGAACACCAGTGGCGA<br>AGGCGGCTCTCTGGTCTGT<br>TACTGACGCTGAGGCTCGA<br>AAGCATGGGTAGCGAACAG<br>GATTAGATACCCTGGTAGT<br>CCATGCCGTAAACGATGAG<br>TGCTAAGTGTTGGGAGGTT<br>TCCGCCTCTCAGTGCTGCA<br>GCTAACGCATTAAGCACTC<br>CGCCTGGGGAGTACGACC<br>GCAAGGTTGAAACTCAAAG<br>GAATTGACGGGGGCCCCGC<br>ACAAGCGGTGGAGCATGTG<br>GTTTAATTCTGAAGCAACGC<br>GAAGAACCTTACCAGGTCT<br>TGACATCTCCTGCAAGCCT<br>AAGAGATTAGGGGTTCCCT<br>TCGGGGACAGGAAGACAG<br>GTGGTGCATGGTTGTCGTC<br>AGCTCGTGTCTGTGAGATGT<br>TGGGTTAAGTCCCGCAACG<br>AGCGCAACCCTTGTTACTA<br>GTTGCCAGCATTAAAGTTGG<br>GCACTCTAGTGAGACTGCC<br>GGTGACAAACCGGAGGAAG<br>GTGGGGACGACGTCAAATC<br>ATCATGCCCCTTATGACCT<br>GGGCTACACACGTGCTACA<br>ATGGATGGTACAATGAGAA<br>GCGAACTCGCGAGGGGAA<br>GCTGATCTCTGAAAACCATT<br>CTCAGTTCGGATTGCAGGC<br>TGCAACTCGCCTGCATGAA<br>GCTGGAATCGCTAGTAATC<br>GCGGATCAGCATGCCGCG |
|--|--|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|              |                    |     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|--------------|--------------------|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|              |                    |     | GTGAATACGTTCCCGGGCC<br>TTGTACACACCGCCC<br>(SEQ ID NO: 38)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|              |                    |     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Silk worm    |                    |     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Enterococcus | <i>Bombyx mori</i> | Gut | AGGTGATCCAGCCGCACCT<br>TCCGATACGGCTACCTTGT<br>TACGACTTCACCCCAATCAT<br>CTATCCCACCTTAGGCGGC<br>TGGCTCCAAAAAGGTTACC<br>TCACCGACTTCGGGTGTTA<br>CAAACCTCTCGTGGTGTGAC<br>GGGCGGTGTGTACAAGGC<br>CCGGGAACGTATTCACCGC<br>GGCGTGCTGATCCGCGATT<br>ACTAGCGATTCCGGCTTCA<br>TGCAGGCGAGTTGCAGCCT<br>GCAATCCGAACTGAGAGAA<br>GCTTTAAGAGATTTGCATGA<br>CCTCGCGGTCTAGCGACTC<br>GTTGTACTTCCCATTGTAGC<br>ACGTGTGTAGCCCAGGTCA<br>TAAGGGGCATGATGATTTG<br>ACGTCATCCCCACCTTCCT<br>CCGGTTTGTACCGGCAGT<br>CTCGCTAGAGTGCCCAACT<br>AAATGATGGCAACTAACAAT<br>AAGGGTTGCGCTCGTTGCG<br>GGACTTAACCCAACATCTC<br>ACGACACGAGCTGACGACA<br>ACCATGCACCACCTGTCAC<br>TTTGTCCCCGAAGGGAAAG<br>CTCTATCTCTAGAGTGGTCA<br>AAGGATGTCAAGACCTGGT<br>AAGGTTCTTCGCGTTGCTT<br>CGAATTAAACCACATGCTC<br>CACCGCTTGTGCGGGCCCC<br>CGTCAATTCCTTTGAGTTTC<br>AACCTTGCGGTCGTACTION<br>CCAGGCGGAGTGCTTAATG |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|--|--|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | CGTTTGCTGCAGCACTGAA<br>GGGCGGAAACCCTCCAACA<br>CTTAGCACTCATCGTTTACG<br>GCGTGGACTACCAGGGTAT<br>CTAATCCTGTTTGCTCCCCA<br>CGCTTTCGAGCCTCAGCGT<br>CAGTTACAGACCAGAGAGC<br>CGCCTTCGCCACTGGTGTT<br>CCTCCATATATCTACGCATT<br>TCACCGCTACACATGGAAT<br>TCCACTCTCCTCTTCTGCAC<br>TCAAGTCTCCCAGTTTCCAA<br>TGACCCTCCCCGGTTGAGC<br>CGGGGGCTTTCACATCAGA<br>CTTAAGAAACCGCCTGCGC<br>TCGCTTTACGCCCAATAAAT<br>CCGGACAACGCTTGCCACC<br>TACGTATTACCGCGGCTGC<br>TGGCACGTAGTTAGCCGTG<br>GCTTTCTGGTTAGATACCGT<br>CAGGGGACGTTCAGTTACT<br>AACGTCCTTGTTCTTCTCTA<br>ACAACAGAGTTTTACGATCC<br>GAAAACCTTCTTCACTCACG<br>CGGCGTTGCTCGGTCAGAC<br>TTTCGTCCATTGCCGAAGAT<br>TCCCTACTGCTGCCTCCCG<br>TAGGAGTCTGGGCCGTGTC<br>TCAGTCCCAGTGTGGCCGA<br>TCACCCTCTCAGGTCGGCT<br>ATGCATCGTGGCCTTGGTG<br>AGCCGTTACCTCACCAACT<br>AGCTAATGCACCGCGGGTC<br>CATCCATCAGCGACACCCG<br>AAAGCGCCTTTCACTCTTAT<br>GCCATGCGGCATAAACTGT<br>TATGCGGTATTAGCACCTG<br>TTTCCAAGTGTTATCCCCCT<br>CTGATGGGTAGGTTACCCA<br>CGTGTTACTCACCCGTCCG |
|--|--|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                |                    |     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|----------------|--------------------|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                |                    |     | CCACTCCTCTTTCCAATTGA<br>GTGCAAGCACTCGGGAGGA<br>AAGAAGCGTTCTGACTTGCA<br>TGTATTAGGCACGCCGCCA<br>GCGTTCGTCCTGAGCCAGG<br>ATCAAACCTCT<br>(SEQ ID NO: 39)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <i>Delftia</i> | <i>Bombyx mori</i> | Gut | CAGAAAGGAGGTGATCCAG<br>CCGCACCTTCCGATACGGC<br>TACCTTGTTACGACTTCACC<br>CCAGTCACGAACCCCGCCG<br>TGGTAAGCGCCCTCCTTGC<br>GGTTAGGCTACCTACTTCT<br>GGCGAGACCCGCTCCCATG<br>GTGTGACGGGCGGTGTGTA<br>CAAGACCCGGGAACGTATT<br>CACCGCGGCATGCTGATCC<br>GCGATTACTAGCGATTCCG<br>ACTTCACGCAGTCGAGTTG<br>CAGACTGCGATCCGGACTA<br>CGACTGGTTTTATGGGATTA<br>GCTCCCCCTCGCGGGTTGG<br>CAACCCTCTGTACCAGCCA<br>TTGTATGACGTGTGTAGCC<br>CCACCTATAAGGGCCATGA<br>GGACTTGACGTCATCCCCA<br>CCTTCCTCCGGTTTGTAC<br>CGGCAGTCTCATTAGAGTG<br>CTCAACTGAATGTAGCAACT<br>AATGACAAGGGTTGCGCTC<br>GTTGCGGGACTTAACCCAA<br>CATCTCACGACACGAGCTG<br>ACGACAGCCATGCAGCACC<br>TGTGTGCAGGTTCTCTTTC<br>GAGCACGAATCCATCTCTG<br>GAAACTTCCTGCCATGTCA<br>AAGGTGGGTAAGGTTTTTC<br>GCGTTGCATCGAATTAAAC<br>CACATCATCCACCGCTTGT<br>GCGGGTCCCCGTCAATTCC |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|--|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | TTTGAGTTTCAACCTTGCGG<br>CCGTACTCCCCAGGCGGTC<br>AACTTCACGCGTTAGCTTC<br>GTTACTGAGAAAATAATTC<br>CCAACAACCAGTTGACATC<br>GTTTAGGGCGTGGACTACC<br>AGGGTATCTAATCCTGTTTG<br>CTCCCCACGCTTTCGTGCA<br>TGAGCGTCAGTACAGGTCC<br>AGGGGATTGCCTTCGCCAT<br>CGGTGTTCTCCGCATATC<br>TACGCATTTCACTGCTACAC<br>GCGGAATTCCATCCCCCTC<br>TACCGTACTCTAGCCATGC<br>AGTCACAAATGCAGTTCCC<br>AGGTTGAGCCCGGGGATTT<br>CACATCTGTCTTACATAACC<br>GCCTGCGCACGCTTTACGC<br>CCAGTAATTCCGATTAACG<br>CTCGCACCTACGTATTAC<br>CGCGGCTGCTGGCACGTA<br>GTTAGCCGGTGCTTATTCTT<br>ACGGTACCGTCATGGGCCC<br>CCTGTATTAGAAGGAGCTTT<br>TTCGTTCCGTACAAAAGCA<br>GTTTACAACCCGAAGGCCT<br>TCATCCTGCACGCGGCATT<br>GCTGGATCAGGCTTTCGCC<br>CATTGTCCAAAATTCCCCAC<br>TGCTGCCTCCCGTAGGAGT<br>CTGGGCCGTGTCTCAGTCC<br>CAGTGTGGCTGGTCGTCCT<br>CTCAGACCAGCTACAGATC<br>GTCGGCTTGGTAAGCTTTT<br>ATCCCACCAACTACCTAATC<br>TGCCATCGGCCGCTCCAAT<br>CGCGCGAGGCCCGAAGGG<br>CCCCCGCTTTCATCCTCAG<br>ATCGTATGCGGTATTAGCTA<br>CTCTTTCGAGTAGTTATCCC |
|--|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                  |                    |     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------|--------------------|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |                    |     | CCACGACTGGGACAGTTCC<br>GATGTATTACTACCCGTTCC<br>GCCACTCGTCAGCGTCCGA<br>AGACCTGTTACCGTTCCGAC<br>TTGCATGTGTAAGGCATGC<br>CGCCAGCGTTCAATCTGAG<br>CCAGGATCAAACCTCTACAG<br>TTCGATCT<br>(SEQ ID NO: 40)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <i>Pelomonas</i> | <i>Bombyx mori</i> | Gut | ATCCTGGCTCAGATTGAAC<br>GCTGGCGGCATGCCTTACA<br>CATGCAAGTCGAACGGTAA<br>CAGGTTAAGCTGACGAGTG<br>GCGAACGGGTGAGTAATAT<br>ATCGGAACGTGCCCAGTCG<br>TGGGGGATAACTGCTCGAA<br>AGAGCAGCTAATACCGCAT<br>ACGACCTGAGGGTGAAAGC<br>GGGGGATCGCAAGACCTC<br>GCNNGATTGGAGCGGCCG<br>ATATCAGATTAGGTAGTTGG<br>TGGGGTAAAGGCCCAACAA<br>GCCAACGATCTGTAGCTGG<br>TCTGAGAGGACGACCAGCC<br>AACTGGGACTGAGACACG<br>GCCCAGACTCCTACGGGAG<br>GCAGCAGTGGGGAATTTTG<br>GACAATGGGCGCAAGCCTG<br>ATCCAGCCATGCCGCGTGC<br>GGGAAGAAGGCCTTCGGGT<br>TGTAACCGCTTTTGTGAG<br>GGAAGAAAAGGTTCTGGTT<br>AATACCTGGGACTCATGAC<br>GGTACCTGAAGAATAAGCA<br>CCGGCTAACTACGTGCCAG<br>CAGCCGCGGTAATACGTAG<br>GGTGCAAGCGTTAATCGGA<br>ATTACTGGGCGTAAAGCGT<br>GCGCAGGCGGTTATGCAAG<br>ACAGAGGTGAAATCCCCGG |

|  |  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|--|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | GCTCAACCTGGGAACTGCC<br>TTTGTGACTGCATAGCTAGA<br>GTACGGTAGAGGGGGATG<br>GAATTCCGCGTGTAGCAGT<br>GAAATGCGTAGATATGCGG<br>AGGAACACCGATGGCGAAG<br>GCAATCCCCTGGACCTGTA<br>CTGACGCTCATGCACGAAA<br>GCGTGGGGAGCAAACAGG<br>ATTAGATACCCTGGTAGTC<br>CACGCCCTAAACGATGTCA<br>ACTGGTTGTTGGGAGGGTT<br>TCTTCTCAGTAACGTANNTA<br>ACGCGTGAAGTTGACCGCC<br>TGGGGAGTACGGCCGCAA<br>GGTTGAAACTCAAAGGAAT<br>TGACGGGGACCCGCACAA<br>GCGGTGGATGATGTGGTTT<br>AATTCGATGCAACGCGAAA<br>AACCTTACCTACCCTTGACA<br>TGCCAGGAATCCTGAAGAG<br>ATTTGGGAGTGCTCGAAAG<br>AGAACCTGGACACAGGTGC<br>TGCATGGCCGTCGTCAGCT<br>CGTGTCGTGAGATGTTGGG<br>TTAAGTCCCGCAACGAGCG<br>CAACCCTTGTCATTAGTTGC<br>TACGAAAGGGCACTCTAAT<br>GAGACTGCCGGTGACAAAC<br>CGGAGGAAGGTGGGGATG<br>ACGTCAGGTCATCATGGCC<br>CTTATGGGTAGGGCTACAC<br>ACGTCATACAATGGCCGGG<br>ACAGAGGGCTGCCAACCCG<br>CGAGGGGGAGCTAATCCCA<br>GAAACCCGGTCGTAGTCCG<br>GATCGTAGTCTGCAACTCG<br>ACTGCGTGAAGTCGGAATC<br>GCTAGTAATCGCGGATCAG<br>CTTGCCGCGGTGAATACGT |
|--|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|  |  |  |                                                                                                                                                                                                                |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  |  | TCCCGGGTCTTGTACACAC<br>CGCCCGTCACACCATGGGA<br>GCGGGTTCTGCCAGAAGTA<br>GTTAGCCTAACCGCAAGGA<br>GGGCGATTACCACGGCAG<br>GGTTCGTGACTGGGGTGAA<br>GTCGTAACAAGGTAGCCGT<br>ATCGGAAGGTGCGGCTGGA<br>TCAC<br>(SEQ ID NO: 41) |
|--|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Any number of bacterial species may be targeted by the compositions or methods described herein. For example, in some instances, the modulating agent may target a single bacterial species. In some instances, the modulating agent may target at least about any of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 500, or more distinct bacterial species. In some instances, the modulating agent may target any one of about 1 to about 5, about 5 to about 10, about 10 to about 20, about 20 to about 50, about 50 to about 100, about 100 to about 200, about 200 to about 500, about 10 to about 50, about 5 to about 20, or about 10 to about 100 distinct bacterial species. In some instances, the modulating agent may target at least about any of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, or more phyla, classes, orders, families, or genera of bacteria.

In some instances, the modulating agent may increase a population of one or more bacteria by at least about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more in the host. In some instances, the modulating agent may reduce the population of one or more bacteria by at least about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in the host. In some instances, the modulating agent may eradicate the population of a bacterium in the host.

In some instances, the modulating agent may alter the bacterial diversity and/or bacterial composition of the host in comparison to a host organism to which the modulating agent has not been administered. In some instances, the modulating agent may increase the bacterial diversity in the host relative to a starting diversity by at least about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in comparison to a host organism to which the modulating agent has not been administered. In some instances, the modulating agent may decrease the bacterial diversity in the host relative to a starting diversity by at least about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in comparison to a host organism to which the modulating agent has not been administered.

In some instances, the modulating agent may alter the function, activity, growth, and/or division of one or more bacterial cells in comparison to a host organism to which the modulating agent has not been administered. For example, the modulating agent may alter the expression of one or genes in the bacteria. In some instances, the modulating agent may alter the function of one or more proteins in the bacteria. In some instances, the modulating agent may alter the function of one or more cellular structures (e.g., the cell wall, the outer or inner membrane) in the bacteria. In some instances, the modulating agent may kill (e.g., lyse) the bacteria.

The target bacterium may reside in one or more parts of the invertebrate host (e.g., insect, mollusk, or nematode). Further, the target bacteria may be intracellular or extracellular. In some instances, the bacteria reside in or on one or more parts of the host gut, including, e.g., the foregut, midgut, and/or hindgut. In some instances, the bacteria reside as an intracellular bacteria within a cell of the host. In some instances, the bacteria reside in a bacteriocyte of the host invertebrate (e.g., insect, mollusk, or nematode).

Changes to the populations of bacteria in the host invertebrate (e.g., insect, mollusk, or nematode) may be determined by any methods known in the art, such as standard culturing techniques, CFU counts, microarray, polymerase chain reaction (PCR), real-time PCR, flow cytometry, fluorescence microscopy, transmission electron microscopy, fluorescence *in situ* hybridization (e.g., FISH), spectrophotometry, matrix-assisted laser desorption ionization-mass spectrometry (MALDI-MS), or DNA sequencing. In some instances, a sample of the host treated with a modulating agent is sequenced (e.g., by metagenomics sequencing of 16S rRNA or rDNA) to determine the microbiome of the host after delivery or administration of the modulating agent. In some instances, a sample of a host that did not receive the modulating agent is also sequenced to provide a reference.

ii. Fungi and Yeast

Exemplary fungi that may be targeted in accordance with the methods and compositions provided herein, include, but are not limited to *Amylostereum areolatum*, *Epichloe spp*, *Pichia pinus*, *Hansenula capsulate*, *Daldinia decipien*, *Ceratocytis spp*, *Ophiostoma spp*, and *Attamyces bromatificus*. Non-limiting examples of yeast and yeast-like symbionts found in invertebrates (e.g., insect, mollusk, or nematode) include *Candida*, *Metschnikowia*, *Leucocoprinu* (e.g., *Leucocoprinus gongylophorus*), *Debaromyces*, *Scheffersomyces shehatae* and *Scheffersomyces stipites*, *Starmerella*, *Pichia*, *Trichosporon*, *Cryptococcus*, *Pseudozyma*, and yeast-like symbionts from the subphylum *Pezizomycotina* (e.g., *Symbiotaphrina buchneri* and *Symbiotaphrina kochii*). Non-limiting examples of yeast that may be targeted by the methods and compositions herein are listed in Table 2.

Table 2: Examples of Yeast in Insects

| Insect Species                | Order: Family         | Yeast Location (Species)                                     |
|-------------------------------|-----------------------|--------------------------------------------------------------|
| <i>Stegobium paniceum</i>     | Coleoptera: Anobiidae | Mycetomes                                                    |
| (= <i>Sitodrepa panicea</i> ) |                       | ( <i>Saccharomyces</i> )                                     |
|                               |                       | Cecae ( <i>Torulopsis buchnerii</i> )                        |
|                               |                       | Mycetome between foregut and midgut                          |
|                               |                       | Mycetomes ( <i>Symbiotaphrina buchnerii</i> )                |
|                               |                       | Mycetomes and digestive tube ( <i>Torulopsis buchnerii</i> ) |
|                               |                       | Gut cecae ( <i>Symbiotaphrina buchnerii</i> )                |

|                                               |                          |                                                                       |
|-----------------------------------------------|--------------------------|-----------------------------------------------------------------------|
| <i>Lasioderma serricorne</i>                  | Coleoptera: Anobiidae    | Mycetome between foregut and midgut                                   |
|                                               |                          | ( <i>Symbiotaphrina kochii</i> )                                      |
|                                               |                          |                                                                       |
| <i>Ernobius abietis</i>                       | Coleoptera: Anobiidae    | Mycetomes ( <i>Torulopsis karawaiewii</i> )                           |
|                                               |                          | ( <i>Candida karawaiewii</i> )                                        |
| <i>Ernobius mollis</i>                        | Coleoptera: Anobiidae    | Mycetomes ( <i>Torulopsis ernobii</i> )                               |
|                                               |                          | ( <i>Candida ernobii</i> )                                            |
| <i>Hemicoelus gibbicollis</i>                 | Coleoptera: Anobiidae    | Larval mycetomes                                                      |
| <i>Xestobium plumbeum</i>                     | Coleoptera: Anobiidae    | Mycetomes ( <i>Torulopsis xestobii</i> )                              |
|                                               |                          | ( <i>Candida xestobii</i> )                                           |
| <i>Criocephalus rusticus</i>                  | Coleoptera: Cerambycidae | Mycetomes                                                             |
| <i>Phoracantha semipunctata</i>               | Coleoptera: Cerambycidae | Alimentary canal ( <i>Candida guilliermondii</i> , <i>C. tenuis</i> ) |
|                                               |                          | Cecae around midgut ( <i>Candida guilliermondii</i> )                 |
| <i>Harpium inquisitor</i>                     | Coleoptera: Cerambycidae | Mycetomes ( <i>Candida rhagii</i> )                                   |
| <i>Harpium mordax</i><br><i>H. sycophanta</i> | Coleoptera: Cerambycidae | Cecae around midgut ( <i>Candida tenuis</i> )                         |
| <i>Gaurotes virginea</i>                      | Coleoptera: Cerambycidae | Cecae around midgut ( <i>Candida rhagii</i> )                         |
|                                               |                          |                                                                       |
| <i>Leptura rubra</i>                          | Coleoptera: Cerambycidae | Cecae around midgut ( <i>Candida tenuis</i> )                         |
|                                               |                          | Cecae around midgut ( <i>Candida parapsilosis</i> )                   |
|                                               |                          |                                                                       |
| <i>Leptura maculicornis</i>                   | Coleoptera: Cerambycidae | Cecae around midgut ( <i>Candida parapsilosis</i> )                   |
|                                               |                          |                                                                       |
| <i>L. cerambyciformis</i>                     |                          |                                                                       |
| <i>Leptura sanguinolenta</i>                  | Coleoptera: Cerambycidae | Cecae around midgut ( <i>Candida</i> sp.)                             |
| <i>Rhagium bifasciatum</i>                    | Coleoptera: Cerambycidae | Cecae around midgut ( <i>Candida tenuis</i> )                         |

|                                  |                          |                                                                                                                                                                                                              |
|----------------------------------|--------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Rhagium inquisitor</i>        | Coleoptera: Cerambycidae | Cecae around midgut ( <i>Candida guilliermondii</i> )                                                                                                                                                        |
| <i>Rhagium mordax</i>            | Coleoptera: Cerambycidae | Cecae around midgut ( <i>Candida</i> )                                                                                                                                                                       |
| <i>Carpophilus hemipterus</i>    | Coleoptera: Nitidulidae  | Intestinal tract (10 yeast species)                                                                                                                                                                          |
| <i>Odontotaenius disjunctus</i>  | Coleoptera: Passalidae   | Hindgut ( <i>Enteroramus dimorphus</i> )                                                                                                                                                                     |
| <i>Odontotaenius disjunctus</i>  | Coleoptera: Passalidae   | Gut ( <i>Pichia stipitis</i> , <i>P. segobiensis</i> , <i>Candida shehatae</i> )                                                                                                                             |
| <i>Verres sternbergianus</i>     |                          | ( <i>C. ergatensis</i> )                                                                                                                                                                                     |
| <i>Scarabaeus semipunctatus</i>  | Coleoptera: Scarabaeidae | Digestive tract (10 yeast species)                                                                                                                                                                           |
| <i>Chironitis furcifer</i>       |                          |                                                                                                                                                                                                              |
| Unknown species                  | Coleoptera: Scarabaeidae | Guts ( <i>Trichosporon cutaneum</i> )                                                                                                                                                                        |
| <i>Dendroctonus and Ips spp.</i> | Coleoptera: Scolytidae   | Alimentary canal (13 yeast species)                                                                                                                                                                          |
| <i>Dendroctonus frontalis</i>    | Coleoptera: Scolytidae   | Midgut ( <i>Candida sp.</i> )                                                                                                                                                                                |
| <i>Ips sexdentatus</i>           | Coleoptera: Scolytidae   | Digestive tract ( <i>Pichia bovis</i> , <i>P. rhodanensis</i> )                                                                                                                                              |
|                                  |                          | Hansenula holstii ( <i>Candida rhagii</i> )                                                                                                                                                                  |
|                                  |                          | Digestive tract                                                                                                                                                                                              |
|                                  |                          | ( <i>Candida pulcherina</i> )                                                                                                                                                                                |
| <i>Ips typographus</i>           | Coleoptera: Scolytidae   | Alimentary canal                                                                                                                                                                                             |
|                                  |                          | Alimentary tracts ( <i>Hansenula capsulata</i> , <i>Candida parapsilosis</i> )                                                                                                                               |
|                                  |                          | Guts and beetle homogenates ( <i>Hansenula holstii</i> , <i>H. capsulata</i> , <i>Candida diddensii</i> , <i>C. mohshtana</i> , <i>C. nitratophila</i> , <i>Cryptococcus albidus</i> , <i>C. laurentii</i> ) |
| <i>Trypodendron lineatum</i>     | Coleoptera: Scolytidae   | Not specified                                                                                                                                                                                                |
| <i>Xyloterinus politus</i>       | Coleoptera: Scolytidae   | Head, thorax, abdomen ( <i>Candida</i> , <i>Pichia</i> , <i>Saccharomycopsis</i> )                                                                                                                           |
|                                  |                          |                                                                                                                                                                                                              |
|                                  |                          |                                                                                                                                                                                                              |
| <i>Periplaneta americana</i>     | Dictyoptera: Blattidae   | Hemocoel ( <i>Candida sp. nov.</i> )                                                                                                                                                                         |

|                                      |                             |                                                                         |
|--------------------------------------|-----------------------------|-------------------------------------------------------------------------|
| <i>Blatta orientalis</i>             | Dictyoptera: Blattidae      | Intestinal tract ( <i>Kluyveromyces blattae</i> )                       |
| <i>Blatella germanica</i>            | Dictyoptera: Blattellidae   | Hemocoel                                                                |
| <i>Cryptocercus punctulatus</i>      | Dictyoptera: Cryptocercidae | Hindgut (1 yeast species)                                               |
| <i>Philophylla heraclei</i>          | Diptera: Tephritidae        | Hemocoel                                                                |
| <i>Aedes</i> (4 species)             | Diptera: Culicidae          | Internal microflora (9 yeast genera)                                    |
| <i>Drosophila pseudoobscura</i>      | Diptera: Drosophilidae      | Alimentary canal (24 yeast species)                                     |
| <i>Drosophila</i> (5 spp.)           | Diptera: Drosophilidae      | Crop (42 yeast species)                                                 |
| <i>Drosophila melanogaster</i>       | Diptera: Drosophilidae      | Crop (8 yeast species)                                                  |
| <i>Drosophila</i> (4 spp.)           | Diptera: Drosophilidae      | Crop (7 yeast species)                                                  |
| <i>Drosophila</i> (6 spp.)           | Diptera: Drosophilidae      | Larval gut (17 yeast species)                                           |
| <i>Drosophila</i> (20 spp.)          | Diptera: Drosophilidae      | Crop (20 yeast species)                                                 |
| <i>Drosophila</i> (8 species groups) | Diptera: Drosophilidae      | Crop ( <i>Kloeckera</i> , <i>Candida</i> , <i>Kluyveromyces</i> )       |
| <i>Drosophila serido</i>             | Diptera: Drosophilidae      | Crop (18 yeast species)                                                 |
| <i>Drosophila</i> (6 spp.)           | Diptera: Drosophilidae      | Intestinal epithelium ( <i>Coccidiascus legeri</i> )                    |
|                                      |                             |                                                                         |
| <i>Protaxymia melanoptera</i>        | Diptera                     | Unknown ( <i>Candida</i> , <i>Cryptococcus</i> , <i>Sporoblomyces</i> ) |
| <i>Astegopteryx styraci</i>          | Homoptera: Aphididae        | Hemocoel and fat body                                                   |
| <i>Tuberaphis sp.</i>                | Homoptera: Aphididae        | Tissue sections                                                         |
| <i>Hamiltonaphis styraci</i>         |                             |                                                                         |
| <i>Glyphinaphis bambusae</i>         |                             |                                                                         |
| <i>Cerataphis sp.</i>                |                             |                                                                         |
| <i>Hamiltonaphis styraci</i>         | Homoptera: Aphididae        | Abdominal hemocoel                                                      |
| <i>Cofana unimaculata</i>            | Homoptera: Cicadellidae     | Fat body                                                                |
|                                      |                             |                                                                         |
|                                      |                             |                                                                         |
| <i>Leofa unicolor</i>                | Homoptera: Cicadellidae     | Fat body                                                                |
| <i>Lecaniines, etc.</i>              | Homoptera: Coccoidea d      | Hemolymph, fatty tissue, etc.                                           |
| <i>Lecanium sp.</i>                  | Homoptera: Coccidae         | Hemolymph, adipose tissue                                               |

|                                  |                         |                                                                                                                                                      |
|----------------------------------|-------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Ceroplastes (4 sp.)</i>       | Homoptera: Coccidae     | Blood smears                                                                                                                                         |
| <i>Laodelphax striatellus</i>    | Homoptera: Delphacidae  | Fat body                                                                                                                                             |
|                                  |                         |                                                                                                                                                      |
|                                  |                         | Eggs                                                                                                                                                 |
|                                  |                         | Eggs ( <i>Candida</i> )                                                                                                                              |
| <i>Nilaparvata lugens</i>        | Homoptera: Delphacidae  | Fat body                                                                                                                                             |
|                                  |                         | Eggs (2 unidentified yeast species)                                                                                                                  |
|                                  |                         | Eggs, nymphs ( <i>Candida</i> )                                                                                                                      |
|                                  |                         | Eggs (7 unidentified yeast species)                                                                                                                  |
|                                  |                         | Eggs ( <i>Candida</i> )                                                                                                                              |
| <i>Nisia nervosa</i>             | Homoptera: Delphacidae  | Fat body                                                                                                                                             |
| <i>Nisia grandiceps</i>          |                         |                                                                                                                                                      |
| <i>Perkinsiella spp.</i>         |                         |                                                                                                                                                      |
| <i>Sardia rostrata</i>           |                         |                                                                                                                                                      |
| <i>Sogatella furcifera</i>       |                         |                                                                                                                                                      |
| <i>Sogatodes orizicola</i>       | Homoptera: Delphacidae  | Fat body                                                                                                                                             |
| <i>Amrasca devastans</i>         | Homoptera: Jassidae     | Eggs, mycetomes, hemolymph                                                                                                                           |
| <i>Tachardina lobata</i>         | Homoptera: Kerriidae    | Blood smears ( <i>Torulopsis</i> )                                                                                                                   |
| <i>Laccifer (=Lakshadia) sp.</i> | Homoptera: Kerriidae    | Blood smears ( <i>Torula variabilis</i> )                                                                                                            |
|                                  |                         |                                                                                                                                                      |
| <i>Comperia merceti</i>          | Hymenoptera: Encyrtidae | Hemolymph, gut, poison gland                                                                                                                         |
| <i>Solenopsis invicta</i>        | Hymenoptera: Formicidae | Hemolymph ( <i>Myrmecomyces annellisae</i> )                                                                                                         |
| <i>S. quinquecuspis</i>          |                         |                                                                                                                                                      |
|                                  |                         |                                                                                                                                                      |
|                                  |                         |                                                                                                                                                      |
| <i>Solenopsis invicta</i>        | Hymenoptera: Formicidae | Fourth instar larvae ( <i>Candida parapsilosis</i> , <i>Yarrowia lipolytica</i> )                                                                    |
|                                  |                         | Gut and hemolymph ( <i>Candida parapsilosis</i> , <i>C. lipolytica</i> , <i>C. guillermondii</i> , <i>C. rugosa</i> , <i>Debaryomyces hansenii</i> ) |
| <i>Apis mellifera</i>            | Hymenoptera: Apidae     | Digestive tracts ( <i>Torulopsis sp.</i> )                                                                                                           |
|                                  |                         | Intestinal tract ( <i>Torulopsis apicola</i> )                                                                                                       |

|                                |                           |                                                 |
|--------------------------------|---------------------------|-------------------------------------------------|
|                                |                           | Digestive tracts (8 yeast species)              |
|                                |                           | Intestinal contents (12 yeast species)          |
|                                |                           | Intestinal contents (7 yeast species)           |
|                                |                           | Intestines (14 yeast species)                   |
|                                |                           | Intestinal tract ( <i>Pichia melissophila</i> ) |
|                                |                           | Intestinal tracts (7 yeast species)             |
|                                |                           | Alimentary canal ( <i>Hansenula silvicola</i> ) |
|                                |                           | Crop and gut (13 yeast species)                 |
| <i>Apis mellifera</i>          | Hymenoptera: Apidae       | Midguts (9 yeast genera)                        |
| <i>Anthophora occidentalis</i> | Hymenoptera:Anthophoridae |                                                 |
| <i>Nomia melanderi</i>         | Hymenoptera:Halictidae    |                                                 |
| <i>Halictus rubicundus</i>     | Hymenoptera:Halictidae    |                                                 |
| <i>Megachile rotundata</i>     | Hymenoptera:Megachilidae  |                                                 |

Any number of fungal species may be targeted by the compositions or methods described herein. For example, in some instances, the modulating agent may target a single fungal species. In some instances, the modulating agent may target at least about any of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 500, or more distinct fungal species. In some instances, the modulating agent may target any one of about 1 to about 5, about 5 to about 10, about 10 to about 20, about 20 to about 50, about 50 to about 100, about 100 to about 200, about 200 to about 500, about 10 to about 50, about 5 to about 20, or about 10 to about 100 distinct fungal species. In some instances, the modulating agent may target at least about any of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, or more phyla, classes, orders, families, or genera of fungi.

In some instances, the modulating agent may increase a population of one or more fungi by at least about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more in the host in comparison to a host organism to which the modulating agent has not been administered. In some instances, the modulating agent may reduce the population of one or more fungi by at least about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more in the host in comparison to a host organism to which the modulating agent has not been administered. In some instances, the modulating agent may eradicate the population of a fungi in the host.

In some instances, the modulating agent may alter the fungal diversity and/or fungal composition of the host. In some instances, the modulating agent may increase the fungal diversity in the host relative to a starting diversity by at least about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in comparison to a host organism to which the modulating agent has not been administered. In some

instances, the modulating agent may decrease the fungal diversity in the host relative to a starting diversity by at least about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in comparison to a host organism to which the modulating agent has not been administered.

In some instances, the modulating agent may alter the function, activity, growth, and/or division of one or more fungi. For example, the modulating agent may alter the expression of one or more genes in the fungus. In some instances, the modulating agent may alter the function of one or more proteins in the fungus. In some instances, the modulating agent may alter the function of one or more cellular components in the fungus. In some instances, the modulating agent may kill the fungus.

Further, the target fungus may reside in one or more parts of the insect. In some instances, the fungus resides in or on one or more parts of the insect gut, including, e.g., the foregut, midgut, and/or hindgut. In some instances, the fungus lives extracellularly in the hemolymph, fat bodies or in specialized structures in the host.

Changes to the population of fungi in the host may be determined by any methods known in the art, such as microarray, polymerase chain reaction (PCR), real-time PCR, flow cytometry, fluorescence microscopy, transmission electron microscopy, fluorescence *in situ* hybridization (e.g., FISH), spectrophotometry, matrix-assisted laser desorption ionization-mass spectrometry (MALDI-MS), and DNA sequencing. In some instances, a sample of the host treated with a modulating agent is sequenced (e.g., by metagenomics sequencing) to determine the microbiome of the host after delivery or administration of the modulating agent. In some instances, a sample of a host that did not receive the modulating agent is also sequenced to provide a reference.

### III. Modulating Agents

The modulating agent of the methods and compositions provided herein may include a polypeptide, a nucleic acid, small molecule, or any combination thereof. In some instances, the modulating agent is a nucleic acid molecule (e.g., DNA molecule or RNA molecule, e.g., mRNA, guide RNA (gRNA), or inhibitory RNA molecule (e.g., siRNA, shRNA, or miRNA), or a hybrid DNA-RNA molecule), a small molecule, a peptide, or a polypeptide (e.g., an antibody molecule, e.g., an antibody or antigen binding fragment thereof). Any of these agents can be used to alter the microbiota of a host invertebrate (e.g., insect, mollusk, or nematode) by targeting pathways in the host and/or resident microorganisms (e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways). For example, any modulating agents described herein may be used to regulate (e.g., to induce or to inhibit) a gene or protein in the host or a microorganism resident in the host (e.g., a protein or a gene encoding a protein listed in Table 7, Table 8, or Table 9).

#### *i. Polypeptides*

The modulating agent described herein may include a polypeptide (e.g., antibody). In some instances, the modulating agent described herein includes a polypeptide or functional fragments or derivative thereof, which target pathways in the host invertebrate (e.g., insect, mollusk, or nematode) and/or resident microorganisms (e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways). In some instances, the agent is a polypeptide listed

in Table 7, Table 8, or Table 9, wherein the primary sequence of the agent polypeptide is provided by reference to its accession number.

A modulating agent including a polypeptide as described herein can be contacted with the target host in an amount and for a time sufficient to: (a) reach a target level (e.g., a predetermined or threshold level) of concentration inside a target host; (b) reach a target level (e.g., a predetermined or threshold level) of concentration inside a target host gut; (c) reach a target level (e.g., a predetermined or threshold level) of concentration inside a target host bacteriocyte; (d) modulate the level, or an activity, of one or more microorganism (e.g., endosymbiont) in the target host; or/and (e) modulate fitness of the target host.

Polypeptides included herein may include naturally occurring polypeptides or recombinantly produced variants. In some instances, the polypeptide may be a functional fragments or variants thereof (e.g., an enzymatically active fragment or variant thereof of polypeptide listed in Table 7, Table 8, or Table 9). Such fragments or variants can be made and screened for similar activity as described herein and would be equivalent hereunder in the methods and compositions disclosed). For example, the polypeptide may be a functionally active variant of any of the polypeptides described herein with at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity, e.g., over a specified region or over the entire sequence, to a sequence of a polypeptide described herein or a naturally occurring polypeptide. In some instances, the polypeptide may have at least 50% (e.g., at least 50%, 60%, 70%, 80%, 90%, 95%, 97%, 99%, or greater) identity to the amino acid sequence of a protein listed in Table 7, Table 8, or Table 9 with reference to the accession number provided.

Methods of making a therapeutic polypeptide are routine in the art. See, in general, *Smales & James (Eds.), Therapeutic Proteins: Methods and Protocols (Methods in Molecular Biology), Humana Press (2005); and Crommelin, Sindelar & Meibohm (Eds.), Pharmaceutical Biotechnology: Fundamentals and Applications, Springer (2013).*

Methods for producing a polypeptide involve expression in mammalian cells, although recombinant proteins can also be produced using insect cells, yeast, bacteria, or other cells under the control of appropriate promoters. Mammalian expression vectors may comprise nontranscribed elements such as an origin of replication, a suitable promoter and enhancer, and other 5' or 3' flanking nontranscribed sequences, and 5' or 3' nontranslated sequences such as necessary ribosome binding sites, a polyadenylation site, splice donor and acceptor sites, and termination sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the other genetic elements required for expression of a heterologous DNA sequence. Appropriate cloning and expression vectors for use with bacterial, fungal, yeast, and mammalian cellular hosts are described in *Green & Sambrook, Molecular Cloning: A Laboratory Manual (Fourth Edition), Cold Spring Harbor Laboratory Press (2012).*

Various mammalian cell culture systems can be employed to express and manufacture a recombinant polypeptide agent (e.g., listed in Tables 4, 5, or 6). Examples of mammalian expression systems include CHO cells, COS cells, HeLA and BHK cell lines. Processes of host cell culture for production of protein therapeutics are described in, e.g., *Zhou and Kantardjieff (Eds.), Mammalian Cell Cultures for Biologics Manufacturing (Advances in Biochemical Engineering/Biotechnology), Springer*

(2014). Purification of protein therapeutics is described in *Franks, Protein Biotechnology: Isolation, Characterization, and Stabilization, Humana Press (2013); and in Cutler, Protein Purification Protocols (Methods in Molecular Biology), Humana Press (2010)*. Formulation of protein therapeutics is described in *Meyer (Ed.), Therapeutic Protein Drug Products: Practical Approaches to formulation in the Laboratory, Manufacturing, and the Clinic, Woodhead Publishing Series (2012)*.

The polypeptide modulating agents discussed hereinafter, namely antibodies, bacteriocins, antimicrobial peptides, and bacteriocyte regulatory peptides, can be used to alter pathways in the host that mediate interactions between the host and microorganisms resident in the host as indicated in the sections for increasing or decreasing the fitness of hosts.

#### (a) Antibodies

In some instances, the modulating agent includes an antibody or antigen binding fragment thereof. For example, an agent described herein may be an antibody that blocks or potentiates activity and/or function of a component of the host immune system pathway or bacteriocyte regulatory pathway listed in Table 8 or Table 9. The antibody may act as an antagonist or agonist of a polypeptide (e.g., enzyme or cell receptor) in the host or microorganisms resident in the host, including any proteins list in Table 7, Table 8, or Table 9.

The making and use of antibodies against a target antigen (e.g., proteins that mediate host-microbiota interactions, e.g., host immune system proteins or bacteriocyte proteins, e.g., proteins in Table 7, Table 8, or Table 9) is known in the art. See, for example, Zhiqiang An (Ed.), *Therapeutic Monoclonal Antibodies: From Bench to Clinic*, 1st Edition, Wiley, 2009 and also Greenfield (Ed.), *Antibodies: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press, 2013, for methods of making recombinant antibodies, including antibody engineering, use of degenerate oligonucleotides, 5'-RACE, phage display, and mutagenesis; antibody testing and characterization; antibody pharmacokinetics and pharmacodynamics; antibody purification and storage; and screening and labeling techniques.

#### (b) Bacteriocins

The modulating agent described herein may include a bacteriocin. In some instances, the bacteriocin is naturally produced by Gram-positive bacteria, such as *Pseudomonas*, *Streptomyces*, *Bacillus*, *Staphylococcus*, or lactic acid bacteria (LAB, such as *Lactococcus lactis*). In some instances, the bacteriocin is naturally produced by Gram-negative bacteria, such as *Hafnia alvei*, *Citrobacter freundii*, *Klebsiella oxytoca*, *Klebsiella pneumonia*, *Enterobacter cloacae*, *Serratia plymithicum*, *Xanthomonas campestris*, *Erwinia carotovora*, *Ralstonia solanacearum*, or *Escherichia coli*. Exemplary bacteriocins include, but are not limited to, Class I-IV LAB antibiotics (such as lantibiotics), colicins, microcins, and pyocins. Non-limiting examples of bacteriocins are listed in Table 3.

Table 3: Examples of Bacteriocins

| Class             | Name          | Producer                        | Targets                                                                                                                                                                                                                                                                                                     | Sequence                                                                         |
|-------------------|---------------|---------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| <b>Class I</b>    | Nisin         | <i>Lactococcus lactis</i>       | Active on Gram-positive bacteria:<br><i>Enterococcus</i> ,<br><i>Lactobacillus</i> ,<br><i>Lactococcus</i> ,<br><i>Leuconostoc</i> ,<br><i>Listeria</i> ,<br><i>Clostridium</i>                                                                                                                             | ITSISLCTPGCKT<br>GALMGCNMKTA<br>TCHCSIHVSK<br>(SEQ ID NO: 42)                    |
|                   | Epidermin     | <i>Staphylococcus epidermis</i> | Gram-positive bacteria                                                                                                                                                                                                                                                                                      | IASKFICTPGCAK<br>TGSFNSYCC<br>(SEQ ID NO: 43)                                    |
| <b>Class II</b>   |               |                                 |                                                                                                                                                                                                                                                                                                             |                                                                                  |
| <b>Class II a</b> | Pediocin PA-1 | <i>Pediococcus acidilactici</i> | <i>Pediococci</i> ,<br><i>Lactobacilli</i> ,<br><i>Leuconostoc</i> ,<br><i>Brochothrix thermosphacta</i> ,<br><i>Propionibacteria</i> ,<br><i>Bacilli</i> ,<br><i>Enterococci</i> ,<br><i>Staphylococci</i> ,<br><i>Listeria clostridia</i> ,<br><i>Listeria monocytogenes</i> ,<br><i>Listeria innocua</i> | KYYGNGVTCG<br>KHSCSVDWGK<br>ATTICIINNGAMA<br>WATGGHQGNHK<br>C<br>(SEQ ID NO: 44) |
| <b>Class II b</b> | Enterocin P   | <i>Enterococcus faecium</i>     | <i>Lactobacillus sakei</i> ,<br><i>Enterococcus faecium</i>                                                                                                                                                                                                                                                 | ATRSYGNGVYC<br>NNSKCWVNWGE<br>AKENIAGIVISGW<br>ASGLAGMGH<br>(SEQ ID NO: 45)      |
| <b>Class II c</b> | Lactococcin G | <i>Streptococcus lactis</i>     | Gram-positive bacteria                                                                                                                                                                                                                                                                                      | GTWDDIGQGIGR<br>VAYWVGKAMGN<br>MSDVNQASRINR<br>KKKH<br>(SEQ ID NO: 46)           |

|                     |                 |                                |                                                                                            |                                                                                                                                                            |
|---------------------|-----------------|--------------------------------|--------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Class II d</b>   | Lactacin-F      | <i>Lactobacillus johnsonii</i> | <i>Lactobacilli, Enterococcus faecalis</i>                                                 | NRWGDTVLSAA<br>SGAGTGIKACKS<br>FGPWGMAICGV<br>GGAAIGGYFGYT<br>HN<br>(SEQ ID NO: 47)                                                                        |
| <b>Class III</b>    |                 |                                |                                                                                            |                                                                                                                                                            |
| <b>Class III a</b>  | Enterocin AS-48 | <i>Enterococcus faecalis</i>   | Broad spectrum: Gram positive and Gram negative bacteria.                                  | MAKEFGIPAAVA<br>GTVLNVVEAGG<br>WVTTIVSILTAVG<br>SGGLSLLAAAGR<br>ESIKAYLKKEIKK<br>KGKRAVIAW<br>(SEQ ID NO: 48)                                              |
| <b>Class III b</b>  | Aureocin A70    | <i>Staphylococcus aureus</i>   | Broad spectrum: Gram positive and Gram negative bacteria.                                  | MSWLNFLKYIAK<br>YGKKAVSAAWK<br>YKGKVLEWLVN<br>GPTLEWVWQKL<br>KKIAGL<br>(SEQ ID NO: 49)                                                                     |
| <b>Class IV</b>     | Garvicin A      | <i>Lactococcus garvieae</i>    | Broad spectrum: Gram positive and Gram negative bacteria.                                  | IGGALGNALNGL<br>GTWANMMNGG<br>GFVNQWQVYAN<br>KGKINQYRPY<br>(SEQ ID NO: 50)                                                                                 |
| <b>Unclassified</b> | Colicin V       | <i>Escherichia coli</i>        | Active against <i>Escherichia coli</i> (also closely related bacteria), Enterobacteriaceae | MRTLTLNELDSV<br>SGGASGRDIAMA<br>IGTLSGQFVAGGI<br>GAAAGGVAGGAI<br>YDYASTHKPNPA<br>MSPSGLGGTIKQ<br>KPEGIPSEAWNY<br>AAGRLCNWSPN<br>NLSDVCL<br>(SEQ ID NO: 51) |

In some instances, the bacteriocin is a colicin, a pyocin, or a microcin produced by Gram-negative bacteria. In some instances, the bacteriocin is a colicin. The colicin may be a group A colicin (e.g., uses the Tol system to penetrate the outer membrane of a target bacterium) or a group B colicin (e.g., uses the Ton system to penetrate the outer membrane of a target bacterium). In some instances,

the bacteriocin is a microcin. The microcin may be a class I microcin (e.g., < 5 kDa, has post-translational modifications) or a class II microcin (e.g., 5-10 kDa, with or without post-translational modifications). In some instances, the class II microcin is a class IIa microcin (e.g., requires more than one genes to synthesize and assemble functional peptides) or a class IIb microcin (e.g., linear peptides with or without post-translational modifications at C-terminus). In some instances, the bacteriocin is a pyocin. In some instances, the pyocin is an R-pyocin, F-pyocin, or S-pyocin.

In some instances, the bacteriocin is a class I, class II, class III, or class IV bacteriocin produced by Gram-positive bacteria. In some instances, the modulating agent includes a Class I bacteriocin (e.g., lanthionine-containing antibiotics (lantibiotics) produced by a Gram-positive bacteria). The class I bacteriocins or lantibiotic may be a low molecular weight peptide (e.g., less than about 5 kDa) and may possess post-translationally modified amino acid residues (e.g., lanthionine,  $\beta$ -methyllanthionine, or dehydrated amino acids).

In some instances, the bacteriocin is a Class II bacteriocin (e.g., non-lantibiotics produced by Gram-positive bacteria). Many are positively charged, non-lanthionine-containing peptides, which unlike lantibiotics, do not undergo extensive post-translational modification. The Class II bacteriocin may belong to one of the following subclasses: "pediocin-like" bacteriocins (e.g., pediocin PA-1 and carnobacteriocin X (Class IIa)); two-peptide bacteriocins (e.g., lactacin F and ABP-118 (Class IIb)); circular bacteriocins (e.g., carnocyclin A and enterocin AS-48 (Class IIc)); or unmodified, linear, non-pediocin-like bacteriocins (e.g., epidermicin NI01 and lactococcin A (Class IId)).

In some instances, the bacteriocin is a Class III bacteriocin (e.g., produced by Gram-positive bacteria). Class III bacteriocins may have a molecular weight greater than 10 kDa and may be heat unstable proteins. The Class III bacteriocins can be further subdivided into Group IIIA and Group IIIB bacteriocins. The Group IIIA bacteriocins include bacteriolytic enzymes which kill sensitive strains by lysis of the cell wall, such as Enterolisin A. Group IIIB bacteriocins include non-lytic proteins, such as Caseicin 80, Helveticin J, and lactacin B.

In some instances, the bacteriocin is a Class IV bacteriocin (e.g., produced by Gram-positive bacteria). Class IV bacteriocins are a group of complex proteins, associated with other lipid or carbohydrate moieties, which appear to be required for activity. They are relatively hydrophobic and heat stable. Examples of Class IV bacteriocins leuconocin S, lactocin 27, and lactocin S.

In some instances, the bacteriocin is an R-type bacteriocin. R-type bacteriocins are contractile bacteriocidal protein complexes. Some R-type bacteriocins have a contractile phage-tail-like structure. The C-terminal region of the phage tail fiber protein determines target-binding specificity. They may attach to target cells through a receptor-binding protein, e.g., a tail fiber. Attachment is followed by sheath contraction and insertion of the core through the envelope of the target bacterium. The core penetration results in a rapid depolarization of the cell membrane potential and prompt cell death. Contact with a single R-type bacteriocin particle can result in cell death. An R-type bacteriocin, for example, may be thermolabile, mild acid resistant, trypsin resistant, sedimentable by centrifugation, resolvable by electron microscopy, or a combination thereof. Other R-type bacteriocins may be complex molecules including multiple proteins, polypeptides, or subunits, and may resemble a tail structure of bacteriophages of the myoviridae family. In naturally occurring R-type bacteriocins, the subunit structures

may be encoded by a bacterial genome, such as that of *C. difficile* or *P. aeruginosa* and form R-type bacteriocins to serve as natural defenses against other bacteria. In some instances, the R-type bacteriocin is a pyocin. In some instances, the pyocin is an R-pyocin, F-pyocin, or S-pyocin.

In some instances, the bacteriocin is a functionally active variant of the bacteriocins described herein. In some instances, the variant of the bacteriocin has at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity, e.g., over a specified region or over the entire sequence, to a sequence of a bacteriocin described herein or a naturally occurring bacteriocin.

In some instances, the bacteriocin may be bioengineered, according to standard methods, to modulate their bioactivity, e.g., increase or decrease or regulate, or to specify their target microorganisms. In other instances, the bacteriocin is produced by the translational machinery (e.g. a ribosome, etc.) of a microbial cell. In some instances, the bacteriocin is chemically synthesized. Some bacteriocins can be derived from a polypeptide precursor. The polypeptide precursor can undergo cleavage (e.g., processing by a protease) to yield the polypeptide of the bacteriocin itself. As such, in some instances, the bacteriocin is produced from a precursor polypeptide. In some other instances, the bacteriocin includes a polypeptide that has undergone post-translational modifications, for example, cleavage, or the addition of one or more functional groups.

The bacteriocins described herein may be formulated in a composition for any of the uses described herein. The compositions disclosed herein may include any number or type (e.g., classes) of bacteriocins, such as at least about any one of 1 bacteriocin, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 100, or more bacteriocins. Suitable concentrations of each bacteriocin in the compositions described herein depends on factors such as efficacy, stability of the bacteriocin, number of distinct bacteriocin types in the compositions, formulation, and methods of application of the composition. In some instances, each bacteriocin in a liquid composition is from about 0.01 ng/ml to about 100 mg/mL. In some instances, each bacteriocin in a solid composition is from about 0.01 ng/g to about 100 mg/g. In some instances, wherein the composition includes at least two types of bacteriocins, the concentration of each type of the bacteriocins may be the same or different. In some instances, the bacteriocin is provided in a composition including a bacterial cell that secretes the bacteriocin. In some instances, the bacteriocin is provided in a composition including a polypeptide (e.g., a polypeptide isolated from a bacterial cell).

Bacteriocins may neutralize (e.g., kill) at least one microorganism other than the individual bacterial cell in which the polypeptide is made, including cells clonally related to the bacterial cell and other microbial cells. As such, a bacterial cell may exert cytotoxic or growth-inhibiting effects on a plurality of microbial organisms by secreting bacteriocins. In some instances, the bacteriocin targets and kills one or more species of bacteria resident in the host via cytoplasmic membrane pore formation, cell wall interference (e.g., peptidoglycanase activity), or nuclease activity (e.g., DNase activity, 16S rRNase activity, or tRNase activity).

In some instances, the bacteriocin has a neutralizing activity. Neutralizing activity of bacteriocins may include, but is not limited to, arrest of microbial reproduction, or cytotoxicity. Some bacteriocins have cytotoxic activity, and thus can kill microbial organisms, for example bacteria, yeast, algae, and the like.

Some bacteriocins can inhibit the reproduction of microbial organisms, for example bacteria, yeast, algae, and the like, for example by arresting the cell cycle.

In some instances, the bacteriocin has killing activity. The killing mechanism of bacteriocins is specific to each group of bacteriocins. In some instances, the bacteriocin has narrow-spectrum  
5 bioactivity. Bacteriocins are known for their very high potency against their target strains. Some bacteriocin activity is limited to strains that are closely related to the bacteriocin producer strain (narrow-spectrum bioactivity). In some instances, the bacteriocin has broad-spectrum bioactivity against a wide range of genera.

In some instances, bacteriocins interact with a receptor molecule or a docking molecule on the  
10 target bacterial cell membrane. For example, nisin is extremely potent against its target bacterial strains, showing antimicrobial activity even at a single-digit nanomolar concentration. The nisin molecule has been shown to bind to lipid II, which is the main transporter of peptidoglycan subunits from the cytoplasm to the cell wall

In some instances, the bacteriocin has anti-fungal activity. A number of bacteriocins with anti-  
15 yeast or anti-fungal activity have been identified. For example, bacteriocins from *Bacillus* have been shown to have neutralizing activity against some yeast strains (see, for example, Adetunji and Olaoye, *Malaysian Journal of Microbiology* 9:130-13, 2013). In another example, an *Enterococcus faecalis* peptide has been shown to have neutralizing activity against *Candida* species (see, for example, Shekh and Roy, *BMC Microbiology* 12:132, 2012). In another example, bacteriocins from *Pseudomonas* have  
20 been shown to have neutralizing activity against fungi, such as *Curvularia lunata*, *Fusarium species*, *Helminthosporium species*, and *Biopolaris species* (see, for example, Shalani and Srivastava, *The Internet Journal of Microbiology* Volume 5 Number 2, 2008). In another example, botrycidin AJ1316 and alirin B1 from *B. subtilis* have been shown to have antifungal activities.

A modulating agent including a bacteriocin as described herein can be contacted with the target  
25 host in an amount and for a time sufficient to: (a) reach a target level (e.g., a predetermined or threshold level) of bacteriocin concentration inside a target host; (b) reach a target level (e.g., a predetermined or threshold level) of bacteriocin concentration inside a target host gut; (c) reach a target level (e.g., a predetermined or threshold level) of bacteriocin concentration inside a target host bacteriocyte; (d) modulate the level, or an activity, of one or more microorganism (e.g., endosymbiont) in the target host;  
30 or/and (e) modulate fitness of the target host.

As illustrated by Example 5, bacteriocins (e.g., colA produced by a transgenic plant) can be used  
as a modulating agent that targets a host pathway (e.g., an insect, e.g., an aphid) that alters the activity, levels, or metabolism of endosymbiotic bacteria resident in the host, such as a *Buchnera* spp., to  
modulate (e.g., decrease) the fitness of the host.

### (c) Antimicrobial Peptides

The modulating agent described herein may include an antimicrobial peptide (AMP). Any AMP  
suitable for inhibiting a microorganism resident in the host may be used. AMPs are a diverse group of molecules, which are divided into subgroups on the basis of their amino acid composition and structure.  
40 The AMP may be derived or produced from any organism that naturally produces AMPs, including AMPs

derived from plants (e.g., copsin), insects (e.g., mastoparan, poneratoxin, cecropin, moricin, melittin), frogs (e.g., magainin, dermaseptin, aurein), and mammals (e.g., cathelicidins, defensins and protegrins). Non-limiting examples of AMPs are listed in Table 4.

5 **Table 4: Examples of Antimicrobial Peptides**

| Type                                              | Characteristic                                                | Example AMP | Sequence                                                                                                                                                                                             |
|---------------------------------------------------|---------------------------------------------------------------|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Anionic peptides                                  | rich in glutamic and aspartic acid                            | dermcidin   | SSLLEKGLDGAKKAVGGLGKL<br>GKDAVEDLESVGKGAVHVDKD<br>VLDSVL<br>(SEQ ID NO: 52)                                                                                                                          |
| Linear cationic $\alpha$ -helical peptides        | lack cysteine                                                 | cecropin A  | KWKLFKKIEKVGQNIRDGIIKAG<br>PAVAVVGQATQIAK<br>(SEQ ID NO: 53)                                                                                                                                         |
|                                                   |                                                               | andropin    | MKYFSVLVVLTLILAIVDQSDAFI<br>NLLDKVEDALHTGAQAGFKLIR<br>PVERGATPKKSEKPEK<br>(SEQ ID NO: 54)                                                                                                            |
|                                                   |                                                               | moricin     | MNILKFFFVFIVAMSLVSCSTAA<br>PAKIPIKAIKTVGKAVGKGLRAI<br>NIASTANDVFNFLKPKKRKH<br>(SEQ ID NO: 55)                                                                                                        |
|                                                   |                                                               | ceratotoxin | MANLKAVFLICIVAFIALQCVVA<br>EPAAEDSVVVKRSIGSALKKAL<br>PVAKKIGKIALPIAKAALPVAAG<br>LVG<br>(SEQ ID NO: 56)                                                                                               |
| Cationic peptide enriched for specific amino acid | rich in proline, arginine, phenylalanine, glycine, tryptophan | abaecin     | MKVVFIFALLATICA AFAYVPLP<br>NVPQPGRRPFPTFPGQGPFNP<br>KIKWPQGY<br>(SEQ ID NO: 57)                                                                                                                     |
|                                                   |                                                               | apidaecins  | KNFALAILVVTFFVAVFGNTNLD<br>PPTRPTRLRREAKPEAEPGNN<br>RPVYIPQPRPPHPRLRREAPE<br>AEPGNNRPVYIPQPRPPHPRL<br>RREAEEAEPGNNRPVYISQP<br>RPPHPRLRREAPEAEPGNNR<br>PVYIPQPRPPHPRLRREAEEA<br>EPGNNRPVYISQPRPPHPRLR |

|                                                                              |                            |              |                                                                                                                                                                                                                                                                                              |
|------------------------------------------------------------------------------|----------------------------|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                              |                            |              | REAEPEAEPGNNRPVYIPQPR<br>PPHPRLRREAEPGNNRP<br>VYIPQPRPPHPRLRREAEP<br>PGNNRPVYIPQPRPPHPRLRR<br>EAKPEAKPGNNRPVYIPQPRP<br>PHPRI<br>(SEQ ID NO: 58)                                                                                                                                              |
|                                                                              |                            | prophenin    | METQRASLCLGRWSLWLLLLA<br>LVVPSASAQALSYREAVLRAVD<br>RLNEQSSEANLYRLLELDQPPK<br>ADEDPGTPKPVSTVKETVCP<br>RPTRRPPELCDFKENGVRVKQC<br>VGTVTLDQIKDPLDITCNEGVR<br>RFPWWWPFLRRPRLRRQAFP<br>PPNVPGPRFPPPNVPGPRFPP<br>PNFPGPRFPPPNFPGRFPPP<br>NFPGPPFPPPIFPGWFPPPPP<br>FRPPFPGPRFPGRR<br>(SEQ ID NO: 59) |
|                                                                              |                            | indolicidin  | MQTQRASLSLGRWSLWLLLLG<br>LVVPSASAQALSYREAVLRAVD<br>QLNELSSEANLYRLLELDPPPK<br>DNEDLGTRKPVSTVKETVCP<br>RTIQQPAEQCDFKEKGRVKQC<br>VGTVTLDPSNDQFDLNCNELQ<br>SVILPWKWPWWPWRRG<br>(SEQ ID NO: 60)                                                                                                   |
| Anionic and cationic peptides that contain cysteine and form disulfide bonds | contain 1-3 disulfide bond | protegrin    | METQRASLCLGRWSLWLLLLA<br>LVVPSASAQALSYREAVLRAVD<br>RLNEQSSEANLYRLLELDQPPK<br>ADEDPGTPKPVSTVKETVCP<br>RPTRQPPELCDFKENGVRVKQC<br>VGTVTLDQIKDPLDITCNEVQG<br>VRGGRLCYCRRRFCVCVGRG<br>(SEQ ID NO: 61)                                                                                             |
|                                                                              |                            | tachyplesins | KWCFRVCYRGICYRRCR<br>(SEQ ID NO: 62)                                                                                                                                                                                                                                                         |
|                                                                              |                            | defensin     | MKCATIVCTIAVVLAATLLNGSV<br>QAAPQEEAALSGGANLNTLLD<br>ELPEETHHAALenyRAKRATC                                                                                                                                                                                                                    |

|  |  |            |                                                                                                      |
|--|--|------------|------------------------------------------------------------------------------------------------------|
|  |  |            | DLASGFGVGSSSLCAAHCIARR<br>YRGGYCNSKAVCVCRN<br>(SEQ ID NO: 63)                                        |
|  |  | drosomycin | MMQIKYLFALFAVLMLVVLGAN<br>EADADCLSGRYKGPCAVWDN<br>ETCRRVCKEEGRSSGHCSPL<br>KCWCEGC<br>(SEQ ID NO: 64) |

The AMP may be active against any number of target microorganisms. In some instances, the AMP may have antibacterial and/or antifungal activities. In some instances, the AMP may have a narrow-spectrum bioactivity or a broad-spectrum bioactivity. For example, some AMPs target and kill only a few species of bacteria or fungi, while others are active against both gram-negative and gram-positive bacteria as well as fungi.

Further, the AMP may function through a number of known mechanisms of action. For example, the cytoplasmic membrane is a frequent target of AMPs, but AMPs may also interfere with DNA and protein synthesis, protein folding, and cell wall synthesis. In some instances, AMPs with net cationic charge and amphipathic nature disrupt bacterial membranes leading to cell lysis. In some instances, AMPs may enter cells and interact with intracellular target to interfere with DNA, RNA, protein, or cell wall synthesis. In addition to killing microorganisms, AMPs have demonstrated a number of immunomodulatory functions that are involved in the clearance of infection, including the ability to alter host gene expression, act as chemokines and/or induce chemokine production, inhibit lipopolysaccharide induced pro-inflammatory cytokine production, promote wound healing, and modulating the responses of dendritic cells and cells of the adaptive immune response.

In some instances, the AMP is a functionally active variant of the AMPs described herein. In some instances, the variant of the AMP has at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity, e.g., over a specified region or over the entire sequence, to a sequence of an AMP described herein or a naturally derived AMP.

In some instances, the AMP may be bioengineered to modulate its bioactivity, e.g., increase or decrease or regulate, or to specify a target microorganism. In some instances, the AMP is produced by the translational machinery (e.g. a ribosome, etc.) of a cell. In some instances, the AMP is chemically synthesized. In some instances, the AMP is derived from a polypeptide precursor. The polypeptide precursor can undergo cleavage (for example, processing by a protease) to yield the polypeptide of the AMP itself. As such, in some instances, the AMP is produced from a precursor polypeptide. In some instances, the AMP includes a polypeptide that has undergone post-translational modifications, for example, cleavage, or the addition of one or more functional groups.

The AMPs described herein may be formulated in a composition for any of the uses described herein. The compositions disclosed herein may include any number or type (e.g., classes) of AMPs, such as at least about any one of 1 AMP, 2, 3, 4, 5, 10, 15, 20, or more AMPs. A suitable concentration of

each AMP in the composition depends on factors such as efficacy, stability of the AMP, number of distinct AMP in the composition, the formulation, and methods of application of the composition. In some instances, each AMP in a liquid composition is from about 0.1 ng/mL to about 100 mg/mL. In some instances, each AMP in a solid composition is from about 0.1 ng/g to about 100 mg/g. In some instances, wherein the composition includes at least two types of AMPs, the concentration of each type of AMP may be the same or different.

A modulating agent including an AMP as described herein can be contacted with the target host in an amount and for a time sufficient to: (a) reach a target level (e.g., a predetermined or threshold level) of AMP concentration inside a target host; (b) reach a target level (e.g., a predetermined or threshold level) of AMP concentration inside a target host gut; (c) reach a target level (e.g., a predetermined or threshold level) of AMP concentration inside a target host bacteriocyte; (d) modulate the level, or an activity, of one or more microorganism (e.g., endosymbiont) in the target host; or/and (e) modulate fitness of the target host.

*(d) Bacteriocyte Regulatory Peptides*

The modulating agent described herein may include a bacteriocyte regulatory peptide (BRP). BRPs are peptides expressed in the bacteriocytes of insects. These genes are expressed first at a developmental time point coincident with the incorporation of symbionts and their bacteriocyte-specific expression is maintained throughout the insect's life. In some instances, the BRP has a hydrophobic amino terminal domain, which is predicted to be a signal peptide. In addition, some BRPs have a cysteine-rich domain. In some instances, the bacteriocyte regulatory peptide is a bacteriocyte-specific cysteine rich (BCR) protein. Bacteriocyte regulatory peptides have a length between about 40 and 150 amino acids. In some instances, the bacteriocyte regulatory peptide has a length in the range of about 45 to about 145, about 50 to about 140, about 55 to about 135, about 60 to about 130, about 65 to about 125, about 70 to about 120, about 75 to about 115, about 80 to about 110, about 85 to about 105, or any range therebetween. Non-limiting examples of BRPs and their activities are listed in Table 5.

**Table 5: Examples of Bacteriocyte Regulatory Peptides**

| Name                                                                  | Peptide Sequence                                                                                                 |
|-----------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|
| Bacteriocyte-specific cysteine rich proteins BCR family, peptide BCR1 | MKLLHGFLIIMLTMHLSIQYAYGGPFLT KYLC DRVCHKLC<br>GDEFVCSCIQYKSLKGLWFP HCPTGKASVVLHNFLTSP<br>(SEQ ID NO: 65)         |
| Bacteriocyte-specific cysteine rich proteins BCR family, peptide BCR2 | MKLLYGFLIIMLTIHLSVQYFESPFETKYNC DTHCNKLCGK<br>IDHCSCIQYHSM EGLWFP HCRTGSAAQMLHD FLSNP<br>(SEQ ID NO: 66)         |
| Bacteriocyte-specific cysteine rich proteins BCR family, peptide BCR3 | MSVRKNVLPTMFVLLIMSPVTPTSVFISAVCYSGCGSLA<br>LVCFVSNGITNGLDYFKSSAPLSTSETSCGEAFDTCTDH<br>CLANFKF<br>(SEQ ID NO: 67) |

|                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-----------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Bacteriocyte-specific cysteine rich proteins BCR family, peptide BCR4 | MRLLYGFLIIMLTIYLSVQDFDPTEFKGPFPTIEICSKYCAV<br>VCNYTSRPCYCVEAAKERDQWFPYCYD<br>(SEQ ID NO: 68)                                                                                                                                                                                                                                                                                                                                                                                                                        |
| Bacteriocyte-specific cysteine rich proteins BCR family, peptide BCR5 | MRLLYGFLIIMLTIHLSVQDIDPNTLRGPYPTKEICSKYCEY<br>NVVCGASLPCICVQDARQLDHWFACCYDGGPEMLM<br>(SEQ ID NO: 69)                                                                                                                                                                                                                                                                                                                                                                                                                |
| Secreted proteins SP family, peptide SP1                              | MKLFVVVVLVAVGIMFVFASDTAAAPT DYEDTNDMISLSS<br>LVGDNSPYVRVSSADSGGSSKTSSKNPILGLLKSVIKLLT<br>KIFGTYSDAAPAMPPIPPALRKNRGMLA<br>(SEQ ID NO: 70)                                                                                                                                                                                                                                                                                                                                                                            |
| Secreted proteins SP family, peptide SP2                              | MVACKVILAVAVVFVAAVQGRPGGEPEWAAPIFAELKSV<br>SDNITNLVGLDNAGEYATAAKNNLNAFAESLKTEAAVFSK<br>SFEGKASASDVFKESTKNFQAVVD TYIKNLPKDLTLKDFT<br>EKSEQALKYMVEHGTEITKKAQGNTETEKEIKEFFKKQIE<br>NLIGQGKALQAKIAEAKKA<br>(SEQ ID NO: 71)                                                                                                                                                                                                                                                                                              |
| Secreted proteins SP family, peptide SP3                              | MKTSSSKVFASCVAI VCLASVANALPVQKSVAATTENPIV<br>EKHGCRAHKNLVRQNVVDLKT YDSMLITNEVVQKQSNE<br>VQSSEQSNEGQNSEQSNEGQNSEQSN EVQSSEHSNEG<br>QNSKQSNEGQNSEQSN EVQSSEHSNEGQNSEQSN EVQ<br>SSEHSNEGQNSKQSNEGQNSKQSN EVQSSEHWN EGQ<br>NSKQSNEDQNSEQSN EGQNSKQSN EGQNSKQSNEDQ<br>NSEQSN EGQNSKQSN EVQSSEQSN EGQNSKQSN EGQS<br>SEQSN EGQNSKQSN EVQSP EEHYDLDP ESSYESEETK<br>GSHESGDDSEHR<br>(SEQ ID NO: 72)                                                                                                                          |
| Secreted proteins SP family, peptide SP4                              | MKTIILGLCLFGALFWSTQSM PVGEVAPAVPAVPSEAVP<br>QKQVEAKPETNAASPVSDAKPESDSKPVDAEVKPTVSEV<br>KAESQKPSGEPKPESDAKPVVASESKPESDPKPAAVVE<br>SKPENDAVAPETNND AKPENAAAPVSENKPATDAKAETE<br>LIAQAKPESKPASDLKAEPEAAKPNSEVPVALPLNPTETK<br>ATQQSVETNQVEQAAPAAAQADPAAAPAADPAPAPAAA<br>PVAAEEAKLSESAPSTENKAAE EPSKPAEQQSAKPVEDA<br>VPAASEISETKVSPAVPAVPEVPASPSAPAVADPVSAPEA<br>EKNAEPAKAANSAEPAVQSEAKPAEDIQKSGAVVSAENP<br>KPV EEQKPAEVAKPAEQSKSEAPAEAPKPTEQSAAE EPK<br>KPESANDEKKEQHSVNKR DATKEKKPTDSIMKKQKQKK<br>AN<br>(SEQ ID NO: 73) |

|                                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Secreted proteins SP family, peptide SP5a | MNGKIVLCFAVFIGQAMSAATGTTPEVEDIKKVAEQMS<br>QTFMSVANHLVGITPNSADAQKSIEKIRTIMNKGFTDMET<br>EANKMKDIVRKNADPKLVEKYDELEKELKKHLSTAKDMF<br>EDKVVKPIGEKVELKKITENVIKTTKDMEATMNKAIDGFKK<br>Q<br>(SEQ ID NO: 74)                                                                                                                                                                                                                                                                                       |
| Secreted proteins SP family, peptide SP6  | MHLFLALGLFIVCGMVDA TFYNPRSQTFNQLMERRQRSI<br>PIPYSGYHYNPIEPSINVLDLSEGLDSRINTFKPIYQNV<br>KMSTQDVNSVPRTQYQPKNSLYDSEYISAKDIPSLFPEE<br>DSYDYKYLGSPLNKYLTRPSTQESGIAINLVAIKETSVFDY<br>GFPTYKSPYSSDSVWNFGSKIPNTVFEDPQSVESDPNTF<br>KVSSPTIKIVKLLPETPEQESIITTTKNYELNYKTTQETPTE<br>AELYPITSEEFQTEDEWHPMVPKENTTKDESSFITTEEPL<br>TEDKSNSITIEKTQTEDESNSIEFNIRTEEKSNSITTEENQ<br>KEDDESMSTTSQETTTAFNLNDTFDTNRYSSSHESLMLR<br>IRELMKNIADQQNKSQFRTVDNIPAKSQSNLSSDESTNQ<br>QFEPQLVNGADTYK<br>(SEQ ID NO: 75) |
| Coleopteracin A, ColA peptide             | MTRTMLFLACVAALYVCISATAGKPEEFAKLSDEAPSND<br>QAMYESIQRYRRFVDGNRYNGGQQQQQQPKQWEVRP<br>DLSRDQRGNTKAQVEINKKGDNHDINAGWGKNINGPDS<br>HKDTWHVGGSVRW<br>(SEQ ID NO: 76)                                                                                                                                                                                                                                                                                                                            |
| RlpA Type I                               | MKETTVVWAKLFLILILAKPLGLKAVNECKRLGNNSCRSH<br>GECCSGFCFIEPGWALGVCKRLGTPKKSDDSNNGKNIEK<br>NNGVHERIDDVFERGVCSYYKGPSITANGDVFDENEMTA<br>AHRTLFPNTMVKVEGMGTSVVVKINDRKTAAADGKVMLLS<br>RAAAESLNIDENTGPVQCQLKFVLDGSGCTPDYGDTCVL<br>HHECCSQNCFREMFSDKGFCLPK<br>(SEQ ID NO: 77)                                                                                                                                                                                                                      |

In some instances, the BRP alters the growth and/or activity of one or more bacteria resident in the bacteriocyte of the host. In some instances, the BRP may be bioengineered to modulate its bioactivity (e.g., increase, decrease, or regulate) or to specify a target microorganism. In some instances, the BRP is produced by the translational machinery (e.g. a ribosome, etc.) of a cell. In some instances, the BRP is chemically synthesized. In some instances, the BRP is derived from a polypeptide precursor. The polypeptide precursor can undergo cleavage (for example, processing by a protease) to yield the polypeptide of the BRP itself. As such, in some instances, the BRP is produced from a precursor polypeptide. In some instances, the BRP includes a polypeptide that has undergone post-translational modifications, for example, cleavage, or the addition of one or more functional groups.

Functionally active variants of the BRPs as described herein are also useful in the compositions and methods described herein. In some instances, the variant of the BRP has at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity, e.g., over a specified region or over the entire sequence, to a sequence of a BRP described herein or naturally derived BRP.

The BRP described herein may be formulated in a composition for any of the uses described herein. The compositions disclosed herein may include any number or type (e.g., classes) of BRPs, such as at least about any one of 1 BRP, 2, 3, 4, 5, 10, 15, 20, or more BRPs. A suitable concentration of each BRP in the composition depends on factors such as efficacy, stability of the BRP, number of distinct BRP, the formulation, and methods of application of the composition. In some instances, each BRP in a liquid composition is from about 0.1 ng/mL to about 100 mg/mL. In some instances, each BRP in a solid composition is from about 0.1 ng/g to about 100 mg/g. In some instances, wherein the composition includes at least two types of BRPs, the concentration of each type of BRP may be the same or different.

A modulating agent including a BRP as described herein can be contacted with the target host in an amount and for a time sufficient to: (a) reach a target level (e.g., a predetermined or threshold level) of BRP concentration inside a target host; (b) reach a target level (e.g., a predetermined or threshold level) of BRP concentration inside a target host gut; (c) reach a target level (e.g., a predetermined or threshold level) of BRP concentration inside a target host bacteriocyte; (d) modulate the level, or an activity, of one or more microorganism (e.g., endosymbiont) in the target host; or/and (e) modulate fitness of the target host.

## *ii. Small Molecules*

In some instances, the modulating agent includes a small molecule. Numerous small molecule agents are useful in the methods and compositions described herein. The small molecules discussed hereinafter can be used to alter pathways in host that mediate interactions between the host and microorganisms resident in the host, as indicated in the sections for decreasing the fitness of insects, such as aphids. Additional small molecule agents can also be screened based on their ability to target components (e.g., polypeptides, e.g., enzymes or cell surface receptors) of pathways in the host invertebrate (e.g., insect, mollusk, or nematode) and/or resident microorganism (e.g., polypeptides that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways). In some instances, the small molecule includes an agonist, antagonist, inhibitor, or an activator. For example, a small molecule described herein may be an agonist, antagonist, inhibitor, or an activator that blocks or potentiates activity and/or function of a component of the host immune system pathway or bacteriocyte regulatory pathway listed in Table 8 or Table 9. The small molecule may act as an antagonist or agonist of a polypeptide (e.g., enzyme or cell receptor) in the host or microorganisms resident in the host, including any proteins list in Table 7, Table 8, or Table 9.

Small molecules include, but are not limited to, small peptides, peptidomimetics (e.g., peptoids), amino acids, amino acid analogs, synthetic polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic and inorganic compounds (including heterorganic and organometallic compounds) generally having a molecular weight less than about 5,000 grams per mole, e.g., organic or

inorganic compounds having a molecular weight less than about 2,000 grams per mole, e.g., organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, e.g., organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds.

5 The small molecule described herein may be formulated in a composition for any of the uses described herein. The compositions disclosed herein may include any number or type (e.g., classes) of small molecules, such as at least about any one of 1 small molecule, 2, 3, 4, 5, 10, 15, 20, or more small molecules. A suitable concentration of each small molecule in the composition depends on factors such as efficacy, stability of the small molecule, number of distinct small molecules, the formulation, and  
10 methods of application of the composition. In some instances, wherein the composition includes at least two types of small molecules, the concentration of each type of small molecule may be the same or different.

A modulating agent including a small molecule as described herein can be contacted with the target host in an amount and for a time sufficient to: (a) reach a target level (e.g., a predetermined or  
15 threshold level) of small molecule concentration inside a target host; (b) reach a target level (e.g., a predetermined or threshold level) of small molecule concentration inside a target host gut; (c) reach a target level (e.g., a predetermined or threshold level) of small molecule concentration inside a target host bacteriocyte; (d) modulate the level, or an activity, of one or more microorganism (e.g., endosymbiont) in the target host; or/and (e) modulate fitness of the target host.

20 In some instances, the small molecule triggers, stimulates, or increases a host's immune response in comparison to a host organism to which the small molecule has not been administered. For example, the small molecule may be peptidoglycan molecule that activates the ROS system in insects by binding to the epithelial cell surface, which in turns induces DUOX enzymatic activity by mobilizing intracellular calcium. In another example, the molecule dihydroxyphenylalanine (DOPA) is an essential  
25 component for cuticle synthesis. Once the cuticle is achieved, DOPA reaches high amounts in insects, which triggers apoptosis and autophagy activation. In another instance, the immune response is effective to reduce the level of an endosymbiont or kill an endosymbiont in comparison to a host organism to which the small molecule has not been administered. In some instances, the small molecule is effective to disrupt or decrease bacteriocyte function in comparison to a host organism to which the modulating agent  
30 has not been administered. For example, molecules that block transport of essential amino acid precursors inside the bacteriocyte also disrupt the production of essential amino acids, e.g., arginine. This alteration ultimately results in death of the endosymbiont, and, eventually, death of the host. Other examples of modulating agents that can be used to stimulate a host's immune system and thereby reduce the levels of endosymbionts resident in the host include lipopolysaccharides, rapamycin, and  $\beta$ -  
35 glucan.

In some instances, the small molecule decreases or increases gene expression of the resident microorganism by binding to non-coding RNA region. For example, the small molecule may be a riboswitch inhibitor, such as ribocil, that binds to a 'riboswitch' regulatory domain in a non-coding region of the messenger RNA that encodes a synthase enzyme involved in riboflavin synthesis, therefore inhibiting  
40 this pathway. In another instance, the small molecule is effective to increase or decrease gene

expression that results in the killing of an endosymbiont. In some instances, the small molecule is effective to disrupt bacteriocyte function.

In some instances, the small molecule alters a host's homeostasis. For example, the small molecule may be an eicosanoid molecule, such as prostaglandin, that activates a fever response to infection as well as in protein exocytosis in salivary glands. Aside from ongoing actions in homeostasis, certain eicosanoid actions occur at crucial points in insect life histories, such as during an infectious challenge and important events in reproduction. Eicosanoids mediate cellular defense reactions in insects. In one example, inhibition of prostaglandin synthesis severely impairs the insects' ability to clear bacteria from the hemolymph. In another instance, the small molecule is effective to increase or decrease an immune response in the host to an endosymbiont or increase or decrease a fever response in the host that results in the killing of an endosymbiont. In some instances, the small molecule is effective to disrupt bacteriocyte function.

As illustrated by Example 14, small molecules (e.g., prostaglandin) can be used as modulating agents that target host pathways that, in turn, alter the activity, levels, or metabolism of endosymbiotic bacteria in the host and thereby modulate (e.g., decrease) the fitness of the host.

### *iii. Nucleic Acids*

Numerous nucleic acids are useful in the compositions and methods described herein. The compositions disclosed herein may include any number or type (e.g., classes) of nucleic acids (e.g., DNA molecule or RNA molecule, e.g., mRNA, guide RNA (gRNA), or inhibitory RNA molecule (e.g., siRNA, shRNA, or miRNA), or a hybrid DNA-RNA molecule), such as at least about 1 class or variant of a nucleic acid, 2, 3, 4, 5, 10, 15, 20, or more classes or variants of nucleic acids. A suitable concentration of each nucleic acid in the composition depends on factors such as efficacy, stability of the nucleic acid, number of distinct nucleic acids, the formulation, and methods of application of the composition. In some instances, wherein the composition includes at least two types of nucleic acid, the concentration of each type of nucleic acid may be the same or different.

A modulating agent including a nucleic acid as described herein can be contacted with the target host in an amount and for a time sufficient to: (a) reach a target level (e.g., a predetermined or threshold level) of nucleic acid concentration inside a target host; (b) reach a target level (e.g., a predetermined or threshold level) of nucleic acid concentration inside a target host gut; (c) reach a target level (e.g., a predetermined or threshold level) of nucleic acid concentration inside a target host bacteriocyte; (d) modulate the level, or an activity, of one or more microorganism (e.g., endosymbiont) in the target host; or/and (e) modulate fitness of the target host.

The nucleic acid modulating agents discussed hereinafter, including nucleic acids encoding polypeptides, synthetic RNA, inhibitory RNA, and gene editing systems, can be used to alter pathways in the host that mediate interactions between the host and microorganisms resident in the host as indicated in the sections for increasing or decreasing the fitness of hosts (e.g., aphids).

#### *(a) Nucleic acids encoding a polypeptide*

In some instances, a composition includes a nucleic acid encoding any one of the polypeptides described herein. Nucleic acids encoding a polypeptide may have a length from about 10 to about

50,000 nucleotides (nts), about 25 to about 100 nts, about 50 to about 150 nts, about 100 to about 200 nts, about 150 to about 250 nts, about 200 to about 300 nts, about 250 to about 350 nts, about 300 to about 500 nts, about 10 to about 1000 nts, about 50 to about 1000 nts, about 100 to about 1000 nts, about 1000 to about 2000 nts, about 2000 to about 3000 nts, about 3000 to about 4000 nts, about 4000 to about 5000 nts, about 5000 to about 6000 nts, about 6000 to about 7000 nts, about 7000 to about 8000 nts, about 8000 to about 9000 nts, about 9000 to about 10,000 nts, about 10,000 to about 15,000 nts, about 10,000 to about 20,000 nts, about 10,000 to about 25,000 nts, about 10,000 to about 30,000 nts, about 10,000 to about 40,000 nts, about 10,000 to about 45,000 nts, about 10,000 to about 50,000 nts, or any range therebetween.

The modulating agent may also include functionally active variants of the nucleic acids described herein. In some instances, the variant of the nucleic acids has at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity, e.g., over a specified region or over the entire sequence, to a sequence of a nucleic acids described herein. In some instances, the invention includes a functionally active polypeptide encoded by a nucleic acid variant as described herein. In some instances, the functionally active polypeptide encoded by the nucleic acid variant has at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity, e.g., over a specified region or over the entire amino acid sequence, to a sequence of a polypeptide described herein or the naturally derived polypeptide sequence.

Some methods for expressing a nucleic acid encoding a protein may involve expression in cells, including host cells (e.g., insect cells, mollusk cells, or nematode cells), yeast, bacteria, or other cells under the control of appropriate promoters. Expression vectors may include nontranscribed elements, such as an origin of replication, a suitable promoter and enhancer, and other 5' or 3' flanking nontranscribed sequences, and 5' or 3' untranslated sequences such as necessary ribosome binding sites, a polyadenylation site, splice donor and acceptor sites, and termination sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the other genetic elements required for expression of a heterologous DNA sequence. Appropriate cloning and expression vectors for use with bacterial, fungal, yeast, and mammalian cellular hosts are described in Green et al., *Molecular Cloning: A Laboratory Manual*, Fourth Edition, Cold Spring Harbor Laboratory Press, 2012.

Genetic modification using recombinant methods is generally known in the art. A nucleic acid sequence coding for a desired gene can be obtained using recombinant methods known in the art, such as, for example by screening libraries from cells expressing the gene, by deriving the gene from a vector known to include the same, or by isolating directly from cells and tissues containing the same, using standard techniques. Alternatively, a gene of interest can be produced synthetically, rather than cloned.

Expression of natural or synthetic nucleic acids is typically achieved by operably linking a nucleic acid encoding the gene of interest to a promoter, and incorporating the construct into an expression vector. Expression vectors can be suitable for replication and expression in bacteria. Expression vectors can also be suitable for replication and integration in eukaryotes. Typical cloning vectors contain

transcription and translation terminators, initiation sequences, and promoters useful for expression of the desired nucleic acid sequence.

Additional promoter elements, e.g., enhancers, regulate the frequency of transcriptional initiation. Typically, these are located in the region 30-110 basepairs (bp) upstream of the start site, although a number of promoters have recently been shown to contain functional elements downstream of the start site as well. The spacing between promoter elements frequently is flexible, so that promoter function is preserved when elements are inverted or moved relative to one another. In the thymidine kinase (tk) promoter, the spacing between promoter elements can be increased to 50 bp apart before activity begins to decline. Depending on the promoter, it appears that individual elements can function either cooperatively or independently to activate transcription.

One example of a suitable promoter is the immediate early cytomegalovirus (CMV) promoter sequence. This promoter sequence is a strong constitutive promoter sequence capable of driving high levels of expression of any polynucleotide sequence operatively linked thereto. Another example of a suitable promoter is Elongation Growth Factor-1 $\alpha$  (EF-1 $\alpha$ ). However, other constitutive promoter sequences may also be used, including, but not limited to the simian virus 40 (SV40) early promoter, mouse mammary tumor virus (MMTV), human immunodeficiency virus (HIV) long terminal repeat (LTR) promoter, MoMuLV promoter, an avian leukemia virus promoter, an Epstein-Barr virus immediate early promoter, a Rous sarcoma virus promoter, as well as human gene promoters such as, but not limited to, the actin promoter, the myosin promoter, the hemoglobin promoter, and the creatine kinase promoter.

Alternatively, the promoter may be an inducible promoter. The use of an inducible promoter provides a molecular switch capable of turning on expression of the polynucleotide sequence which it is operatively linked when such expression is desired, or turning off the expression when expression is not desired. Examples of inducible promoters include, but are not limited to a metallothionine promoter, a glucocorticoid promoter, a progesterone promoter, and a tetracycline promoter.

The expression vector to be introduced can also contain either a selectable marker gene or a reporter gene or both to facilitate identification and selection of expressing cells from the population of cells sought to be transfected or infected through viral vectors. In other aspects, the selectable marker may be carried on a separate piece of DNA and used in a co-transfection procedure. Both selectable markers and reporter genes may be flanked with appropriate regulatory sequences to enable expression in the host cells. Useful selectable markers include, for example, antibiotic-resistance genes, such as neo and the like.

Reporter genes may be used for identifying potentially transformed cells and for evaluating the functionality of regulatory sequences. In general, a reporter gene is a gene that is not present in or expressed by the recipient source and that encodes a polypeptide whose expression is manifested by some easily detectable property, e.g., enzymatic activity. Expression of the reporter gene is assayed at a suitable time after the DNA has been introduced into the recipient cells. Suitable reporter genes may include genes encoding luciferase, beta-galactosidase, chloramphenicol acetyl transferase, secreted alkaline phosphatase, or the green fluorescent protein gene (e.g., Ui-Tei et al., *FEBS Letters* 479:79-82, 2000). Suitable expression systems are well known and may be prepared using known techniques or obtained commercially. In general, the construct with the minimal 5' flanking region showing the highest

level of expression of reporter gene is identified as the promoter. Such promoter regions may be linked to a reporter gene and used to evaluate agents for the ability to modulate promoter-driven transcription.

In some instances, an organism may be genetically modified to alter expression of one or more proteins. Expression of the one or more proteins may be modified for a specific time, e.g., development or differentiation state of the organism. In one instances, the invention includes a composition to alter expression of one or more proteins, e.g., proteins that affect activity, structure, or function. Expression of the one or more proteins may be restricted to a specific location(s) or widespread throughout the organism.

#### (b) *Synthetic mRNA*

The modulating agent may include an mRNA molecule, e.g., a synthetic mRNA molecule encoding a polypeptide. In some instances, the mRNA molecule increases the level (e.g., protein and/or mRNA level) and/or activity of an agent, e.g., a positive regulator of function, e.g., a gene or gene product listed in Table 7, Table 8, or Table 9. In some instances, the mRNA molecule encodes a polypeptide agent or a fragment thereof. For example, the mRNA molecule may encode a polypeptide having at least 50% (e.g., at least 50%, 60%, 70%, 80%, 90%, 95%, 97%, 99%, or greater) identity to the amino acid sequence of an agent listed in Table 7, Table 8, or Table 9 all with reference to accession number provided. In other examples, the mRNA molecule has at least 50% (e.g., at least 50%, 60%, 70%, 80%, 90%, 95%, 97%, 99%, or greater) identity to the nucleic acid sequence encoding an agent listed in Table 7, Table 8, or Table 9. In some instances, the mRNA molecule encodes an amino acid sequence differing by no more than 30 (e.g., no more than 30, 20, 10, 5, 4, 3, 2, or 1) amino acids to the amino acid sequence of an agent listed in Table 7, Table 8, or Table 9 all with reference to accession number provided. In some instances, the mRNA molecule includes a sequence encoding a fragment of a gene or gene product listed in Table 7, Table 8, or Table 9 all with reference to accession number provided. For example, the fragment includes 10-20, 20-40, 40-60, 60-80, 80-100, 100-120, 120-140, 140-160, 160-180, 180-200, 200-250, 250-300, 300-400, 400-500, 500-600, or more amino acids in length. In some instances, the fragment is a functional fragment, e.g., having at least 20%, e.g., at least 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or greater, of an activity of a full length gene or gene product listed in Table 7, Table 8, or Table 9 all with reference to accession numbers provided. In some instances, the mRNA molecule increases the level and/or activity of or encodes an agent (or fragment thereof).

An exemplary mRNA molecule includes an RNA encoding any polypeptide selected from Table 7, Table 8, or Table 9.

The synthetic mRNA molecule can be modified, e.g., chemically. The mRNA molecule can be chemically synthesized or transcribed *in vitro*. The mRNA molecule can be disposed on a plasmid, e.g., a viral vector, bacterial vector, or eukaryotic expression vector. In some examples, the mRNA molecule can be delivered to cells by transfection, electroporation, or transduction (e.g., adenoviral or lentiviral transduction).

In some instances, the modified RNA agent of interest described herein has modified nucleosides or nucleotides. Such modifications are known and are described, e.g., in WO 2012/019168. Additional

modifications are described, e.g., in WO 2015/038892; WO 2015/038892; WO 2015/089511; WO 2015/196130; WO 2015/196118 and WO 2015/196128 A2.

In some instances, the modified RNA encoding a polypeptide of interest described herein has one or more terminal modification, e.g., a 5' cap structure and/or a poly-A tail (e.g., of between 100-200 nucleotides in length). The 5' cap structure may be selected from the group consisting of CapO, CapI, ARCA, inosine, NI-methyl-guanosine, 2'fluoro- guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, and 2-azido- guanosine. In some cases, the modified RNAs also contain a 5' UTR including at least one Kozak sequence, and a 3' UTR. Such modifications are known and are described, e.g., in WO 2012/135805 and WO 2013/052523. Additional terminal modifications are described, e.g., in WO 2014/164253 and WO 2016/011306, WO 2012/045075, and WO 2014/093924.

Chimeric enzymes for synthesizing capped RNA molecules (e.g., modified mRNA) which may include at least one chemical modification are described in WO 2014/028429.

In some instances, a modified mRNA may be cyclized, or concatemerized, to generate a translation competent molecule to assist interactions between poly-A binding proteins and 5'-end binding proteins. The mechanism of cyclization or concatemerization may occur through at least 3 different routes: 1) chemical, 2) enzymatic, and 3) ribozyme catalyzed. The newly formed 5'-/3'- linkage may be intramolecular or intermolecular. Such modifications are described, e.g., in WO 2013/151736.

Methods of making and purifying modified RNAs are known and disclosed in the art. For example, modified RNAs are made using only in vitro transcription (IVT) enzymatic synthesis. Methods of making IVT polynucleotides are known in the art and are described in WO 2013/151666, WO 2013/151668, WO 2013/151663, WO 2013/151669, WO 2013/151670, WO 2013/151664, WO 2013/151665, WO 2013/151671, WO 2013/151672, WO 2013/151667 and WO 2013/151736.S Methods of purification include purifying an RNA transcript including a polyA tail by contacting the sample with a surface linked to a plurality of thymidines or derivatives thereof and/or a plurality of uracils or derivatives thereof (polyT/U) under conditions such that the RNA transcript binds to the surface and eluting the purified RNA transcript from the surface (WO 2014/152031); using ion (e.g., anion) exchange chromatography that allows for separation of longer RNAs up to 10,000 nucleotides in length via a scalable method (WO 2014/144767); and subjecting a modified mRNA sample to DNase treatment (WO 2014/152030).

Formulations of modified RNAs are known and are described, e.g., in WO 2013/090648. For example, the formulation may be, but is not limited to, nanoparticles, poly(lactic-co-glycolic acid)(PLGA) microspheres, lipidoids, lipoplex, liposome, polymers, carbohydrates (including simple sugars), cationic lipids, fibrin gel, fibrin hydrogel, fibrin glue, fibrin sealant, fibrinogen, thrombin, rapidly eliminated lipid nanoparticles (reLNPs) and combinations thereof.

Modified RNAs encoding polypeptides in the fields of human disease, antibodies, viruses, and a variety of in vivo settings are known and are disclosed in for example, Table 6 of International Publication Nos. WO 2013/151666, WO 2013/151668, WO 2013/151663, WO 2013/151669, WO 2013/151670, WO 2013/151664, WO 2013/151665, WO 2013/151736; Tables 6 and 7 International Publication No. WO 2013/151672; Tables 6, 178 and 179 of International Publication No. WO 2013/151671; Tables 6, 185 and 186 of International Publication No WO 2013/151667. Any of the foregoing may be synthesized as

an IVT polynucleotide, chimeric polynucleotide or a circular polynucleotide, and each may include one or more modified nucleotides or terminal modifications.

*(c) Inhibitory RNA*

In some instances, the modulating agent includes an inhibitory RNA molecule, e.g., that acts via the RNA interference (RNAi) pathway. For example, an inhibitory RNA molecule may include a short interfering RNA, short hairpin RNA, and/or a microRNA that targets host pathways (e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways, e.g., proteins or genes encoding proteins listed in Table 8 or Table 9, all with reference to accession number provided) in the host invertebrate (e.g., insect, mollusk, or nematode) and/or pathways in the resident microorganisms (e.g., proteins or genes encoding proteins listed in Table 7, all with reference to accession number provided). Certain RNA molecules can inhibit gene expression through the biological process of RNA interference (RNAi). RNAi molecules include RNA or RNA-like structures typically containing 15-50 base pairs (such as about 18-25 base pairs) and having a nucleobase sequence identical (complementary) or nearly identical (substantially complementary) to a coding sequence in an expressed target gene within the cell. RNAi molecules include, but are not limited to: short interfering RNAs (siRNAs), double-strand RNAs (dsRNA), short hairpin RNAs (shRNA), meroduplexes, dicer substrates, and multivalent RNA interference (U.S. Pat. Nos. 8,084,599 8,349,809, 8,513,207 and 9,200,276). A shRNA is a RNA molecule comprising a hairpin turn that decreases expression of target genes via RNAi. shRNAs can be delivered to cells in the form of plasmids, e.g., viral or bacterial vectors, e.g., by transfection, electroporation, or transduction). A microRNA is a non-coding RNA molecule that typically has a length of about 22 nucleotides. MiRNAs bind to target sites on mRNA molecules and silence the mRNA, e.g., by causing cleavage of the mRNA, destabilization of the mRNA, or inhibition of translation of the mRNA. In some instances, the inhibitory RNA molecule decreases the level and/or activity of a negative regulator of function. In other embodiments, the inhibitor RNA molecule decreases the level and/or activity of an inhibitor of a positive regulator of function.

RNAi molecules include a sequence substantially complementary, or fully complementary, to all or a fragment of a target gene. RNAi molecules may complement sequences at the boundary between introns and exons to prevent the maturation of newly-generated nuclear RNA transcripts of specific genes into mRNA for transcription. RNAi molecules complementary to specific genes can hybridize with the mRNA for a target gene and prevent its translation. The antisense molecule can be DNA, RNA, or a derivative or hybrid thereof. Examples of such derivative molecules include, but are not limited to, peptide nucleic acid (PNA) and phosphorothioate-based molecules such as deoxyribonucleic guanidine (DNG) or ribonucleic guanidine (RNG).

RNAi molecules can be provided as "ready-to-use" RNA synthesized in vitro or as an antisense gene transfected into cells which will yield RNAi molecules upon transcription. Hybridization with mRNA results in degradation of the hybridized molecule by RNase H and/or inhibition of the formation of translation complexes. Both result in a failure to produce the product of the original gene.

The length of the RNAi molecule that hybridizes to the transcript of interest may be around 10 nucleotides, between about 15 or 30 nucleotides, or about 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26,

27, 28, 29, 30 or more nucleotides. The degree of identity of the antisense sequence to the targeted transcript may be at least 75%, at least 80%, at least 85%, at least 90%, or at least 95.

RNAi molecules may also include overhangs, i.e., typically unpaired, overhanging nucleotides which are not directly involved in the double helical structure normally formed by the core sequences of the herein defined pair of sense strand and antisense strand. RNAi molecules may contain 3' and/or 5' overhangs of about 1-5 bases independently on each of the sense strands and antisense strands. In some instances, both the sense strand and the antisense strand contain 3' and 5' overhangs. In some instances, one or more of the 3' overhang nucleotides of one strand base pairs with one or more 5' overhang nucleotides of the other strand. In other instances, the one or more of the 3' overhang nucleotides of one strand base do not pair with the one or more 5' overhang nucleotides of the other strand. The sense and antisense strands of an RNAi molecule may or may not contain the same number of nucleotide bases. The antisense and sense strands may form a duplex wherein the 5' end only has a blunt end, the 3' end only has a blunt end, both the 5' and 3' ends are blunt ended, or neither the 5' end nor the 3' end are blunt ended. In another instance, one or more of the nucleotides in the overhang contains a thiophosphate, phosphorothioate, deoxynucleotide inverted (3' to 3' linked) nucleotide or is a modified ribonucleotide or deoxynucleotide.

Small interfering RNA (siRNA) molecules include a nucleotide sequence that is identical to about 15 to about 25 contiguous nucleotides of the target mRNA. In some instances, the siRNA sequence commences with the dinucleotide AA, includes a GC-content of about 30-70% (about 30-60%, about 40-60%, or about 45%-55%), and does not have a high percentage identity to any nucleotide sequence other than the target in the genome in which it is to be introduced, for example as determined by standard BLAST search.

siRNAs and shRNAs resemble intermediates in the processing pathway of the endogenous microRNA (miRNA) genes (Bartel, *Cell* 116:281-297, 2004). In some instances, siRNAs can function as miRNAs and vice versa (Zeng et al., *Mol. Cell* 9:1327-1333, 2002; Doench et al., *Genes Dev.* 17:438-442, 2003). Exogenous siRNAs downregulate mRNAs with seed complementarity to the siRNA (Birmingham et al., *Nat. Methods* 3:199-204, 2006). Multiple target sites within a 3' UTR give stronger downregulation (Doench et al., *Genes Dev.* 17:438-442, 2003).

Known effective siRNA sequences and cognate binding sites are also well represented in the relevant literature. RNAi molecules are readily designed and produced by technologies known in the art. In addition, there are computational tools that increase the chance of finding effective and specific sequence motifs (Pei et al., *Nat. Methods* 3(9):670-676, 2006; Reynolds et al., *Nat. Biotechnol.* 22(3):326-330, 2004; Khvorova et al., *Nat. Struct. Biol.* 10(9):708-712, 2003; Schwarz et al., *Cell* 115(2):199-208, 2003; Ui-Tei et al., *Nucleic Acids Res.* 32(3):936-948, 2004; Heale et al., *Nucleic Acids Res.* 33(3):e30, 2005; Chalk et al., *Biochem. Biophys. Res. Commun.* 319(1):264-274, 2004; and Amarzguoui et al., *Biochem. Biophys. Res. Commun.* 316(4):1050-1058, 2004).

The RNAi molecule modulates expression of RNA encoded by a gene. Because multiple genes can share some degree of sequence homology with each other, in some instances, the RNAi molecule can be designed to target a class of genes with sufficient sequence homology. In some instances, the RNAi molecule can contain a sequence that has complementarity to sequences that are shared amongst

different gene targets or are unique for a specific gene target. In some instances, the RNAi molecule can be designed to target conserved regions of an RNA sequence having homology between several genes thereby targeting several genes in a gene family (e.g., different gene isoforms, splice variants, mutant genes, etc.). In some instances, the RNAi molecule can be designed to target a sequence that is unique to a specific RNA sequence of a single gene.

In some instances, the inhibitory RNA molecule decreases the level and/or activity of a host component (e.g., component in pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways) and/or microbial component, including genes encoding proteins listed in Table 7, Table 8, or Table 9, all with reference to accession number provided. In some instances, the inhibitory RNA molecule inhibits expression of a host component (e.g., component in pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways) or microbial component, e.g., genes encoding proteins listed in Table 7, Table 8, or Table 9 (e.g., inhibits translation to protein). In other instances, the inhibitor RNA molecule increases degradation of a host component (e.g., component in pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways) and/or microbial component, e.g., genes encoding proteins listed in Table 7, Table 8, or Table 9 and/or decreases the stability (i.e., half-life) of the pathway component. The inhibitory RNA molecule can be chemically synthesized or transcribed in vitro.

In some instances, the compositions described herein include an RNAi, e.g., siRNA, to regulate (e.g., inhibit) expression of a gene encoding any of the components described herein that regulate a host's immune system. In some instances, a composition includes an RNAi, e.g., siRNA, to inhibit expression of any one of the genes described herein that regulate (e.g., inhibit) the development or function of a bacteriocyte in the host. In some instances, regulation of the host immune system leads to a reduction or killing of an endosymbiotic microorganism in the host, and in turn, reduces the fitness of the host. In some instances, regulation of bacteriocyte development and/or function leads to a reduction or killing of an endosymbiotic microorganism in the host, and in turn, reduces the fitness of the host. In some instances, the RNAi (e.g., siRNA) may inhibit expression of *Ubx* to disrupt symbiont localization of bacteriocytes. In some instances, the RNAi (e.g., siRNA) inhibits expression of *abd-A* and *Antp* to disrupt bacteriome integrity and positioning.

In some instances, one or more RNAi molecules target any gene encoding any protein described herein, e.g., see Table 7, Table 8, or Table 9. In some instances, one or more RNAi molecules target a bacterial gene as described herein. In some instances, one or more RNAi molecules target an endosymbiont gene as described herein. In some instances, one or more RNAi molecules target a bacteriocyte gene as described herein. In some instances, one or more RNAi molecules target a host gene as described herein. In some instances, one or more RNAi molecules target an immune system gene in a host as described herein.

An inhibitory RNA molecule can be modified, e.g., to contain modified nucleotides, e.g., 2'-fluoro, 2'-o-methyl, 2'-deoxy, unlocked nucleic acid, 2'-hydroxy, phosphorothioate, 2'-thiouridine, 4'-thiouridine, 2'-deoxyuridine. Without being bound by theory, it is believed that such modifications can increase nuclease resistance and/or serum stability, or decrease immunogenicity.

In some instances, the RNAi molecule is linked to a delivery polymer via a physiologically labile bond or linker. The physiologically labile linker is selected such that it undergoes a chemical transformation (e.g., cleavage) when present in certain physiological conditions, (e.g., disulfide bond cleaved in the reducing environment of the cell cytoplasm). Release of the molecule from the polymer, by  
 5 cleavage of the physiologically labile linkage, facilitates interaction of the molecule with the appropriate cellular components for activity.

The RNAi molecule-polymer conjugate may be formed by covalently linking the molecule to the polymer. The polymer is polymerized or modified such that it contains a reactive group A. The RNAi molecule is also polymerized or modified such that it contains a reactive group B. Reactive groups A and  
 10 B are chosen such that they can be linked via a reversible covalent linkage using methods known in the art.

Conjugation of the RNAi molecule to the polymer can be performed in the presence of an excess of polymer. Because the RNAi molecule and the polymer may be of opposite charge during conjugation, the presence of excess polymer can reduce or eliminate aggregation of the conjugate. Alternatively, an  
 15 excess of a carrier polymer, such as a polycation, can be used. The excess polymer can be removed from the conjugated polymer prior to administration of the conjugate. Alternatively, the excess polymer can be co-administered with the conjugate.

Injection of double-stranded RNA (dsRNA) into mother insects efficiently suppresses their offspring's gene expression during embryogenesis, see for example, Khila et al., *PLoS Genet.*  
 20 5(7):e1000583, 2009; and Liu et al., *Development* 131(7):1515-1527, 2004. Matsuura et al. (*PNAS* 112(30):9376-9381, 2015) has shown that suppression of *Ubx* eliminates bacteriocytes and the symbiont localization of bacteriocytes.

The making and use of inhibitory agents based on non-coding RNA such as ribozymes, RNase P, siRNAs, and miRNAs are also known in the art, for example, as described in *Sioud, RNA  
 25 Therapeutics: Function, Design, and Delivery (Methods in Molecular Biology). Humana Press (2010).*

Other examples of nucleic acid modulating agents that can be used herein include dsRNAs having at least 50% (e.g., at least 50%, 60%, 70%, 80%, 90%, 95%, 97%, 99%, or greater) identity to the sequence of any one of SEQ ID NOs: 148-150.

As illustrated by Examples 1-4 and 7-9, inhibitory RNA (e.g., dsRNA or PNA) can be used as a  
 30 modulating agent that targets a host pathway (e.g., an insect, e.g., an aphid) that, in turn, alters the activity, levels, or metabolism of endosymbiotic bacteria, such as a *Buchnera* spp., resident in the host and thereby modulates (e.g., decreases) the fitness of the host.

#### (d) Gene Editing

The modulating agents described herein may include a component of a gene editing system. For example, the agent may introduce an alteration (e.g., insertion, deletion (e.g., knockout), translocation, inversion, single point mutation, or other mutation) in a gene related to the immune system or bacteriocyte of a host invertebrate (e.g., insect, mollusk, or nematode) or in a gene in a microorganism resident in the host invertebrate, e.g., an enzyme or receptor gene described in Table 7, Table 8, or Table  
 40 9 all with reference to accession number provided. Exemplary gene editing systems include the zinc

finger nucleases (ZFNs), Transcription Activator-Like Effector-based Nucleases (TALEN), and the clustered regulatory interspaced short palindromic repeat (CRISPR) system. ZFNs, TALENs, and CRISPR-based methods are described, e.g., in Gaj et al., *Trends Biotechnol.* 31(7):397-405, 2013.

In a typical CRISPR/Cas system, an endonuclease is directed to a target nucleotide sequence (e.g., a site in the genome that is to be sequence-edited) by sequence-specific, non-coding “guide RNAs” that target single- or double-stranded DNA sequences. Three classes (I-III) of CRISPR systems have been identified. The class II CRISPR systems use a single Cas endonuclease (rather than multiple Cas proteins). One class II CRISPR system includes a type II Cas endonuclease such as Cas9, a CRISPR RNA (“crRNA”), and a trans-activating crRNA (“tracrRNA”). The crRNA contains a “guide RNA,” i.e., typically an about 20-nucleotide RNA sequence that corresponds to a target DNA sequence. The crRNA also contains a region that binds to the tracrRNA to form a partially double-stranded structure which is cleaved by RNase III, resulting in a crRNA/tracrRNA hybrid. The RNAs serve as guides to direct Cas proteins to silence specific DNA/RNA sequences, depending on the spacer sequence. See, e.g., Horvath et al., *Science* 327:167-170, 2010; Makarova et al., *Biology Direct* 1:7, 2006; Pennisi, *Science* 341:833-836, 2013. The target DNA sequence must generally be adjacent to a “protospacer adjacent motif” (“PAM”) that is specific for a given Cas endonuclease; however, PAM sequences appear throughout a given genome. CRISPR endonucleases identified from various prokaryotic species have unique PAM sequence requirements; examples of PAM sequences include 5'-NGG (SEQ ID NO: 78) (*Streptococcus pyogenes*), 5'-NNAGAA (SEQ ID NO: 79) (*Streptococcus thermophilus* CRISPR1), 5'-NGGNG (SEQ ID NO: 80) (*Streptococcus thermophilus* CRISPR3), and 5'-NNNGATT (SEQ ID NO: 81) (*Neisseria meningitidis*). Some endonucleases, e.g., Cas9 endonucleases, are associated with G-rich PAM sites, e.g., 5'-NGG (SEQ ID NO: 78), and perform blunt-end cleaving of the target DNA at a location 3 nucleotides upstream from (5' from) the PAM site. Another class II CRISPR system includes the type V endonuclease Cpf1, which is smaller than Cas9; examples include AsCpf1 (from *Acidaminococcus* sp.) and LbCpf1 (from *Lachnospiraceae* sp.). Cpf1-associated CRISPR arrays are processed into mature crRNAs without the requirement of a tracrRNA; in other words a Cpf1 system requires only the Cpf1 nuclease and a crRNA to cleave the target DNA sequence. Cpf1 endonucleases, are associated with T-rich PAM sites, e.g., 5'-TTN. Cpf1 can also recognize a 5'-CTA PAM motif. Cpf1 cleaves the target DNA by introducing an offset or staggered double-strand break with a 4- or 5-nucleotide 5' overhang, for example, cleaving a target DNA with a 5-nucleotide offset or staggered cut located 18 nucleotides downstream from (3' from) from the PAM site on the coding strand and 23 nucleotides downstream from the PAM site on the complimentary strand; the 5-nucleotide overhang that results from such offset cleavage allows more precise genome editing by DNA insertion by homologous recombination than by insertion at blunt-end cleaved DNA. See, e.g., Zetsche et al., *Cell* 163:759-771, 2015.

For the purposes of gene editing, CRISPR arrays can be designed to contain one or multiple guide RNA sequences corresponding to a desired target DNA sequence; see, for example, Cong et al., *Science* 339:819-823, 2013; Ran et al., *Nature Protocols* 8:2281-2308, 2013. At least about 16 or 17 nucleotides of gRNA sequence are required by Cas9 for DNA cleavage to occur; for Cpf1 at least about 16 nucleotides of gRNA sequence is needed to achieve detectable DNA cleavage. In practice, guide RNA sequences are generally designed to have a length of between 17-24 nucleotides (e.g., 19, 20, or

21 nucleotides) and complementarity to the targeted gene or nucleic acid sequence. Custom gRNA generators and algorithms are available commercially for use in the design of effective guide RNAs. Gene editing has also been achieved using a chimeric “single guide RNA” (“sgRNA”), an engineered (synthetic) single RNA molecule that mimics a naturally occurring crRNA-tracrRNA complex and contains both a tracrRNA (for binding the nuclease) and at least one crRNA (to guide the nuclease to the sequence targeted for editing). Chemically modified sgRNAs have also been demonstrated to be effective in genome editing; see, for example, Hendel et al., *Nature Biotechnol.* 985-991, 2015.

Whereas wild-type Cas9 generates double-strand breaks (DSBs) at specific DNA sequences targeted by a gRNA, a number of CRISPR endonucleases having modified functionalities are available, for example: a “nickase” version of Cas9 generates only a single-strand break; a catalytically inactive Cas9 (“dCas9”) does not cut the target DNA but interferes with transcription by steric hindrance. dCas9 can further be fused with an effector to repress (CRISPRi) or activate (CRISPRa) expression of a target gene. For example, Cas9 can be fused to a transcriptional repressor (e.g., a KRAB domain) or a transcriptional activator (e.g., a dCas9–VP64 fusion). A catalytically inactive Cas9 (dCas9) fused to FokI nuclease (“dCas9-FokI”) can be used to generate DSBs at target sequences homologous to two gRNAs. See, e.g., the numerous CRISPR/Cas9 plasmids disclosed in and publicly available from the Addgene repository (Addgene, 75 Sidney St., Suite 550A, Cambridge, MA 02139; [addgene.org/crispr/](http://addgene.org/crispr/)). A “double nickase” Cas9 that introduces two separate double-strand breaks, each directed by a separate guide RNA, is described as achieving more accurate genome editing by Ran et al., *Cell* 154:1380-1389, 2013.

CRISPR technology for editing the genes of eukaryotes is disclosed in US Patent Application Publications US 2016/0138008 A1 and US 2015/0344912 A1, and in US Patents 8,697,359, 8,771,945, 8,945,839, 8,999,641, 8,993,233, 8,895,308, 8,865,406, 8,889,418, 8,871,445, 8,889,356, 8,932,814, 8,795,965, and 8,906,616. Cpf1 endonuclease and corresponding guide RNAs and PAM sites are disclosed in US Patent Application Publication 2016/0208243 A1.

In some instances, the desired genome modification involves homologous recombination, wherein one or more double-stranded DNA breaks in the target nucleotide sequence is generated by the RNA-guided nuclease and guide RNA(s), followed by repair of the break(s) using a homologous recombination mechanism (“homology-directed repair”). In such instances, a donor template that encodes the desired nucleotide sequence to be inserted or knocked-in at the double-stranded break is provided to the cell or subject; examples of suitable templates include single-stranded DNA templates and double-stranded DNA templates (e.g., linked to the polypeptide described herein). In general, a donor template encoding a nucleotide change over a region of less than about 50 nucleotides is provided in the form of single-stranded DNA; larger donor templates (e.g., more than 100 nucleotides) are often provided as double-stranded DNA plasmids. In some instances, the donor template is provided to the cell or subject in a quantity that is sufficient to achieve the desired homology-directed repair but that does not persist in the cell or subject after a given period of time (e.g., after one or more cell division cycles). In some instances, a donor template has a core nucleotide sequence that differs from the target nucleotide sequence (e.g., a homologous endogenous genomic region) by at least 1, at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, or more nucleotides. This core sequence is flanked by “homology arms” or regions of high sequence identity with the targeted nucleotide sequence; in some instances, the

regions of high identity include at least 10, at least 50, at least 100, at least 150, at least 200, at least 300, at least 400, at least 500, at least 600, at least 750, or at least 1000 nucleotides on each side of the core sequence. In some instances where the donor template is in the form of a single-stranded DNA, the core sequence is flanked by homology arms including at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, or at least 100 nucleotides on each side of the core sequence. In instances, where the donor template is in the form of a double-stranded DNA, the core sequence is flanked by homology arms including at least 500, at least 600, at least 700, at least 800, at least 900, or at least 1000 nucleotides on each side of the core sequence. In one instance, two separate double-strand breaks are introduced into the cell or subject's target nucleotide sequence with a "double nickase" Cas9 (see Ran et al., *Cell* 154:1380-1389, 2013), followed by delivery of the donor template.

In some instances, the composition includes a gRNA and a targeted nuclease, e.g., a Cas9, e.g., a wild type Cas9, a nickase Cas9 (e.g., Cas9 D10A), a dead Cas9 (dCas9), eSpCas9, Cpf1, C2C1, or C2C3, or a nucleic acid encoding such a nuclease. The choice of nuclease and gRNA(s) is determined by whether the targeted mutation is a deletion, substitution, or addition of nucleotides, e.g., a deletion, substitution, or addition of nucleotides to a targeted sequence. Fusions of a catalytically inactive endonuclease e.g., a dead Cas9 (dCas9, e.g., D10A; H840A) tethered with all or a portion of (e.g., biologically active portion of) an (one or more) effector domain create chimeric proteins that can be linked to the polypeptide to guide the composition to specific DNA sites by one or more RNA sequences (sgRNA) to modulate activity and/or expression of one or more target nucleic acids sequences.

In instances, the agent includes a guide RNA (gRNA) for use in a CRISPR system for gene editing. In some instances, the agent includes a zinc finger nuclease (ZFN), or a mRNA encoding a ZFN, that targets (e.g., cleaves) a nucleic acid sequence (e.g., DNA sequence) of a gene related to pathways in the host invertebrate (e.g., insect, mollusk, or nematode) and/or resident microorganisms (e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways, e.g., a gene encoding a protein listed in Table 7, Table 8, or Table 9, all with reference to accession number provided). In some instances, the agent includes a TALEN, or an mRNA encoding a TALEN, that targets (e.g., cleaves) a nucleic acid sequence (e.g., DNA sequence) in a gene related to pathways in the host invertebrate (e.g., insect, mollusk, or nematode) and/or resident microorganisms (e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways, e.g., a gene encoding a protein listed in Table 7, Table 8, or Table 9, all with reference to accession number provided).

For example, the gRNA can be used in a CRISPR system to engineer an alteration in a gene related to pathways in the host invertebrate (e.g., insect, mollusk, or nematode) and/or resident microorganisms (e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways, e.g., a gene encoding a protein listed in Table 7, Table 8, or Table 9, all with reference to accession number provided). In other examples, the ZFN and/or TALEN can be used to engineer an alteration in a gene related to pathways in the host invertebrate (e.g., insect, mollusk, or nematode) and/or resident microorganisms (e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways, e.g., a gene encoding a protein listed in Table 7, Table 8, or Table 9, all with reference to accession number provided). Exemplary alterations

include insertions, deletions (e.g., knockouts), translocations, inversions, single point mutations, or other mutations. The alteration can be introduced in the gene in a cell, e.g., in vitro, ex vivo, or in vivo. In some examples, the alteration increases the level and/or activity of a gene related to pathways in the host invertebrate (e.g., insect, mollusk, or nematode) and/or resident microorganisms (e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways, e.g., a gene encoding a protein listed in Table 7, Table 8, or Table 9, all with reference to accession number provided). In other examples, the alteration decreases the level and/or activity of (e.g., knocks down or knocks out) a gene related to pathways in the host invertebrate (e.g., insect, mollusk, or nematode) and/or resident microorganisms (e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways, e.g., a gene encoding a protein listed in Table 7, Table 8, or Table 9, all with reference to accession number provided). In yet another example, the alteration corrects a defect (e.g., a mutation causing a defect), in a gene related to pathways in the host invertebrate (e.g., insect, mollusk, or nematode) and/or resident microorganisms (e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways, e.g., a gene encoding protein listed in Table 7, Table 8, or Table 9, all with reference to accession number provided).

In some instances, the CRISPR system is used to edit (e.g., to add or delete a base pair) a target gene (e.g., a gene related to pathways in the host invertebrate (e.g., insect, mollusk, or nematode) and/or resident microorganisms, e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways, e.g., a gene encoding a protein listed in Table 7, Table 8, or Table 9, all with reference to accession number provided). In other instances, the CRISPR system is used to introduce a premature stop codon, e.g., thereby decreasing the expression of a target gene. In yet other instances, the CRISPR system is used to turn off a target gene in a reversible manner, e.g., similarly to RNA interference. In some instances, the CRISPR system is used to direct Cas to a promoter of a gene, thereby blocking an RNA polymerase sterically.

In some instances, a CRISPR system can be generated to edit a gene related to pathways in the host invertebrate (e.g., insect, mollusk, or nematode) and/or resident microorganisms (e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways, e.g., genes that encode a protein listed in Table 7, Table 8, or Table 9, all with reference to accession number provided), using technology described in, e.g., U.S. Publication No. 20140068797, Cong, *Science* 339: 819-823, 2013; Tsai, *Nature Biotechnol.* 32:6 569-576, 2014; U.S. Patent No.: 8,871,445; 8,865,406; 8,795,965; 8,771,945; and 8,697,359.

In some instances, the CRISPR interference (CRISPRi) technique can be used for transcriptional repression of specific genes, e.g., a gene encoding a host immune system or bacteriocyte component (e.g., an enzyme or receptor described herein). In CRISPRi, an engineered Cas9 protein (e.g., nuclease-null dCas9, or dCas9 fusion protein, e.g., dCas9-KRAB or dCas9-SID4X fusion) can pair with a sequence specific guide RNA (sgRNA). The Cas9-gRNA complex can block RNA polymerase, thereby interfering with transcription elongation. The complex can also block transcription initiation by interfering with transcription factor binding. The CRISPRi method is specific with minimal off-target effects and is

multiplexable, e.g., can simultaneously repress more than one gene (e.g., using multiple gRNAs). Also, the CRISPRi method permits reversible gene repression.

In some instances, CRISPR-mediated gene activation (CRISPRa) can be used for transcriptional activation, e.g., of one or more genes described herein (e.g., a gene encoding any of the proteins listed in Table 7, Table 8, or Table 9). In the CRISPRa technique, dCas9 fusion proteins recruit transcriptional activators. For example, dCas9 can be fused to polypeptides (e.g., activation domains) such as VP64 or the p65 activation domain (p65D) and used with sgRNA (e.g., a single sgRNA or multiple sgRNAs), to activate a gene or genes, e.g., endogenous gene(s) in the host or microorganism resident in the host. Multiple activators can be recruited by using multiple sgRNAs – this can increase activation efficiency. A variety of activation domains and single or multiple activation domains can be used. In addition to engineering dCas9 to recruit activators, sgRNAs can also be engineered to recruit activators. For example, RNA aptamers can be incorporated into a sgRNA to recruit proteins (e.g., activation domains) such as VP64. In some examples, the synergistic activation mediator (SAM) system can be used for transcriptional activation. In SAM, MS2 aptamers are added to the sgRNA. MS2 recruits the MS2 coat protein (MCP) fused to p65AD and heat shock factor 1 (HSF1).

The CRISPRi and CRISPRa techniques are described in greater detail, e.g., in Dominguez et al., *Nat. Rev. Mol. Cell Biol.* 17:5-15, 2016, incorporated herein by reference. In addition, dCas9-mediated epigenetic modifications and simultaneous activation and repression using CRISPR systems, as described in Dominguez et al., can be used to modulate a component of a host or microbial pathway described herein (e.g., pathways that mediate host-microbiota interactions, e.g., host immune system pathways or bacteriocyte pathways, e.g., a gene encoding a protein listed in Table 7, Table 8, or Table 9, all with reference to accession number provided).

#### *iv. Target genes and proteins*

Any of the modulating agents described herein can be used to alter (e.g., increase or decrease) gene expression, alter (e.g., increase or decrease) a target protein activity, and/or alter function in the host or a microorganism resident in the host. Proteins or genes that are involved in a variety of processes may be targeted, including any of the functional classes listed in Table 6.

**Table 6: Functional classes of target genes**

| Functional class | Effect                                                                                                                                                                                                                                         |
|------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Homeostasis      | Downregulation or knockout of these genes will disrupt inner target microorganism homeostatic balance generating in the host a cellular malfunction and therefore a decrease in their fitness.                                                 |
| Information      | Downregulation or knockout of these genes will stop or slow protein synthesis therefore affecting proliferation of the microorganism. This in turn will generate in the host a cellular malfunction and therefore a decrease in their fitness. |

|            |                                                                                                                                                                                                                                                                       |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Metabolism | Downregulation or knockout of these genes will stop the catabolism and anabolism of nutrients affecting proliferation of the microorganism. This in turn will generate in the host a cellular malfunction and therefore a decrease in their fitness.                  |
| Transport  | Downregulation or knockout of these genes will stop or decrease the transport of amino acids or their precursors therefore affecting microorganism survival. This in turn will generate in the host a cellular malfunction and therefore a decrease in their fitness. |

(a) *Target microbial genes and proteins*

Any of the modulating agents described herein can be used to alter gene expression or target proteins in a microorganism resident in the host. In some instances, the modulating agent (e.g., an antibody) directly targets a protein in a microorganism resident in the host, including any one of the proteins listed in Table 7. In other instances, the modulating agent (e.g., nucleic acid, e.g., RNAi) alters gene expression (e.g., increases or decreases gene expression) in a microorganism resident in the host, including genes that encode any of the proteins listed in Table 7, in comparison to a host organism to which the modulating agent has not been administered.

**Table 7: Target bacterial proteins**

| Functional class | Protein name                               | Sequence Accession No. | Organism                                                             |
|------------------|--------------------------------------------|------------------------|----------------------------------------------------------------------|
| Homeostasis      | chaperonin GroEL                           | NP_239860.1            | <i>Buchnera aphidicola</i> str. APS<br>( <i>Acyrtosiphon pisum</i> ) |
| Homeostasis      | DnaK protein                               | NP_239985.1            | <i>Buchnera aphidicola</i> str. APS<br>( <i>Acyrtosiphon pisum</i> ) |
| Homeostasis      | heat shock protein GrpE1                   | NP_240076.1            | <i>Buchnera aphidicola</i> str. APS<br>( <i>Acyrtosiphon pisum</i> ) |
| Homeostasis      | ATP-dependent protease ATP-binding subunit | NP_240382.1            | <i>Buchnera aphidicola</i> str. APS<br>( <i>Acyrtosiphon pisum</i> ) |
| Homeostasis      | heat shock protein GrpE2                   | NP_240015.1            | <i>Buchnera aphidicola</i> str. APS<br>( <i>Acyrtosiphon pisum</i> ) |

|             |                                            |             |                                                          |
|-------------|--------------------------------------------|-------------|----------------------------------------------------------|
| Information | methionine aminopeptidase                  | NP_240059.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | 30S ribosomal protein S1                   | NP_240132.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | DnaJ protein                               | NP_239984.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | polynucleotide phosphorylase/polyadenylase | NP_240191.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | DNA gyrase subunit B                       | NP_239852.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | tryptophanyl-tRNA synthetase               | NP_240343.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | threonyl-tRNA synthetase                   | NP_239957.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | alanyl-tRNA synthetase                     | NP_240220.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | asparaginyl-tRNA synthetase                | NP_240178.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | tRNA (guanine-N1)-methyltransferase        | NP_240213.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | 50S ribosomal protein L30                  | NP_240313.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | replicative DNA helicase                   | NP_240352.2 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |

|             |                                                  |             |                                                          |
|-------------|--------------------------------------------------|-------------|----------------------------------------------------------|
| Information | glycyl-tRNA synthetase subunit alpha             | NP_239968.2 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | A/G-specific adenine glycosylase                 | NP_240358.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | tRNA pseudouridine 55 synthase                   | NP_240193.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | methionyl-tRNA synthetase                        | NP_239942.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | lysyl-tRNA synthetase                            | NP_240385.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | glutamyl-tRNA synthetase                         | NP_240227.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | DNA polymerase I                                 | NP_240243.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | ribosomal large subunit pseudouridine synthase C | NP_240167.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | histidyl-tRNA synthetase                         | NP_240112.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | DNA polymerase III subunits gamma and tau        | NP_240292.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | tRNA modification GTPase TrmE                    | NP_239858.2 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | DNA polymerase III beta chain                    | NP_239853.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |

|             |                                                                      |             |                                                          |
|-------------|----------------------------------------------------------------------|-------------|----------------------------------------------------------|
| Information | RNA polymerase sigma factor RpoD                                     | NP_239892.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | arginyl-tRNA synthetase                                              | NP_240071.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | seryl-tRNA synthetase                                                | NP_240135.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | DNA polymerase III alpha chain                                       | NP_240067.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | aspartyl-tRNA synthetase                                             | NP_240138.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | leucyl-tRNA synthetase                                               | NP_240256.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Information | phenylalanyl-tRNA synthetase beta chain                              | NP_239962.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Metabolism  | 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase | NP_239871.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Metabolism  | bifunctional aspartokinase I/homoserine dehydrogenase I              | NP_240025.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Metabolism  | sulfate adenylate transferase subunit 1                              | NP_240235.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Metabolism  | adenylosuccinate synthetase                                          | NP_240370.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Metabolism  | phosphoadenosine phosphosulfate reductase                            | NP_240238.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |

|            |                                                                  |             |                                                          |
|------------|------------------------------------------------------------------|-------------|----------------------------------------------------------|
| Metabolism | phosphoserine aminotransferase                                   | NP_240134.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Metabolism | F0F1 ATP synthase subunit gamma                                  | NP_239849.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Transport  | cell division inhibitor MinC                                     | NP_240149.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Transport  | 23S rRNA m(2)G2445 methyltransferase                             | NP_240181.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |
| Transport  | tRNA uridine 5-carboxymethylaminomethyl modification enzyme GidA | NP_239843.1 | <i>Buchnera aphidicola str. APS (Acyrtosiphon pisum)</i> |

*(b) Target genes and proteins in hosts*

Any of the modulating agents described herein can be used to alter gene expression or target proteins in the host. In some instances, the modulating agent (e.g., an antibody) directly targets a protein in the host, including any one of the proteins listed in Table 8 or Table 9. In other instances, the modulating agent (e.g., nucleic acid, e.g., RNAi) targets a gene in the host, including genes that encode any of the proteins listed in Table 8 or Table 9. For example, the nucleic acids described herein can be used to alter (e.g., increase or decrease) gene expression in a host (e.g., genes that regulate bacteriocyte function or development (e.g., bacteriocyte regulatory peptides) or genes that regulate the immune system) including, but not limited to, any of the genes listed in Table 8 and Table 9, in comparison to a host organism to which the modulating agent has not been administered.

**Table 8: Proteins that regulate bacteriocyte function**

| Functional class | Protein name                                  | Sequence accession number | Organism                  |
|------------------|-----------------------------------------------|---------------------------|---------------------------|
| Homeostasis      | ns1 binding protein                           | ACYPI000701-PA            | <i>Acyrtosiphon pisum</i> |
| Homeostasis      | Ubx transcription factor                      | ACYPI001856-RA            | <i>Acyrtosiphon pisum</i> |
| Metabolism       | ATPase                                        | ACYPI002584-PA            | <i>Acyrtosiphon pisum</i> |
| Metabolism       | Succinic semialdehyde dehydrogenase isoform 1 | ACYPI002063-PA            | <i>Acyrtosiphon pisum</i> |
| Metabolism       | Vacuolar proton atpases isoform 1             | ACYPI007445-PA            | <i>Acyrtosiphon pisum</i> |
| Metabolism       | Membrane alanyl aminopeptidase N              | ACYPI006675-PA            | <i>Acyrtosiphon pisum</i> |

|            |                                                                                                                                   |                |                           |
|------------|-----------------------------------------------------------------------------------------------------------------------------------|----------------|---------------------------|
| Metabolism | Protease m1 zinc metalloprotease                                                                                                  | ACYPI009427-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Zinc metalloprotease                                                                                                              | ACYPI000580-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Purine biosynthesis protein 6, pur6                                                                                               | ACYPI010114-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Myo inositol monophosphatase                                                                                                      | ACYPI009018-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Phenylalanine hydroxylase                                                                                                         | ACYPI007803-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | 4-nitrophenylphosphatase isoform 1                                                                                                | ACYPI005939-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Cytosolic purine 5-nucleotidase                                                                                                   | ACYPI007730-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Amine oxidase                                                                                                                     | ACYPI006507-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Phosphoserine aminotransferase                                                                                                    | ACYPI004666-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | 4-Hydroxybutyrate CoA-transferase                                                                                                 | ACYPI001782-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Dihydropyrimidine dehydrogenase                                                                                                   | ACYPI004747-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Glycine dehydrogenase, mitochondrial                                                                                              | ACYPI005060-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Ribose-phosphate pyrophosphokinase 1, putative                                                                                    | ACYPI006288-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Phosphoserine phosphatase isoform 1                                                                                               | ACYPI000304-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Adenine phosphoribosyltransferase                                                                                                 | ACYPI003436-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Pantothenate kinase 4 (Pantothenic acid kinase 4) (hPanK4)                                                                        | ACYPI003518-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Phosphoenolpyruvate carboxykinase                                                                                                 | ACYPI001978-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Cystathionine beta-lyase, partial                                                                                                 | ACYPI000593-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Putative 5-nucleotidase, partial                                                                                                  | ACYPI002452-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Zipper CG15792-PD                                                                                                                 | ACYPI004129-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Aconitase                                                                                                                         | ACYPI008211-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Prophenoloxidase                                                                                                                  | ACYPI001367-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Glycinamide ribonucleotide synthetase-aminoimidazole ribonucleotide synthetase-glycinamide ribonucleotide transformylase, partial | ACYPI009448-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Aldehyde dehydrogenase                                                                                                            | ACYPI003925-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Metalloprotease                                                                                                                   | ACYPI008675-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Prophenoloxidase                                                                                                                  | ACYPI004484-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | 5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase                                                | ACYPI008919-PA | <i>Acyrtosiphon pisum</i> |

|            |                                                                       |                |                           |
|------------|-----------------------------------------------------------------------|----------------|---------------------------|
| Metabolism | Sec24B protein, putative                                              | ACYPI005848-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Gmp synthase                                                          | ACYPI006177-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | mCG117402                                                             | ACYPI000180-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Lambda-crystallin                                                     | ACYPI001738-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Glyoxylate/hydroxypyruvate reductase                                  | ACYPI001693-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Fructose-1,6-bisphosphatase                                           | ACYPI002694-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Imaginal disk growth factor                                           | ACYPI001365-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Lysosomal alpha-mannosidase (mannosidase alpha class 2b member 1)     | ACYPI000371-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | 5-oxoprolinase (ATP-hydrolysing)                                      | ACYPI004211-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Aspartate ammonia lyase                                               | ACYPI006003-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Aldo-keto reductase                                                   | ACYPI005685-PA | <i>Acyrtosiphon pisum</i> |
| Metabolism | Aminomethyltransferase                                                | ACYPI002795-PA | <i>Acyrtosiphon pisum</i> |
| Regulatory | Bacteriocyte-specific cysteine rich proteins BCR family, peptide BCR1 | ACYPI32128     | <i>Acyrtosiphon pisum</i> |
| Regulatory | Bacteriocyte-specific cysteine rich proteins BCR family, peptide BCR2 | ACYPI38738     | <i>Acyrtosiphon pisum</i> |
| Regulatory | Bacteriocyte-specific cysteine rich proteins BCR family, peptide BCR3 | ACYPI44142     | <i>Acyrtosiphon pisum</i> |
| Regulatory | Bacteriocyte-specific cysteine rich proteins BCR family, peptide BCR6 | ACYPI49532     | <i>Acyrtosiphon pisum</i> |
| Regulatory | Bacteriocyte-specific cysteine rich proteins BCR family, peptide BCR8 | ACYPI45157     | <i>Acyrtosiphon pisum</i> |
| Regulatory | Secreted proteins SP family, peptide SP1                              | ACYPI008389    | <i>Acyrtosiphon pisum</i> |
| Regulatory | Secreted proteins SP family, peptide SP2                              | ACYPI000294    | <i>Acyrtosiphon pisum</i> |
| Regulatory | Secreted proteins SP family, peptide SP3                              | ACYPI005168    | <i>Acyrtosiphon pisum</i> |
| Regulatory | Secreted proteins SP family, peptide SP4                              | ACYPI009984    | <i>Acyrtosiphon pisum</i> |
| Regulatory | Secreted proteins SP family, peptide SP5a                             | ACYPI004796    | <i>Acyrtosiphon pisum</i> |
| Regulatory | Secreted proteins SP family, peptide SP6                              | ACYPI001839    | <i>Acyrtosiphon pisum</i> |
| Signaling  | Glean peptide GLEAN_28598                                             | ACYPI48598-PA  | <i>Acyrtosiphon pisum</i> |
| Signaling  | Glean peptide GLEAN_33885                                             | ACYPI53885-PA  | <i>Acyrtosiphon pisum</i> |
| Signaling  | past-1                                                                | ACYPI007266-PA | <i>Acyrtosiphon pisum</i> |
| Transport  | Putative vacuolar ATP synthase subunit E isoform 1                    | ACYPI006090-PA | <i>Acyrtosiphon pisum</i> |
| Transport  | Vacuolar ATP synthase subunit H                                       | ACYPI002312-PA | <i>Acyrtosiphon pisum</i> |
| Transport  | Vacuolar ATPase subunit C                                             | ACYPI006545-PA | <i>Acyrtosiphon pisum</i> |

|           |                                                                    |                |                           |
|-----------|--------------------------------------------------------------------|----------------|---------------------------|
| Transport | Vacuolar ATP synthase 16 kDa proteolipid subunit (Ductin) (VHA16K) | ACYPI003545-PA | <i>Acyrtosiphon pisum</i> |
| Transport | Potassium/chloride symporter, putative, partial                    | ACYPI000507-PA | <i>Acyrtosiphon pisum</i> |
| Transport | Cationic amino acid transporter                                    | ACYPI008904-PA | <i>Acyrtosiphon pisum</i> |

In some instances, the methods or compositions provided herein may be effective to decrease the host's resistance to parasites or pathogens (e.g., fungal, bacterial, or viral pathogens or parasites) in comparison to a host organism to which the modulating agent has not been administered. In some instances, the methods or compositions provided herein may be effective to decrease the host's resistance to a pathogen or parasite (e.g., fungal, bacterial, or viral pathogens; or parasitic mites) by about 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or greater than 100% relative to a reference level (e.g., a level found in a host that does not receive a modulating agent).

In some instances, the modulating agent is effective to alter the innate immune system of a host to indirectly change microbial diversity in the host relative to a host organism to which the modulating agent has not been administered. Invertebrates exhibit multiple immune reactions, some of which are homologous to immune mechanisms found in mammals. General principles of innate immunity in insects have been summarized by other reviews (Lemaitre et al., *Annu. Rev. Immunol.* 25:697-743, 2007; Charroux et al., *Fly* 4:40-47, 2010; Ganesan et al., *Curr. Top. Microbiol. Immunol.* 349:25-60, 2011; Chambers et al., *Curr. Opin. Immunol.* 24:10-14, 2012). For example, in *D. melanogaster*, there are two major inducible responses enabling local immunity at the intestinal epithelial cell layer: production of AMPs and synthesis of reactive oxygen species (ROS). While both of these induced responses might be seen as classical resistance mechanisms, they both include negative feedback loops and modulatory components, which can confer host tolerance toward the commensal gut microbiota.

Colonization of the gut by commensal bacteria could induce immune priming events resulting in the activation or alteration of the immune response not only toward recurrent colonization of commensal bacteria, but also against pathogens. Gut microbiota is essential not only for priming the immune system of the host to bacteria, e.g., mosquitoes to *Plasmodium*, but also for eliciting the priming response upon rechallenging the hosts with the bacteria. The bacteria-dependent priming response in mosquitoes is characterized by differentiation of prohemocytes into granulocytes and the presence of increased numbers of circulating granulocytes with changed morphology and binding properties. In another example, tsetse fly bacterial symbionts, including the gut-inhabiting Gammaproteobacterium *S. glossinidius*, are essential during larval development in order that the adult flies could present a trypanosome-refractory phenotype. In this case, the bacteria seem to influence the formation and integrity of the adult peritrophic matrix, thereby indirectly regulating the fly's ability to detect and respond to the presence of trypanosomes.

In some instances, an immune response is modulated by production of a modulating agent. For example, in the systemic immune response of *D. melanogaster*, Toll and IMD are the two major signaling pathways inducing antimicrobial peptide (AMP) production. Upon pathogen exposure, only the IMD pathway is active and triggers a local AMP response. Activation occurs by binding different variants of bacterial peptidoglycan (PGN) to extra- or intracellular epithelial receptors belonging to the peptidoglycan

recognition protein (PGRP) family. The protein Pirk sequesters specific PGN-binding receptors (PGRP-LC) in the cytoplasm, thereby reducing the number of these receptors localizing to the cell surface and retarding IMD pathway signaling. PGRP-LE ensures immune tolerance to the commensal microbiota via the up-regulation of amidases and Pirk. Downstream signaling via the IMD pathway results in activation of the transcription factor Relish, which in turn induces expression of several AMPs and other immunity-related genes. PGRP-LE induces a Relish-dependent immune response to pathogenic bacteria. PGRP-LE also ensures immune tolerance to the commensal microbiota via the up-regulation of amidases and Pirk.

The homeobox transcription factor Caudal specifically represses AMP gene transcription in the gut by binding to promoter regions. Caudal-deficiency causes a constitutive AMP production to occur and a shift in the gut microbiota. In one embodiment, an IMD pathway is inactivated to restrict expression of one or more AMPs and/or other immunity-related genes, e.g., caudal deficiency, thereby activating an immune response to a gut microbiota.

In some instances, exposure to pathogens in the gut triggers the generation of ROS via the membrane-associated dual oxidase (DUOX) system. For example, in *D. melanogaster*, PGN-dependent and PGN-independent signaling pathways produce ROS that causes oxidative stress not only on the bacteria but also on the host's epithelial cells. *D. melanogaster* eliminates excessive ROS by activating immune responsive catalases. This catalase production results in increased tolerance, due to a decrease in self-harm caused by the bacteria-induced immune response, possibly through locally restricted catalase activity, e.g., to the proximity of the epithelial surface. In one embodiment, immune responsive catalases are inactivated or repressed to maintain ROS production via the DUOX system and oxidative stress via PGN-dependent and PGN-independent signaling pathways to activate an immune response to a gut microbiota.

In some instances, the modulating agent is effective to alter (e.g., increase or decrease) gene expression in a host to increase or decrease the host's immune system response or immunoregulatory signaling, e.g., immune system response to microorganisms resident in the host (e.g., microorganisms resident in host bacteriocytes) in comparison to a host organism to which the modulating agent has not been administered. Nonlimiting examples of immune system related genes/proteins in bacteriocytes are shown in Table 9.

**Table 9: Immunoregulatory proteins that modulate bacteriocytes**

| Functional class | Protein name                                | Sequence accession number | Organism               |
|------------------|---------------------------------------------|---------------------------|------------------------|
| Regulatory       | BicD (Protein bicaudal D)                   | CG6605                    | <i>D. melanogaster</i> |
| Regulatory       | PGRP SA (Peptidoglycan recognition protein) | CG11709                   | <i>D. melanogaster</i> |
| Regulatory       | Relish (transcription factor)               | CG11992                   | <i>D. melanogaster</i> |
| Regulatory       | Pirk (poor lmd response upon knock-in)      | CG15678                   | <i>D. melanogaster</i> |
| Regulatory       | DUOX (Dual oxidase)                         | CG3131                    | <i>D. melanogaster</i> |
| Regulatory       | p38c (p38c MAP kinase)                      | CG33338                   | <i>D. melanogaster</i> |
| Regulatory       | MKP3 (Dual phosphatase)                     | CG14080                   | <i>D. melanogaster</i> |

|            |                                                   |         |                        |
|------------|---------------------------------------------------|---------|------------------------|
| Regulatory | CanB (calcineurin B )                             | CG4209  | <i>D. melanogaster</i> |
| Regulatory | Plad (phospholipase D)                            | CG12110 | <i>D. melanogaster</i> |
| Regulatory | Cact (cactus)                                     | CG5848  | <i>D. melanogaster</i> |
| Regulatory | DIF (Dorsal related immunity factor)              | CG6794  | <i>D. melanogaster</i> |
| Regulatory | dIAP2 (Drosophila Inhibitor of APoptosis2)        | CG8293  | <i>D. melanogaster</i> |
| Regulatory | Toll (Toll Interacting Protein )                  | CG5490  | <i>D. melanogaster</i> |
| Regulatory | gnbp1 (Gram Negative Binding Protein1)            | CG6895  | <i>D. melanogaster</i> |
| Regulatory | LysC (c-type lysozyme)                            | CR9111  | <i>D. melanogaster</i> |
| Regulatory | imd (immune deficiency)                           | CG5576  | <i>D. melanogaster</i> |
| Regulatory | Diap2 (Death-associated inhibitor of apoptosis 2) | CG8293  | <i>D. melanogaster</i> |
| Regulatory | ecsit                                             | CG10610 | <i>D. melanogaster</i> |

#### *v. Bacteria as modulating agents*

In some instances, the modulating agent described herein includes one or more bacteria. Numerous bacteria are useful in the compositions and methods described herein. In some instances, the agent is a bacterial species endogenously found in the host. In some instances, the bacterial modulating agent is an endosymbiotic bacterial species. In some instances, the bacterial modulating agent is a pathogen in the host. Non-limiting examples of bacteria that may be used as modulating agents include all bacterial species described herein in Section II of the detailed description and those listed in Table 1. For example, the modulating agent may be a bacterial species from any bacterial phyla present in host (e.g., insect, mollusk, or nematode) guts, including *Gammaproteobacteria*, *Alphaproteobacteria*, *Betaproteobacteria*, *Bacteroidetes*, *Firmicutes* (e.g., *Lactobacillus* and *Bacillus* spp.), *Clostridia*, *Actinomycetes*, *Spirochetes*, *Verrucomicrobia*, and *Actinobacteria*.

In some instances, the bacteria may be used as a modulating agent to stimulate an immune response in the host that leads to a decrease in the level, diversity, or metabolism of one or more microorganisms resident in the host. In some instances, the bacteria are delivered as live bacterial cells (e.g., *E. coli* cells). Alternatively, the bacteria may be delivered as heat-killed bacterial cells (e.g., heat-killed *E. coli* cells). In other instances, the bacteria may be delivered as lysates (e.g., prepared from lysed whole cells), or fractions thereof.

In some instances, the modulating agent includes genetically modified or transformed bacteria as described herein. In some instances, the bacteria are genetically modified. For example, the bacteria may be modified via the introduction of genetic material into a bacterium using standard methods in the art (e.g., transduction, transformation, or conjugation), thereby modifying or altering the cellular physiology, and/or biochemistry of the bacterium. Through the introduction of genetic material, the modified bacteria may acquire new functions or properties. In some instances, the genetically modified bacteria may produce and secrete a modulating agent described herein. For example, genetically modified bacteria may produce polypeptides, small molecules, or nucleic acids that target specific host endosymbionts or other microorganisms resident in the host. In some instances, the genetically modified bacteria may be used to produce a product that stimulates a host immune response.

In some instances, the genetically engineered or transformed bacteria are provided to impart new functionalities to the host. New functionalities may include, for example, the ability to degrade pesticides (e.g., insecticides, molluscicides, or nematocides; e.g., a pesticide listed in Table 11), plant allelochemicals, or produce nutrients. For example, genetically modified bacteria may be generated from naturally occurring bacteria isolated from pesticide (e.g., insecticide, molluscicide, or nematocides; e.g., a pesticide listed in Table 11) resistant pests, e.g. *Burkholderia* strains from the insect *R. pedestris*. When bacteria from insecticide resistant pests are cultured with commensal bacteria isolated from a host of interest, e.g., bees, genes imparting insecticide resistance (e.g., the ability to use the insecticide as a carbon source) may be transferred to the commensal bacteria of the host of interest, e.g., bees. The genetically modified bacteria are then reintroduced into the host, e.g., bees, and insecticide resistant insects are selected by breeding them in an environment rich in the insecticide. Any bacterial phyla that are commonly present in a host may be modified to have a new functionality, including, but not limited to, *Gammaproteobacteria*, *Alphaproteobacteria*, *Betaproteobacteria*, *Bacteroidetes*, *Firmicutes* including *Lactobacillus* and *Bacillus* species, *Clostridia*, *Actinomycetes*, *Spirochetes*, *Verrucomicrobia*, *Actinobacteria*, and others.

The bacterium may be genetically modified to decrease the fitness of the host or to kill the host. In some instances, a bacterium may be genetically modified to improve its survival in host cells, to increase its toxicity to the host cell, and/or to provide a function that decreases the fitness of the host (e.g., decreased resistance to a pesticide, e.g., a pesticide listed in Table 11). In some instances, the bacterium is modified to no longer synthesize an essential molecule that it typically provides to the host cell. In some instances, the bacterium is genetically modified so that the population of the bacteria is increased to a level that negatively impacts the host, e.g., by excessive utilization of essential materials or machinery or by competing with beneficial microorganisms resident in the host. In some instances, the bacteria only kill the host following further human intervention.

The genetically modified bacteria described herein can be generated using any methods known in the art. For example, methods for the delivery of nucleic acids to bacteria include, but are not limited to, different chemical, electrochemical and biological approaches, for example, heat shock transformation, electroporation, transfection, for example liposome-mediated transfection, DEAE-Dextran-mediated transfection or calcium phosphate transfection. In some instances, a nucleic acid construct, for example, an expression construct including at least one nucleic acid sequence is introduced into the bacteria using a vehicle, or vector, for transferring genetic material. Vectors for transferring genetic material to bacteria are well known to those of skill in the art and include, for example, plasmids, artificial chromosomes, and viral vectors. Methods for the construction of nucleic acid constructs, including expression constructs including constitutive or inducible heterologous promoters, knockout and knockdown constructs, as well as methods and vectors for the delivery of a nucleic acid or nucleic acid construct to bacteria are well known to those of skill in the art, and are described, for example, in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 3rd edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2001; Amberg et al., *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*, Cold Spring Harbor Laboratory Press, 2005; Abelson et al., *Guide to Yeast Genetics and Molecular Biology*, Part A, Volume 194 (Methods in Enzymology Series, 194), Academic Press, 2004; Guthrie et al., *Guide to*

*Yeast Genetics and Molecular and Cell Biology*, 1st edition, Part B, Volume 350 (Methods in Enzymology, Vol 350), Academic Press, 2002; Stephanopoulos et al., *Metabolic Engineering: Principles and Methodologies*, 1st edition, Academic Press, 1998; and Smolke, *The Metabolic Pathway Engineering Handbook: Fundamentals*, 1st edition, CRC Press, 2009, all of which are incorporated by reference  
 5 herein in their entireties.

#### *vi. Fungi as modulating agents*

In some instances, the modulating agent described herein includes one or more fungi. Numerous fungi are useful in the compositions and methods. For example, the fungal modulating agent may be  
 10 yeast such as *Sacchromyces cerevisiae* or *Pichia pastoris*.

In some instances, the fungi (e.g., *S. cerevisiae* or *P. pastoris*) may be used as modulating agents to stimulate an immune response in the host that leads to a decrease in the levels, diversity, or metabolism of one or more microorganisms (e.g., bacteria or fungi) resident in the host. In some instances, the fungi are delivered as live fungal cells (e.g., *P. pastoris* cells, see Example 15).

15 Alternatively, the fungi may be delivered as heat-killed fungal cells (e.g., heat-killed *S. cerevisiae* cells). In other instances, the fungi may be delivered as lysates (e.g., prepared from lysed whole cells), or fractions thereof.

In some instances, the modulating agent includes a genetically modified or transformed fungus as described herein. In some instances, the fungus is genetically modified. For example, the fungus may  
 20 be modified via the introduction of genetic material into a fungus using standard methods in the art, thereby modifying or altering the cellular physiology and/or biochemistry of the fungus. Through the introduction of genetic material, the modified fungus may acquire new functions or properties. In some instances, the genetically modified fungus may produce and secrete a modulating agent described herein. In some instances, the genetically modified fungus may be used to produce a product that  
 25 stimulates a host immune response.

In some instances, the genetically engineered or transformed fungus is provided to impart new functionalities to the host. New functionalities may include, for example, the ability to degrade pesticides (e.g., insecticides), plant allelochemicals, or produce nutrients. Any fungal phyla that are commonly present in a host may be genetically modified to have a new functionality, including, but not limited to,  
 30 *Candida*, *Metschnikowia*, *Debaromyces*, *Scheffersomyces shehatae* and *Scheffersomyces stipites*, *Starmerella*, *Pichia*, *Trichosporon*, *Cryptococcus*, *Pseudozyma*, and yeast-like symbionts from the subphylum Pezizomycotina (e.g., *Symbiotaphrina buchneri* and *Symbiotaphrina kochii*).

As illustrated by Example 15, yeast, such as *P. pastoris*, can be used as a modulating agent that stimulates a host immune response (e.g., in an insect, e.g., an aphid) that, in turn, alters the activity,  
 35 levels, or metabolism of endosymbiotic bacteria, such as a *Buchnera* spp., resident in the host and thereby modulates (e.g., decreases) the fitness of the host.

#### *vii. Modifications to modulating agents*

In some instances, the nucleic acid molecule, peptide, polypeptide, or antibody molecule can be  
 40 modified. For example, the modification can be a chemical modification, e.g., conjugation to a marker,

e.g., fluorescent marker or a radioactive marker. In other examples, the modification can include conjugation or operational linkage to a moiety that enhances the stability, delivery, targeting, bioavailability, or half-life of the agent, e.g., a lipid, a glycan, a polymer (e.g., PEG), a cation moiety.

(a) *Fusions*

5 Any of the modulating agents described herein may be fused or linked to an additional moiety. In some instances, the modulating agent includes a fusion of one or more additional moieties (e.g., 1 additional moiety, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more additional moieties). In some instances, the additional moiety is any one of the modulating agents described herein (e.g., a peptide, polypeptide, small molecule, or antibiotic). Alternatively, the additional moiety may not act as modulating agent itself but may instead  
10 serve a secondary function. For example, the additional moiety may to help the modulating agent access, bind, or become activated at a target site in the host (e.g., at a host gut or a host bacteriocyte) or at a target microorganism resident in the host.

In some instances, the additional moiety may help the modulating agent penetrate a target host cell or target microorganism resident in the host. For example, the additional moiety may include a cell  
15 penetrating peptide. Cell penetrating peptides (CPPs) may be natural sequences derived from proteins; chimeric peptides that are formed by the fusion of two natural sequences; or synthetic CPPs, which are synthetically designed sequences based on structure–activity studies. In some instances, CPPs have the capacity to ubiquitously cross cellular membranes (e.g., prokaryotic and eukaryotic cellular membranes) with limited toxicity. Further, CPPs may have the capacity to cross cellular membranes via energy-  
20 dependent and/or independent mechanisms, without the necessity of a chiral recognition by specific receptors. Non-limiting examples of CPPs are listed in Table 10.

**Table 10: Examples of Cell Penetrating Peptides (CPPs)**

| Peptide                | Origin                  | Sequence                                       |
|------------------------|-------------------------|------------------------------------------------|
| <u>Protein-derived</u> |                         |                                                |
| Penetratin             | Antennapedia            | RQIKIWFQNRRMKWKK<br>(SEQ ID NO: 82)            |
| Tat peptide            | Tat                     | GRKKRRQRRRPPQ<br>(SEQ ID NO: 83)               |
| pVEC                   | Cadherin                | LLIILRRRIRKQAHASK<br>(SEQ ID NO: 84)           |
| <u>Chimeric</u>        |                         |                                                |
| Transportan            | Galanine/Mastoparan     | GWTLNSAGYLLGKINLKALAALAKKIL<br>(SEQ ID NO: 85) |
| MPG                    | HIV-gp41/SV40 T-antigen | GALFLGFLGAAGSTMGAWSQPKKKRKV<br>(SEQ ID NO: 86) |

|                               |                                          |                                          |
|-------------------------------|------------------------------------------|------------------------------------------|
| Pep-1                         | HIV-reverse transcriptase/SV40 T-antigen | KETWWETWWTEWSQPKKKRKV<br>(SEQ ID NO: 87) |
|                               |                                          |                                          |
| <u>Synthetic</u>              |                                          |                                          |
| Polyarginines                 | Based on Tat peptide                     | $(R)_n$ ; $6 < n < 12$                   |
| MAP                           | de novo                                  | KLALKLALKALKKAALKLA<br>(SEQ ID NO: 88)   |
| R <sub>6</sub> W <sub>3</sub> | Based on penetratin                      | RRWWRRWRR<br>(SEQ ID NO: 89)             |

In some instances, the additional moiety is a peptide nucleic acid (PNAs). Peptide nucleic acids (PNAs) include one or more nucleic acid side chains linked to an amide backbone. One or more amino acid units in the PNA have an amide containing backbone, e.g., aminoethyl-glycine, similar to a peptide backbone, with a nucleic acid side chain in place of the amino acid side chain. PNAs are known to hybridize complementary DNA and RNA with higher affinity than their oligonucleotide counterparts. This character of PNA not only makes the polypeptide of the invention a stable hybrid with the nucleic acid side chains, but at the same time, the neutral backbone and hydrophobic side chains result in a hydrophobic unit within the polypeptide. Examples of PNA moieties include a molecule that includes a peptide, such as a CPP (e.g., an amino acid sequence having at least at least 80% (e.g., 80%, 90%, 95%, 97%, 99%, or greater) identity to SEQ ID NO: 106) and a nucleotide sequence having at least 80% (e.g., 80%, 90%, 95%, 97%, 99%, or greater) identity to the sequence of any one of SEQ ID NOs: 105 or 151-153.

The nucleic acid side chain includes, but is not limited to, a purine or a pyrimidine side chain such as adenine, cytosine, guanine, thymine, and uracil. In one instances, the nucleic acid side chain includes a nucleoside analog, such as 5-fluorouracil; 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 4-methylbenzimidazole, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, dihydrouridine, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil (acp<sup>3</sup> U), 2,6-diaminopurine, 3-nitropyrrole, inosine, thiouridine, queuosine, wyosine, diaminopurine, isoguanine, isocytosine, diaminopyrimidine, 2,4-difluorotoluene, isoquinoline, pyrrolo[2,3-β]pyridine, and any others that can base pair with a purine or a pyrimidine side chain.

In other instances, the additional moiety helps the modulating agent bind a target microorganism (e.g., a fungi or bacterium) resident in the host. The additional moiety may include one or more targeting

domains. In some instances, the targeting domain may target the modulating agent to one or more microorganisms (e.g., bacterium or fungus) resident in the gut of the host. In some instances, the targeting domain may target the modulating agent to a specific region of the host (e.g., host gut or bacteriocyte) to access microorganisms that are generally present in said region of the host. For example, the targeting domain may target the modulating agent to the foregut, midgut, or hindgut of the host. In other instances, the targeting domain may target the modulating agent to a bacteriocyte in the host and/or one or more specific bacteria resident in a host bacteriocyte.

*(b) Pre- or Pro-domains*

In some instances, the modulating agent may include a pre- or pro- amino acid sequence. For example, the modulating agent may be an inactive protein or peptide that can be activated by cleavage or post-translational modification of a pre- or pro-sequence. In some instances, the modulating agent is engineered with an inactivating pre- or pro-sequence. For example, the pre- or pro-sequence may obscure an activation site on the modulating agent, e.g., a receptor binding site, or may induce a conformational change in the modulating agent. Thus, upon cleavage of the pre- or pro-sequence, the modulating agent is activated.

Alternatively, the modulating agent may include a pre- or pro-small molecule, e.g., an antibiotic. The modulating agent may be an inactive small molecule described herein that can be activated in a target environment inside the host. For example, the small molecule may be activated upon reaching a certain pH in the host gut.

*(c) Linkers*

In instances where the modulating agent is connected to an additional moiety, the modulating agent may further include a linker. For example, the linker may be a chemical bond, e.g., one or more covalent bonds or non-covalent bonds. In some instances, the linker may be a peptide linker (e.g., 2, 3, 4, 5, 6, 8, 10, 12, 14, 16, 20, 25, 30, 35, 40, or more amino acids longer). The linker may include any flexible, rigid, or cleavable linkers described herein.

A flexible peptide linker may include any of those commonly used in the art, including linkers having sequences having primarily Gly and Ser residues ("GS" linker). Flexible linkers may be useful for joining domains that require a certain degree of movement or interaction and may include small, non-polar (e.g. Gly) or polar (e.g. Ser or Thr) amino acids.

Alternatively, a peptide linker may be a rigid linker. Rigid linkers are useful to keep a fixed distance between moieties and to maintain their independent functions. Rigid linkers may also be useful when a spatial separation of the domains is critical to preserve the stability or bioactivity of one or more components in the fusion. Rigid linkers may, for example, have an alpha helix-structure or Pro-rich sequence, (XP)<sub>n</sub>, with X designating any amino acid, preferably Ala, Lys, or Glu.

In yet other instances, a peptide linker may be a cleavable linker. In some instances, linkers may be cleaved under specific conditions, such as the presence of reducing reagents or proteases. *In vivo* cleavable linkers may utilize the reversible nature of a disulfide bond. One example includes a thrombin-sensitive sequence (e.g., PRS) between two Cys residues. *In vitro* thrombin treatment of CPRSC results

in the cleavage of the thrombin-sensitive sequence, while the reversible disulfide linkage remains intact. Such linkers are known and described, e.g., in Chen et al., *Adv. Drug Deliv. Rev.* 65(10):1357-1369, 2013. Cleavage of linkers in fusions may also be carried out by proteases that are expressed *in vivo* under conditions in specific cells or tissues of the host or microorganisms resident in the host. In some instances, cleavage of the linker may release a free functional, modulating agent upon reaching a target site or cell.

Fusions described herein may alternatively be linked by a linking molecule, including a hydrophobic linker, such as a negatively charged sulfonate group; lipids, such as a poly (--CH<sub>2</sub>--) hydrocarbon chains, such as polyethylene glycol (PEG) group, unsaturated variants thereof, hydroxylated variants thereof, amidated or otherwise N-containing variants thereof, non-carbon linkers; carbohydrate linkers; phosphodiester linkers, or other molecule capable of covalently linking two or more molecules, e.g., two modulating agents. Non-covalent linkers may be used, such as hydrophobic lipid globules to which the modulating agent is linked, for example, through a hydrophobic region of the modulating agent or a hydrophobic extension of the modulating agent, such as a series of residues rich in leucine, isoleucine, valine, or perhaps also alanine, phenylalanine, or even tyrosine, methionine, glycine, or other hydrophobic residue. The modulating agent may be linked using charge-based chemistry, such that a positively charged moiety of the modulating agent is linked to a negative charge of another modulating agent or an additional moiety.

#### IV. Formulations and Compositions

The compositions described herein may be formulated either in pure form (e.g., the composition contains only the modulating agent) or together with one or more additional agents (such as excipient, delivery vehicle, carrier, diluent, stabilizer, etc.) to facilitate application or delivery of the compositions. Examples of suitable excipients and diluents include, but are not limited to, lactose, dextrose, sucrose, sorbitol, mannitol, starches, gum acacia, calcium phosphate, alginates, tragacanth, gelatin, calcium silicate, microcrystalline cellulose, polyvinylpyrrolidone, cellulose, water, saline solution, syrup, methylcellulose, methyl- and propylhydroxybenzoates, talc, magnesium stearate, and mineral oil.

In some instances, the composition includes a delivery vehicle or carrier. In some instances, the delivery vehicle includes an excipient. Exemplary excipients include, but are not limited to, solid or liquid carrier materials, solvents, stabilizers, slow-release excipients, colorings, and surface-active substances (surfactants). In some instances, the delivery vehicle is a stabilizing vehicle. In some instances, the stabilizing vehicle includes a stabilizing excipient. Exemplary stabilizing excipients include, but are not limited to, epoxidized vegetable oils, antifoaming agents, e.g. silicone oil, preservatives, viscosity regulators, binding agents and tackifiers. In some instances, the stabilizing vehicle is a buffer suitable for the modulating agent. In some instances, the composition is microencapsulated in a polymer bead delivery vehicle. In some instances, the stabilizing vehicle protects the modulating agent against UV and/or acidic conditions. In some instances, the delivery vehicle contains a pH buffer. In some instances, the composition is formulated to have a pH in the range of about 4.5 to about 9.0, including for example pH ranges of about any one of 5.0 to about 8.0, about 6.5 to about 7.5, or about 6.5 to about 7.0.

Depending on the intended objectives and prevailing circumstances, the composition may be formulated into emulsifiable concentrates, suspension concentrates, directly sprayable or dilutable solutions, coatable pastes, diluted emulsions, spray powders, soluble powders, dispersible powders, wettable powders, dusts, granules, encapsulations in polymeric substances, microcapsules, foams, aerosols, carbon dioxide gas preparations, tablets, resin preparations, paper preparations, nonwoven fabric preparations, or knitted or woven fabric preparations. In some instances, the composition is a liquid. In some instances, the composition is a solid. In some instances, the composition is an aerosol, such as in a pressurized aerosol can. In some instances, the composition is present in the waste (such as feces) of the pest. In some instances, the composition is present in or on a live pest.

In some instances, the delivery vehicle is the food or water of the host. In other instances, the delivery vehicle is a food source for the host. In some instances, the delivery vehicle is a food bait for the host. In some instances, the composition is a comestible agent consumed by the host. In some instances, the composition is delivered by the host to a second host, and consumed by the second host. In some instances, the composition is consumed by the host or a second host, and the composition is released to the surrounding of the host or the second host via the waste (such as feces) of the host or the second host. In some instances, the modulating agent is included in food bait intended to be consumed by a host or carried back to its colony.

In some instances, the delivery vehicle is a bacterial vector. The modulating agent can be incorporated in a bacterial vector using any suitable cloning methods and reagents known in the art, such as described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989. "Bacterial vector" as used herein refers to any genetic element, such as plasmids, bacteriophage vectors, transposons, cosmids, and chromosomes, which is capable of replication inside bacterial cells and which is capable of transferring genes between cells. Exemplary bacterial vectors include, but are not limited to, lambda vector system gtl 1, gt WES.tB, Charon 4, and plasmid vectors such as pBR322, pBR325, pACYC177, pACYC184, pUC8, pUC9, pUC18, pUC19, pLG339, pR290, pKC37, pKCIOL, SV 40, pBluescript II SK +/- or KS +/- (see "Stratagene Cloning Systems" Catalog, Stratagene, La Jolla, California, 1993), pQE, pIH821, pGEX, pET series (see Studier et al., "Use of T7 RNA Polymerase to Direct Expression of Cloned Genes," *Gene Expression Technology*, Vol. 185, 1990), and any derivatives thereof.

Each bacterial vector may encode one or more modulating agents. In some instances, the bacterial vector includes a nucleic acid molecule encoding a polypeptide to be expressed in the target symbiotic bacterium or a host bacterium. In some instances, the bacterial vector includes a nucleic acid molecule encoding a bacteriocin to be expressed in the target bacterium. In some instances, the bacterial vector further includes one or more regulatory elements, such as promoters, termination signals, and transcription and translation elements. In some instances, the regulatory sequence is operably linked to a nucleic acid encoding a gene (e.g., any of the nucleic acids described herein) to be expressed in the target symbiotic bacterium.

In some instances, the bacterial vector is introduced into a bacterium to be consumed by the host or a member in the colony of the host. In some instances, the bacterium is the target symbiotic bacterium. In some instances, the bacterium is a naturally occurring bacterium of the gut of the host, or a

genetically modified derivative thereof, which can be easily introduced to the host through ingestion.

Exemplary bacteria for use in carrying the bacterial vector include, but are not limited to, *Proteobacter*, including the genus *Pseudomonas*; *Actinobacter*, including *Priopionibacterium* and *Corynebacterium*;

*Firmicutes*, including the any species of the genera *Mycoplasma*, *Bacillus*, *Streptococcus*,

5 *Staphylococcus*; *Fibrobacteres*; *Spirochaetes*, including *Treponema* and *Borrelia*; *Bacteroides*, including the genera *Bacteroides* and *Flavobacterium*. Also suitable are any bacteria of the *Enterobacteriaceae*, including the genus *Serratia*, including, but not limited to *S. marcescens*, *S. entomophila*, *S.*

*proteamaculans*; any species of *Enterobacter*, including, but not limited to, *E. cloacae*, *E. aerogenes*, *E. dissolvens*, *E. agglomerans*, *E. hafiii*; and any species belonging to the following genera: *Citrobacter*,

10 *Escherichia*, *Klebsiella*, *Kluyvera*, *Panotaea*, *Proteus*, *Salmonella*, *Xenorhabdus*, and *Yokenella*.

In some instances, the modulating agent may make up about 0.1% to about 100% of the composition, such as any one of about 0.01% to about 100%, about 1% to about 99.9%, about 0.1% to about 10%, about 1% to about 25%, about 10% to about 50%, about 50% to about 99%, or about 0.1% to about 90% of active ingredients (e.g., a polypeptide, nucleic acid, small molecule, or combinations thereof). In some instances, the composition includes at least any of 0.1%, 0.5%, 1%, 5%, 10%, 20%, 15 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or more active ingredients (e.g., a polypeptide, nucleic acid, small molecule, or combinations thereof). In some instances, the concentrated agents are preferred as commercial products, the final user normally uses diluted agents, which have a substantially lower concentration of active ingredient.

20 Any of the formulations described herein may be used in the form of a bait, a coil, an electric mat, a smoking preparation, a fumigant, or a sheet.

#### *i. Liquid Formulations*

The compositions provided herein may be in a liquid formulation. Liquid formulations are generally mixed with water, but in some instances may be used with crop oil, diesel fuel, kerosene or other light oil as a carrier. The amount of active ingredient often ranges from about 0.5 to about 80 percent by weight.

An emulsifiable concentrate formulation may contain a liquid active ingredient, one or more petroleum-based solvents, and an agent that allows the formulation to be mixed with water to form an emulsion. Such concentrates may be used in agricultural, ornamental and turf, forestry, structural, food processing, livestock, and public health pest formulations. These may be adaptable to application equipment from small portable sprayers to hydraulic sprayers, low-volume ground sprayers, mist blowers, and low-volume aircraft sprayers. Some active ingredients are readily dissolve in a liquid carrier. When mixed with a carrier, they form a solution that does not settle out or separate, e.g., a homogenous solution. Formulations of these types may include an active ingredient, a carrier, and one or more other ingredients. Solutions may be used in any type of sprayer, indoors and outdoors.

In some instances, the composition may be formulated as an invert emulsion. An invert emulsion is a water-soluble active ingredient dispersed in an oil carrier. Invert emulsions require an emulsifier that allows the active ingredient to be mixed with a large volume of petroleum-based carrier, usually fuel oil. Invert emulsions aid in reducing drift. With other formulations, some spray drift results when water

droplets begin to evaporate before reaching target surfaces; as a result the droplets become very small and lightweight. Because oil evaporates more slowly than water, invert emulsion droplets shrink less and more active ingredient reaches the target. Oil further helps to reduce runoff and improve rain resistance. It further serves as a sticker-spreader by improving surface coverage and absorption. Because droplets are relatively large and heavy, it is difficult to get thorough coverage on the undersides of foliage. Invert emulsions are most commonly used along rights-of-way where drift to susceptible non-target areas can be a problem.

A flowable or liquid formulation combines many of the characteristics of emulsifiable concentrates and wettable powders. Manufacturers use these formulations when the active ingredient is a solid that does not dissolve in either water or oil. The active ingredient, impregnated on a substance such as clay, is ground to a very fine powder. The powder is then suspended in a small amount of liquid. The resulting liquid product is quite thick. Flowables and liquids share many of the features of emulsifiable concentrates, and they have similar disadvantages. They require moderate agitation to keep them in suspension and leave visible residues, similar to those of wettable powders.

Flowables/liquids are easy to handle and apply. Because they are liquids, they are subject to spilling and splashing. They contain solid particles, so they contribute to abrasive wear of nozzles and pumps. Flowable and liquid suspensions settle out in their containers. Because flowable and liquid formulations tend to settle, packaging in containers of five gallons or less makes remixing easier.

Aerosol formulations contain one or more active ingredients and a solvent. Most aerosols contain a low percentage of active ingredients. There are two types of aerosol formulations—the ready-to-use type commonly available in pressurized sealed containers and those products used in electrical or gasoline-powered aerosol generators that release the formulation as a smoke or fog.

Ready to use aerosol formulations are usually small, self-contained units that release the formulation when the nozzle valve is triggered. The formulation is driven through a fine opening by an inert gas under pressure, creating fine droplets. These products are used in greenhouses, in small areas inside buildings, or in localized outdoor areas. Commercial models, which hold five to 5 pounds of active ingredient, are usually refillable.

Smoke or fog aerosol formulations are not under pressure. They are used in machines that break the liquid formulation into a fine mist or fog (aerosol) using a rapidly whirling disk or heated surface.

#### *ii. Dry or Solid Formulations*

Dry formulations can be divided into two types: ready-to-use and concentrates that must be mixed with water to be applied as a spray. Most dust formulations are ready to use and contain a low percentage of active ingredients (less than about 10 percent by weight), plus a very fine, dry inert carrier made from talc, chalk, clay, nut hulls, or volcanic ash. The size of individual dust particles varies. A few dust formulations are concentrates and contain a high percentage of active ingredients. Mix these with dry inert carriers before applying. Dusts are always used dry and can easily drift to non-target sites.

*iii. Granule or Pellet Formulations*

In some instances, the composition is formulated as granules. Granular formulations are similar to dust formulations, except granular particles are larger and heavier. The coarse particles may be made from materials such as clay, corncobs, or walnut shells. The active ingredient either coats the outside of the granules or is absorbed into them. The amount of active ingredient may be relatively low, usually ranging from about 0.5 to about 15 percent by weight. Granular formulations are most often used to apply to the soil, insects, mollusks, or nematodes living in the soil, or absorption into plants through the roots. Granular formulations are sometimes applied by airplane or helicopter to minimize drift or to penetrate dense vegetation. Once applied, granules may release the active ingredient slowly. Some granules require soil moisture to release the active ingredient. Granular formulations also are used to control larval mosquitoes and other aquatic pests. Granules are used in agricultural, structural, ornamental, turf, aquatic, right-of-way, and public health (biting insect) pest-control operations.

In some instances, the composition is formulated as pellets. Most pellet formulations are very similar to granular formulations; the terms are used interchangeably. In a pellet formulation, however, all the particles are the same weight and shape. The uniformity of the particles allows use with precision application equipment.

*iv. Powders*

In some instances, the composition is formulated as a powder. In some instances, the composition is formulated as a wettable powder. Wettable powders are dry, finely ground formulations that look like dusts. They usually must be mixed with water for application as a spray. A few products, however, may be applied either as a dust or as a wettable powder—the choice is left to the applicator. Wettable powders have about 1 to about 95 percent active ingredient by weight; in some cases more than about 50 percent. The particles do not dissolve in water. They settle out quickly unless constantly agitated to keep them suspended. They can be used for most pest problems and in most types of spray equipment where agitation is possible. Wettable powders have excellent residual activity. Because of their physical properties, most of the formulation remains on the surface of treated porous materials such as concrete, plaster, and untreated wood. In such cases, only the water penetrates the material.

In some instances, the composition is formulated as a soluble powder. Soluble powder formulations look like wettable powders. However, when mixed with water, soluble powders dissolve readily and form a true solution. After they are mixed thoroughly, no additional agitation is necessary. The amount of active ingredient in soluble powders ranges from about 15 to about 95 percent by weight; in some cases more than about 50 percent. Soluble powders have all the advantages of wettable powders and none of the disadvantages, except the inhalation hazard during mixing.

In some instances, the composition is formulated as a water-dispersible granule. Water-dispersible granules, also known as dry flowables, are like wettable powders, except instead of being dust-like, they are formulated as small, easily measured granules. Water-dispersible granules must be mixed with water to be applied. Once in water, the granules break apart into fine particles similar to wettable powders. The formulation requires constant agitation to keep it suspended in water. The percentage of active ingredient is high, often as much as 90 percent by weight. Water-dispersible

granules share many of the same advantages and disadvantages of wettable powders, except they are more easily measured and mixed. Because of low dust, they cause less inhalation hazard to the applicator during handling

#### 5 *v. Bait*

In some instances, the composition includes a bait. The bait can be in any suitable form, such as a solid, paste, pellet or powdered form. The bait can also be carried away by the host back to a population of said host (e.g., a colony or hive). The bait can then act as a food source for other members of the colony, thus providing an effective modulating agent for a large number of hosts and potentially an  
10 entire host colony.

The baits can be provided in a suitable "housing" or "trap." Such housings and traps are commercially available and existing traps can be adapted to include the compositions described herein. The housing or trap can be box-shaped for example, and can be provided in pre-formed condition or can be formed of foldable cardboard for example. Suitable materials for a housing or trap include plastics and  
15 cardboard, particularly corrugated cardboard. The inside surfaces of the traps can be lined with a sticky substance in order to restrict movement of the host once inside the trap. The housing or trap can contain a suitable trough inside which can hold the bait in place. A trap is distinguished from a housing because the host cannot readily leave a trap following entry, whereas a housing acts as a "feeding station" which provides the host with a preferred environment in which they can feed and feel safe from predators.

20

#### *vi. Attractants*

In some instances, the composition includes an attractant (e.g., a chemoattractant). The attractant may attract an adult host or immature host (e.g., larva) to the vicinity of the composition. Attractants include pheromones, a chemical that is secreted by an animal, especially a host (e.g., insect,  
25 mollusk, or nematode), which influences the behavior or development of others of the same species. Other attractants include sugar and protein hydrolysate syrups, yeasts, and rotting meat. Attractants also can be combined with an active ingredient and sprayed onto foliage or other items in the treatment area.

Various attractants are known which influence host behavior as a host's search for food, oviposition or mating sites, or mates. Attractants useful in the methods and compositions described  
30 herein include, for example, eugenol, phenethyl propionate, ethyl dimethylisobutyl-cyclopropane carboxylate, propyl benzodioxancarboxylate, cis-7,8-epoxy-2-methyloctadecane, trans-8,trans-0-dodecadienol, cis-9-tetradecenol (with cis-11-hexadecenol), trans-11-tetradecenol, cis-11-hexadecenol, (Z)-11,12-hexadecadienal, cis-7-dodecenyl acetate, cis-8-dodecenyl acetate, cis-9-dodecenyl acetate, cis-9-tetradecenyl acetate, cis-11-tetradecenyl acetate, trans-11-tetradecenyl acetate (with cis-11), cis-  
35 9,trans-11-tetradecadienyl acetate (with cis-9,trans-12), cis-9,trans-1 2-tetradecadienyl acetate, cis-7,cis-11-hexadecadienyl acetate (with cis-7,trans-11), cis-3,cis-13-octadecadienyl acetate, trans-3,cis-13-octadecadienyl acetate, anethole and isoamyl salicylate.

Means other than chemoattractants may also be used to attract a host (e.g., insect, mollusk, or nematode), including lights in various wavelengths or colors.

40

*vii. Nanocapsules/Microencapsulation/Liposomes*

In some instances, the composition is provided in a microencapsulated formulation.

Microencapsulated formulations are mixed with water and sprayed in the same manner as other sprayable formulations. After spraying, the plastic coating breaks down and slowly releases the active ingredient.

*viii. Carriers*

Any of the compositions described herein may be formulated to include the modulating agent described herein and an inert carrier. Such carrier can be a solid carrier, a liquid carrier, a gel carrier, and/or a gaseous carrier. In certain instances, the carrier can be a seed coating. The seed coating is any non-naturally occurring formulation that adheres, in whole or part, to the surface of the seed. The formulation may further include an adjuvant or surfactant. The formulation can also include one or more modulating agents to enlarge the action spectrum.

A solid carrier used for formulation includes finely-divided powder or granules of clay (e.g. kaolin clay, diatomaceous earth, bentonite, Fubasami clay, acid clay, etc.), synthetic hydrated silicon oxide, talc, ceramics, other inorganic minerals (e.g., sericite, quartz, sulfur, activated carbon, calcium carbonate, hydrated silica, etc.), a substance which can be sublimated and is in the solid form at room temperature (e.g., 2,4,6-triisopropyl-1,3,5-trioxane, naphthalene, p-dichlorobenzene, camphor, adamantan, etc.); wool; silk; cotton; hemp; pulp; synthetic resins (e.g., polyethylene resins such as low-density polyethylene, straight low-density polyethylene and high-density polyethylene; ethylene-vinyl ester copolymers such as ethylene-vinyl acetate copolymers; ethylene-methacrylic acid ester copolymers such as ethylene-methyl methacrylate copolymers and ethylene-ethyl methacrylate copolymers; ethylene-acrylic acid ester copolymers such as ethylene-methyl acrylate copolymers and ethylene-ethyl acrylate copolymers; ethylene-vinylcarboxylic acid copolymers such as ethylene-acrylic acid copolymers; ethylene-tetracyclododecene copolymers; polypropylene resins such as propylene homopolymers and propylene-ethylene copolymers; poly-4-methylpentene-1, polybutene-1, polybutadiene, polystyrene; acrylonitrile-styrene resins; styrene elastomers such as acrylonitrile-butadiene-styrene resins, styrene-conjugated diene block copolymers, and styrene-conjugated diene block copolymer hydrides; fluororesins; acrylic resins such as poly(methyl methacrylate); polyamide resins such as nylon 6 and nylon 66; polyester resins such as polyethylene terephthalate, polyethylene naphthalate, polybutylene terephthalate, and polycyclohexylenedimethylene terephthalate; polycarbonates, polyacetals, polyacrylsulfones, polyarylates, hydroxybenzoic acid polyesters, polyetherimides, polyester carbonates, polyphenylene ether resins, polyvinyl chloride, polyvinylidene chloride, polyurethane, and porous resins such as foamed polyurethane, foamed polypropylene, or foamed ethylene, etc.), glasses, metals, ceramics, fibers, cloths, knitted fabrics, sheets, papers, yarn, foam, porous substances, and multifilaments.

A liquid carrier may include, for example, aromatic or aliphatic hydrocarbons (e.g., xylene, toluene, alkyl naphthalene, phenylxylethane, kerosine, gas oil, hexane, cyclohexane, etc.), halogenated hydrocarbons (e.g., chlorobenzene, dichloromethane, dichloroethane, trichloroethane, etc.), alcohols (e.g., methanol, ethanol, isopropyl alcohol, butanol, hexanol, benzyl alcohol, ethylene glycol, etc.), ethers (e.g., diethyl ether, ethylene glycol dimethyl ether, diethylene glycol monomethyl ether, diethylene glycol

monoethyl ether, propylene glycol monomethyl ether, tetrahydrofuran, dioxane, etc.), esters (e.g., ethyl acetate, butyl acetate, etc.), ketones (e.g., acetone, methyl ethyl ketone, methyl isobutyl ketone, cyclohexanone, etc.), nitriles (e.g., acetonitrile, isobutyronitrile, etc.), sulfoxides (e.g., dimethyl sulfoxide, etc.), amides (e.g., N,N-dimethylformamide, N,N-dimethylacetamide, cyclic imides (e.g. N-methylpyrrolidone) alkylidene carbonates (e.g., propylene carbonate, etc.), vegetable oil (e.g., soybean oil, cottonseed oil, etc.), vegetable essential oils (e.g., orange oil, hyssop oil, lemon oil, etc.), or water.

A gaseous carrier may include, for example, butane gas, flon gas, liquefied petroleum gas (LPG), dimethyl ether, and carbon dioxide gas.

#### *ix. Adjuvants*

In some instances, the composition provided herein may include an adjuvant. Adjuvants are chemicals that do not possess activity. Adjuvants are either pre-mixed in the formulation or added to the spray tank to improve mixing or application or to enhance performance. They are used extensively in products designed for foliar applications. Adjuvants can be used to customize the formulation to specific needs and compensate for local conditions. Adjuvants may be designed to perform specific functions, including wetting, spreading, sticking, reducing evaporation, reducing volatilization, buffering, emulsifying, dispersing, reducing spray drift, and reducing foaming. No single adjuvant can perform all these functions, but compatible adjuvants often can be combined to perform multiple functions simultaneously.

Among nonlimiting examples of adjuvants included in the formulation are binders, dispersants and stabilizers, specifically, for example, casein, gelatin, polysaccharides (e.g., starch, gum arabic, cellulose derivatives, alginic acid, etc.), lignin derivatives, bentonite, sugars, synthetic water-soluble polymers (e.g., polyvinyl alcohol, polyvinylpyrrolidone, polyacrylic acid, etc.), PAP (acidic isopropyl phosphate), BHT (2,6-di-t-butyl-4-methylphenol), BHA (a mixture of 2-t-butyl-4-methoxyphenol and 3-t-butyl-4-methoxyphenol), vegetable oils, mineral oils, fatty acids and fatty acid esters.

#### *x. Surfactants*

In some instances, the composition provided herein includes a surfactant. Surfactants, also called wetting agents and spreaders, physically alter the surface tension of a spray droplet. For a formulation to perform its function properly, a spray droplet must be able to wet the foliage and spread out evenly over a leaf. Surfactants enlarge the area of formulation coverage, thereby increasing the pest's exposure to the chemical. Surfactants are particularly important when applying a formulation to waxy or hairy leaves. Without proper wetting and spreading, spray droplets often run off or fail to cover leaf surfaces adequately. Too much surfactant, however, can cause excessive runoff and reduce efficacy.

Surfactants are classified by the way they ionize or split apart into electrically charged atoms or molecules called ions. A surfactant with a negative charge is anionic. One with a positive charge is cationic, and one with no electrical charge is nonionic. Formulation activity in the presence of a nonionic surfactant can be quite different from activity in the presence of a cationic or anionic surfactant. Selecting the wrong surfactant can reduce the efficacy of a pesticide product and injure the target plant. Anionic surfactants are most effective when used with contact pesticides (pesticides that control the pest by direct

contact rather than being absorbed systemically). Cationic surfactants should never be used as stand-alone surfactants because they usually are phytotoxic.

Nonionic surfactants, often used with systemic pesticides, help pesticide sprays penetrate plant cuticles. Nonionic surfactants are compatible with most pesticides, and most EPA-registered pesticides that require a surfactant recommend a nonionic type. Adjuvants include, but are not limited to, stickers, extenders, plant penetrants, compatibility agents, buffers or pH modifiers, drift control additives, defoaming agents, and thickeners.

Among nonlimiting examples of surfactants included in the compositions described herein are alkyl sulfate ester salts, alkyl sulfonates, alkyl aryl sulfonates, alkyl aryl ethers and polyoxyethylenated products thereof, polyethylene glycol ethers, polyvalent alcohol esters and sugar alcohol derivatives.

#### *xi. Combinations*

In formulations and in the use forms prepared from these formulations, the modulating agent may be in a mixture with other active compounds, such as pesticidal agents (e.g., insecticides, sterilants, acaricides, nematocides, molluscicides, or fungicides; e.g., a pesticide listed in Table 11), attractants, growth-regulating substances, or herbicides. As used herein, the term “pesticidal agent” refers to any substance or mixture of substances intended for preventing, destroying, repelling, or mitigating any pest. A pesticide can be a chemical substance or biological agent used against pests including insects, mollusks, pathogens, weeds, nematodes, and microbes that compete with humans for food, destroy property, spread disease, or are a nuisance. The term “pesticidal agent” may further encompass other bioactive molecules such as antibiotics, antivirals pesticides, antifungals, antihelminthics, nutrients, pollen, sucrose, and/or agents that stun or slow insect movement.

In instances where the modulating agent is applied to plants, a mixture with other known compounds, such as herbicides, fertilizers, growth regulators, safeners, semiochemicals, or else with agents for improving plant properties is also possible.

### **V. Delivery**

A host described herein can be exposed to any of the compositions described herein in any suitable manner that permits delivering or administering the composition to the host invertebrate (e.g., insect, mollusk, or nematode). The modulating agent may be delivered either alone or in combination with other active or inactive substances and may be applied by, for example, spraying, microinjection, through plants, pouring, dipping, in the form of concentrated liquids, gels, solutions, suspensions, sprays, powders, pellets, briquettes, bricks and the like, formulated to deliver an effective concentration of the modulating agent. Amounts and locations for application of the compositions described herein are generally determined by the habits of the host, the lifecycle stage at which the microorganisms of the host can be targeted by the modulating agent, the site where the application is to be made, and the physical and functional characteristics of the modulating agent. The modulating agents described herein may be administered to the host invertebrate (e.g., insect, mollusk, or nematode) by oral ingestion, but may also be administered by means which permit penetration through the cuticle or penetration of the host (e.g., insect, mollusk, or nematode) respiratory system.

In some instances, the invertebrate host (e.g., insect, mollusk, or nematode) can be simply “soaked” or “sprayed” with a solution including the modulating agent. Alternatively, the modulating agent can be linked to a food component (e.g., comestible) of the invertebrate host (e.g., insect, mollusk, or nematode) for ease of delivery and/or in order to increase uptake of the modulating agent by the host.

5 Methods for oral introduction include, for example, directly mixing a modulating agent with the host’s food, spraying the modulating agent in the host’s habitat or field, as well as engineered approaches in which a species that is used as food is engineered to express a modulating agent, then fed to the host to be affected. In some instances, for example, the modulating agent composition can be incorporated into, or overlaid on the top of, the host’s diet. For example, the modulating agent composition can be sprayed  
10 onto a field of crops which a host inhabits.

In some instances, the composition is sprayed directly onto a plant e.g., crops, by e.g., backpack spraying, aerial spraying, crop spraying/dusting etc. In instances where the modulating agent is delivered to a plant, the plant receiving the modulating agent may be at any stage of plant growth. For example, formulated modulating agents can be applied as a seed-coating or root treatment in early stages of plant  
15 growth or as a total plant treatment at later stages of the crop cycle. In some instances, the modulating agent may be applied as a topical agent to a plant, such that the host invertebrate (e.g., insect, mollusk, or nematode) ingests or otherwise comes in contact with the plant upon interacting with the plant.

Further, the modulating agent may be applied (e.g., in the soil in which a plant grows, or in the water that is used to water the plant) as a systemic agent that is absorbed and distributed through the  
20 tissues of a plant or animal host, such that a host invertebrate (e.g., insect, mollusk, or nematode) feeding thereon will obtain an effective dose of the modulating agent. In some instances, plants or food organisms may be genetically transformed to express the modulating agent such that a host feeding upon the plant or food organism will ingest the modulating agent.

Delayed or continuous release can also be accomplished by coating the modulating agent or a  
25 composition with the modulating agent(s) with a dissolvable or bioerodable coating layer, such as gelatin, which coating dissolves or erodes in the environment of use, to then make the modulating agent available, or by dispersing the agent in a dissolvable or erodable matrix. Such continuous release and/or dispensing means devices may be advantageously employed to consistently maintain an effective concentration of one or more of the modulating agents described herein in a specific host habitat.

30 Alternatively, the modulating agent is expressed in a bacterial or fungal cell and the bacterial or fungal cell is taken up or eaten by the host invertebrate (e.g., insect, mollusk, or nematode) species. Bacteria or fungi can be engineered to produce any of the modulating agents described herein. In other instances, a virus such as a baculovirus which specifically infects host invertebrates (e.g., insect, mollusk, or nematode) may also be used. This ensures safety for mammals, especially humans and animals,  
35 since the virus will not infect mammals.

The modulating agent can also be incorporated into the medium in which the host invertebrate (e.g., insect, mollusk, or nematode) grows, lives, reproduces, feeds, or infests. For example, a modulating agent can be incorporated into a food container, feeding station, protective wrapping, or a hive. For some applications the modulating agent may be bound to a solid support for application in  
40 powder form or in a “trap” or “feeding station.” As an example, for applications where the composition is

to be used in a trap or as bait for a particular host invertebrate (e.g., insect, mollusk, or nematode), the compositions may also be bound to a solid support or encapsulated in a time-release material. For example, the compositions described herein can be administered by delivering the composition to at least one habitat where an agricultural pest (e.g., aphid) grows, lives, reproduces, or feeds.

5 *i. Engineered Plants*

The terms “genetically engineered plant” or “transgenic plant” refer to a plant cell or a plant that expresses a modulating agent. The transgenic plants are also meant to include progeny (decendent, offspring, etc.) of any generation of such a transgenic plant or a seed of any generation of all such transgenic plants wherein said progeny or seed includes a modulating agent.

10 The skilled artisan will recognize that a wide variety of transformation techniques exist in the art, and new techniques are continually becoming available. Any technique that is suitable for the target host plant can be employed. For example, the constructs can be introduced in a variety of forms including, but not limited to as a strand of DNA, in a plasmid, or in an artificial chromosome. The introduction of the constructs into the target plant cells can be accomplished by a variety of techniques, including, but not  
15 limited to *Agrobacterium*-mediated transformation, electroporation, microinjection, microprojectile bombardment calcium-phosphate-DNA co-precipitation or liposome-mediated transformation of a heterologous nucleic acid. The transformation of the plant is preferably permanent, i.e. by integration of the introduced expression constructs into the host plant genome, so that the introduced constructs are passed onto successive plant generations.

20 Any plant species may be transformed to create a transgenic plant. The transgenic plant may be a dicotyledonous plant or a mono- cotyledonous plant. For example and without limitation, transgenic plants of the compositions and methods described herein may be derived from any of the following dicotyledonous plant families: Leguminosae, including plants such as pea, alfalfa and soybean; Umbelliferae, including plants such as carrot and celery; Solanaceae, including the plants such as  
25 tomato, potato, aubergine, tobacco, and pepper; Cruciferae, particularly the genus Brassica, which includes plant such as oilseed rape, beet, cabbage, cauliflower and broccoli); and *Arabidopsis thaliana*; Compositae, which includes plants such as lettuce; Malvaceae, which includes cotton; Fabaceae, which includes plants such as peanut, and the like. Transgenic plants of the invention may be derived from monocotyle- donous plants, such as, for example, wheat, barley, sorghum, millet, rye, triticale, maize,  
30 rice, oats, switchgrass, miscanthus, and sugarcane. Transgenic plants of the invention are also embodied as trees such as apple, pear, quince, plum, cherry, peach, nectarine, apricot, papaya, mango, and other woody species including coniferous and deciduous trees such as poplar, pine, sequoia, cedar, oak, willow, and the like.

Any promoter capable of driving expression in the plant of interest may be used. The promoter  
35 may be native or analogous or foreign or heterologous to the plant host. The choice of promoters to be included depends upon several factors, including, but not limited to, efficiency, selectability, inducibility, desired expression level, and cell- or tissue-preferential expression. It is a routine matter for one of skill in the art to modulate the expression of a sequence by appropriately selecting and positioning promoters and other regulatory regions relative to that sequence.

Promoters active in photosynthetic tissue in order to drive transcription in green tissues such as leaves and stems are of particular interest. Most suitable are promoters that drive expression only or predominantly in such tissues. The promoter may confer expression constitutively throughout the plant, or differentially with respect to the green tissues, or differentially with respect to the developmental stage of the green tissue in which expression occurs, or in response to external stimuli.

Examples of such promoters include the ribulose-1,5-bisphosphate carboxylase (RbcS) promoters such as the RbcS promoter from eastern larch (*Larix laricina*), the pine cab6 promoter (Yamamoto et al., *Plant Cell Physiol.* 35:773-778, 1994), the Cab-1 gene promoter from wheat (Fejes et al., *Plant Mol. Biol.* 15:921-932, 1990), the CAB-1 promoter from spinach (Lubberstedt et al., *Plant Physiol.* 104:997-1006, 1994), the cab1 R promoter from rice (Luan et al., *Plant Cell* 4:971-981, 1992), the pyruvate orthophosphate dikinase (PPDK) promoter from corn (Matsuoka et al., *Proc. Natl. Acad. Sci. USA* 90:9586-9590, 1993), the tobacco Lhcb1\*2 promoter (Cerdan et al. (1997) *Plant Mol. Biol.* 33:245-255), the *Arabidopsis thaliana* SUC2 sucrose-H<sup>+</sup> symporter promoter (Truernit et al. *Planta* 196:564-570, 1995), and thylakoid membrane protein promoters from spinach (psaD, psaF, psaE, PC, FNR, atpC, atpD, cab, rbcS. Other promoters that drive transcription in stems, leaves and green tissue are described in U.S. Patent Publication No. 2007/0006346. The TrpA promoter is a pith preferred promoter and has been described in U.S. Pat. No. 6,018,104.

A maize gene encoding phosphoenol carboxylase (PEPC) has been described by Hudspeth et al. (*Plant Molec. Biol.* 12:579-589, 1989). Using standard molecular biological techniques the promoter for this gene can be used to drive the expression of any gene in a green tissue-specific manner in transgenic plants.

In some other instances, inducible promoters may be desired. Inducible promoters drive transcription in response to external stimuli such as chemical agents or environmental stimuli. For example, inducible promoters can confer transcription in response to hormones such as gibberellic acid or ethylene, or in response to light or drought.

## VI. Screening

Any of the modulating agents described herein may be isolated from a screening assay, wherein a library of modulating agents (e.g., a mixture of variants of a starting modulating agent) is screened for modulating agents (e.g., modulating agent variants) that are effective to alter the microbiota of a host (e.g., insect/mollusk/nematode) and thereby modulate (e.g., increase or decrease) host fitness.

For example, the screening assays provided herein may be effective to identify one or more modulating agents (e.g., a polypeptide, nucleic acid, small molecule, or combinations thereof) that target symbiotic microorganisms resident in the host and thereby decrease the fitness of the host. For example, the identified modulating agent (e.g., a polypeptide, nucleic acid, small molecule, or combinations thereof) may be effective to decrease the viability of pesticide- or allelochemical-degrading microorganisms (e.g., bacteria e.g., a bacterium that degrades a pesticide listed in Table 11), thereby increasing the host's sensitivity to a pesticide (e.g., sensitivity to a pesticide listed in Table 11) or allelochemical agent.

Alternatively, a screening assay may be used to identify a modulating agent effective to increase host fitness (e.g., insect, mollusk, or nematode fitness). For example, the screening assay may be used

to identify one or more modulating agents that target specific microorganisms and/or specific hosts.

Further, the screening assays may be used to identify modulating agents that provide one or more microorganisms with enhanced functionalities. For example, the screening assay may be effective to isolate modulating agents that provide one or more microorganisms with an enhanced ability to

5 metabolize (e.g., degrade) a pesticide (e.g., insecticide, e.g., neonicotinoid) or plant allelochemical (e.g., caffeine, soyacystatin, fenitrothion, monoterpenes, diterpene acids, or phenolic compounds (e.g., tannins, flavonoids)). Delivery and colonization of an isolated microorganism in the host may increase the host's resistance to the pesticide or plant allelochemical, thereby increasing host fitness. The methods may also be useful for the isolation of modulating agents that provide microorganisms with an enhanced ability to

10 colonize any of the hosts described herein.

**Table 11: Pesticides**

|                             |                                 |                   |
|-----------------------------|---------------------------------|-------------------|
| Aclonifen                   | Fenchlorazole-ethyl             | Pendimethalin     |
| Acetamiprid                 | Fenothiocarb                    | Penflufen         |
| Alanycarb                   | Fentirothion                    | Penflufen         |
| Amidosulfuron               | Fenpropidin                     | Pentachlorbenzene |
| Aminocyclopyrachlor         | Fluazolate                      | Penthiopyrad      |
| Amisulbrom                  | Flufenoxuron                    | Penthiopyrad      |
| Anthraquinone               | Flumetralin                     | Pirimiphos-methyl |
| Asulam, sodium salt         | Fluxapyroxad                    | Prallethrin       |
| Benfuracarb                 | Fuberidazole                    | Profenofos        |
| Bensulide                   | Glufosinate-ammonium            | Proquinazid       |
| beta-HCH; beta-BCH          | Glyphosate                      | Prothiofos        |
| Bioresmethrin               | Group: Borax, borate salts (see | Pyraclofos        |
| Blasticidin-S               | Group: Paraffin oils, Mineral   | Pyrazachlor       |
| Borax; disodium tetraborate | Halfenprox                      | Pyrazophos        |
| Boric acid                  | Imiprothrin                     | Pyridaben         |
| Bromoxynil heptanoate       | Imidacloprid                    | Pyridalyl         |
| Bromoxynil octanoate        | Ipconazole                      | Pyridiphenthion   |
| Carbosulfan                 | Isopyrazam                      | Pyrifenox         |
| Chlorantraniliprole         | Isopyrazam                      | Quinmerac         |
| Chlordimeform               | Lenacil                         | Rotenone          |
| Chlorfluazuron              | Magnesium phosphide             | Sedaxane          |
| Chlorphropham               | Metaflumizone                   | Sedaxane          |
| Climbazole                  | Metazachlor                     | Silafluofen       |
| Clopyralid                  | Metazachlor                     | Sintofen          |
| Copper (II) hydroxide       | Metobromuron                    | Spinetoram        |
| Cyflufenamid                | Metoxuron                       | Sulfoxaflor       |
| Cyhalothrin                 | Metsulfuron-methyl              | Temephos          |
| Cyhalothrin, gamma          | Milbemectin                     | thiocloprid       |
| Decahydrate                 | Naled                           | Thiamethoxam      |
| Diafenthiuron               | Napropamide                     | Tolfenpyrad       |

|                   |                |                       |
|-------------------|----------------|-----------------------|
| Dimefuron         | Nicosulfuron   | Tralomethrin          |
| Dimoxystrobin     | Nitenpyram     | Tributyltin compounds |
| Dinotefuran       | Nitrobenzene   | Tridiphane            |
| Diquat dichloride | o-phenylphenol | Triflumizole          |
| Dithianon         | oils           | Validamycin           |
| E-Phosphamidon    | Oxadiargyl     | Zinc phosphide        |
| EPTC              | Oxycarboxin    |                       |
| Ethaboxam         | Paraffin oil   |                       |
| Ethirimol         | Penconazole    |                       |

### EXAMPLES

The following is an example of the methods of the invention. It is understood that various other embodiments may be practiced, given the general description provided above.

5

#### **Example 1: Generation of a cDNA library from corn leaf aphid larvae (*Rhopalosiphum maidis*).**

This Example demonstrates the production of a cDNA library from corn leaf aphid larvae (*Rhopalosiphum maidis*).

#### 10 *Experimental design:*

To generate the library, RNA from 0.9 g whole first-instar larvae (4 to 5 days post-hatch; held at 16°C) is purified using the following phenol/TRI REAGENT®-based method (MOLECULAR RESEARCH CENTER, Cincinnati, Ohio). Larvae are homogenized at room temperature in a 15 mL homogenizer with 10 mL of TRI REAGENT® until a homogenous suspension is obtained. Following 5 min incubation at room temperature, the homogenate is dispensed into 1.5 mL microfuge tubes (1 mL per tube), 200 µL of chloroform is added, and the mixture is vigorously shaken for 15 seconds. After allowing the extraction to sit at room temperature for 10 min, the phases are separated by centrifugation at 12,000×g at 4° C. The upper phase (including about 0.6 mL) is carefully transferred into another sterile 1.5 mL tube, and an equal volume of room temperature isopropanol is added. After incubation at room temperature for 5 to 10 min, the mixture is centrifuged 8 min at 12,000×g (4°C or 25°C).

The supernatant is carefully removed and discarded, and the RNA pellet is washed twice by vortexing with 75% ethanol, with recovery by centrifugation for 5 min at 7,500×g (4°C or 25°C) after each wash. The ethanol is carefully removed, the pellet is allowed to air-dry for 3 to 5 min, and then is dissolved in nuclease-free sterile water. RNA concentration is determined by measuring the absorbance (A) at 260 nm and 280 nm. The RNA extracted is stored at -80°C until further processed, and RNA quality is determined by running an aliquot through a 1% agarose gel.

The larval total RNA is converted into a cDNA library using random priming. The larval cDNA library is sequenced at ½ plate scale by GS FLX 454 Titanium™ series chemistry at EUROFINs MWG Operon, which results in over 600,000 reads with an average read length of 348 bp. 350,000 reads are assembled into over 50,000 contigs. Both the unassembled reads and the contigs are converted into BLASTable databases using the publicly available program, FORMATDB (available from NCBI).

30

**Example 2: Production and purification of *Bcr1* dsRNA**

This Example demonstrates the production and purification of a synthetic dsRNA from a cDNA library.

*Experimental design:*

5 *Bcr1* gene (ACYPI32128) is an essential gene for bacteriocyte regulation and function in insects. *Bcr1* cDNA is prepared from the larval total RNA described in Example 1 and for *Bcr1* dsRNA synthesis prepared by PCR using the primer pairs: Forward 5'-aaactgctgcatggctttct-3' (SEQ ID NO: 90) and reverse 5'-acaggcctttcaggctttta-3' (SEQ ID NO: 91). For the target gene region, two separate PCR amplifications are performed. The first PCR amplification introduces a T7 promoter sequence at the 5' end  
10 ((TTAATACGACTCACTATAGGGAGA; SEQ ID NO: 92) of the amplified sense strands. The second reaction incorporates the T7 promoter sequence at the 5' ends of the antisense strands. The two PCR amplified fragments for each region of the target *Bcr1* gene are then mixed in equal amounts, and the mixture is used as a transcription template for dsRNA production. Double-stranded RNA for insect bioassay is synthesized and purified using an AMBION® MEGASCRIP® RNAi kit following the  
15 manufacturer's instructions (INVITROGEN) or HiScribe® T7 *In Vitro* Transcription Kit following the manufacturer's instructions (New England Biolabs, Ipswich, Mass.). The concentration of *Bcr1* dsRNA is measured using a NANODROP™ 8000 spectrophotometer (THERMO SCIENTIFIC, Wilmington, Del.) and the purified *Bcr1* dsRNA is prepared in TE buffer.

*Bcr1* dsRNA hairpin sequence of one strand:

20 AUGAAACUGCUGCAUGGCUUUCUGAUUAUUAUGCUGACCAUGCAUCUGAGCAUUCAGUAUGCGU  
AUGGCGGCCCGUUUCUGACCAAUAUCUGUGCGAUCGCGUGUGCCAUAACUGUGCGGCGAUGA  
AUUUGUGUGCAGCUGCAUUCAGUAUAAAAGCCUGAAAGGCCUGUGGUUUCGCAUUGCCCGACC  
GGCAAAGCGAGCGUGGUGCUGCAUAACUUUCUGACCAGCCCGUUUUUUUUUUCGGGCUGGUCAG  
AAAGUUAUGCAGCACCACGCUCGCUUUGCCGGUCGGGCAUUGCGGAACACAGGCCUUUCAGG  
25 CUUUUAUACUGAAUGCAGCUGCACACAAAUUCAUCGCCGCACAGUUUAUGGCACACGCGAUCGCA  
CAGAUUUUGGUCAGAAACGGGCGCCAUACGCAUACUGAAUGCUCAGAUUGCAUGGUCAGCAUAA  
UAAUCAGAAAGCCAUGCAGCAGUUUCAU (SEQ ID NO: 93)

**Example 3: Treatment of aphids (*Rhopalosiphum maidis*) with *Bcr1* dsRNA**

30 This Example demonstrates the ability to kill or decrease the fitness of aphids, *Rhopalosiphum maidis*, through treatment with a dsRNA solution by targeting expression of the *Bcr1* gene (ACYPI32128), which is an essential gene for bacteriocyte regulation and function in insects.

Aphids are one of the most important agricultural insect pests. They cause direct feeding damage to crops and serve as vectors of plant viruses. In addition, aphid honeydew promotes the growth  
35 of sooty mold and attracts nuisance ants. The use of chemical treatments, unfortunately still widespread, leads to the selection of resistant individuals whose eradication becomes increasingly difficult.

*Therapeutic design:* dsRNA solutions are formulated with 0 (negative control), 0.5, 1, or 5 µg/ml of *Bcr1* dsRNA from Example 2 in 10 mL of TE buffer with 0.5% sucrose and essential amino acids.

40 *Experimental design:*

To prepare for the treatment, aphids are grown in a lab environment and medium. In a climate-controlled room (16 h light photoperiod; 60±5% RH; 20±2 °C), fava bean plants are grown in a mixture of vermiculite and perlite and are infested with aphids. To limit maternal effects or health differences between plants, 5–10 adults from different plants are distributed among 10 two-week-old plants, and allowed to multiply to high density for 5–7 days. For experiments, second and third instar aphids are collected from healthy plants and divided into treatments so that each treatment receives approximately the same number of individuals from each of the collection plants.

Wells of a 96-well plate are filled with 200 µl of artificial aphid diet (Febvay et al., *Canadian Journal of Zoology* 66(11):2449-2453, 1988) and the plate is covered with parafilm to make a feeding sachet. Artificial diet is either mixed with the solution of TE buffer (Tris HCl (1 mM) plus EDTA (0.1 mM) buffer, pH 7.2.) with 0.5% sucrose and essential amino acids only as a negative control, or mixed with dsRNA solutions diluted in TE buffer containing varying concentrations of dsRNA. dsRNA solutions are mixed with artificial diet to obtain final concentrations between 0.5 to 5 µg/ml.

For each replicate treatment, 30–50 second and third instar aphids are placed individually in wells of a 96-well plate and a feeding sachet plate is inverted above them, allowing the insects to feed through the parafilm and keeping them restricted to individual wells. Experimental aphids are kept under the same environmental conditions as aphid colonies. After the aphids are fed for 24 hr, the feeding sachet is replaced with a new one containing sterile artificial diet and a new sterile sachet is provided every 24 h for 4 days. At the time that the sachet is replaced, aphids are also checked for mortality. An aphid is counted as dead if it had turned brown or is at the bottom of the well and does not move during the observation. If an aphid is on the parafilm of the feeding sachet but not moving, it is assumed to be feeding and alive.

The survival rates of aphids treated with *Bcr1* dsRNA are compared to the aphids treated with the negative control. The survival rate of aphids treated with *Bcr1* dsRNA is decreased as compared to the control treated aphids.

#### **Example 4: Production of transgenic grass expressing *Bcr1* dsRNAs**

This Example demonstrates genetic modification and production of *Bcr1* dsRNA in a transgenic grass for delivery to aphids.

Transgenic forage grass blue grama, *Bouteloua gracilis*, that produces the *Bcr1* dsRNA molecules, through expression of a chimeric gene stably-integrated into the plant genome, is produced by microprojectile bombardment, using a system based on the highly chlorophyllous and embryogenic cell line 'TIANSJ98' (Aguado-Santacruz et al., *Theoretical and Applied Genetics* 104(5):763-771, 2002).

*'TIANSJ98' cell line establishment and maintenance:* The embryogenic, highly chlorophyllous 'TIANSJ98' cell line is obtained from culturing shoot apice-derived green calli in liquid MPC medium as described in (Aguado-Santacruz et al., *Plant Cell Rep.* 20:131-136, 2001). This cell line is subcultured every 20 days, transferring 1 ml of the cell suspension into 24 ml of fresh MPC medium. The reasons for utilizing the finely dispersed condition of the embryogenic calli are 1) to synchronize the physiological stage of the target cells, 2) to maximize the distribution of the totipotent material on the paper filters

(optimizing the shoot cover of the bombarded plasmids), and 3) to facilitate the identification of independent transformation events (green spots) within the dispersed cell clusters under selection.

*Microprojectile bombardment of embryogenic cells:*

5 A binary transformation vector is used with the template fragment for *Bcr1* dsRNA expression. This plasmid contains the inverted repeat of the target *Bcr1* gene under the control of a double 35S Cauliflower Mosaic Virus promoter, and a leader sequence from Alfalfa Mosaic Virus (Aguado-Santacruz et al., *Theoretical and Applied Genetics* 104(5):763-771, 2002).

10 The highly chlorophyllous embryogenic cell line 'TIAN SJ98' is used as the target for the microprojectile delivery experiments. The cells are distributed onto 2.0-cm diameter paper-filter disks (approximately 2 g FW cells). Bombardment mixtures are as follows: 50  $\mu$ L of M10 tungsten particles (15 mg/ml), 10  $\mu$ L of DNA (1  $\mu$ g/ml), 50  $\mu$ L of 2.5 M  $\text{CaCl}_2$  and 20  $\mu$ L of 0.1 M spermidine are mixed in sequential order, vortexed for 5 min and then briefly sonicated. The mixture is centrifuged at 10,000 rpm for 10 s. 60  $\mu$ L of the supernatant are removed and the rest is dispensed into 5- $\mu$ L aliquots for individual  
15 shoots. Bombardments are performed using the Particle Inflow Gun (Finer et al., *Plant Cell Rep.* 11:323-328, 1992). The particle/DNA mixture is placed in the center of the syringe filter unit. Embryogenic cells are covered with a 500- $\mu$ m baffle, placed at a distance of 10 cm from the screen filter unit containing the particles, and bombarded once in the vacuum chamber at 60 mmHg. Two different osmotic media for pre- and post-bombardment treatments (0.4 and 1 M mannitol supplied in solidified MPC medium) and  
20 three bombardment pressures (60, 80 and 100 PSI) are tested. Pre-bombardment treatment is applied 24 hr before shooting. After discharge, the paper filters supporting the embryogenic cells are maintained for 3-days more on the same osmotic medium used in the pre-bombardment treatment. Thus, a total of nine treatments are evaluated with ten dishes bombarded per treatment. As a control, filters with suspension material are bombarded using particles without DNA.

25  
*Selection of stable transformed clones and recovery of plants:*  
After the 3-day post-bombardment osmotic treatment on MPC medium containing 0.4 or 1 M mannitol, but lacking antibiotic, the paper filter disks supporting the bombarded cells are transferred onto MPC medium containing 140 mg/l of kanamycin and incubated at  $30 \pm 1$  °C in white light provided by cold  
30 fluorescent lamps. The same procedure is followed for cells bombarded but not subjected to osmotic treatment. The kanamycin concentration is raised 2-months later to 150 or 160 mg/l. The cells are sub cultured every 3 weeks and maintained for 8 months in selection. After this period, kanamycin-resistant clones are transferred to regeneration medium containing full-strength MS medium, 3% sucrose, 2.5% phytigel (Sigma, St. Louis, Missouri) but no antibiotic. The regenerated shoots are transferred for rooting  
35 to 1/2 MS containing 3.0  $\mu$ M (0.56 mg/l)  $\alpha$ -naphthaleneacetic acid, 2.5  $\mu$ M (0.51 mg/l) indole-3-butyric acid and 2.5% phytigel, and incubated at  $30 \pm 1$  °C under continuous fluorescent light. Later, rooted plantlets are transferred to pots, hardened off, and grown to maturity in a greenhouse.

*PCR analysis for transformation verification:*

Total genomic DNA is prepared from kanamycin-resistant and untransformed control plants using the following protocol: Approximately 250 mg of cells are collected in 2-ml Eppendorf tubes and ground to a fine powder in liquid N<sub>2</sub> using a glass pestle attached to a homogenizer (Caframo, Stirrer type RZR). Powdered cells are re-suspended with 500 µL of extraction buffer (7.0 M urea, 0.35 M NaCl, 0.05 M Tris-HCl, pH 8.0, 0.02 M EDTA, 1% sarcosine) for at least 45 min. The cell homogenate is extracted with 1 vol of phenol/ chloroform. The aqueous phase is separated by centrifugation and then precipitated using an equal volume of isopropyl alcohol. The precipitated DNA is washed once with 70% ethanol and resuspended in TE buffer (0.01 M Tris-HCl, 0.01 M EDTA, pH 8.0).

For PCR analysis, 100 to 150 ng, are used for genomic DNA amplifications, in 25-µL reactions. Primers forward 5'-aaactgctgcatggcttct-3' (SEQ ID NO: 91) and reverse 5'-acaggcctttcaggctttta-3' (SEQ ID NO: 92) are designed for amplifying an internal 169-bp fragment of the *Bcr1* sense anti sense inserted fragment. PCR reactions are carried out using a Perkin Elmer thermocycler for 30 cycles. Reaction temperatures are denaturation 95°C (2 min), annealing 56°C (30 s), and extension 72°C (30 s). The 25-µL reaction volumes contain: 1 × PCR buffer, 0.25 mM of dNTPs, 2 mM MgCl<sub>2</sub>, 0.2 µM of primers and 2.5 u of Taq. The amplification products are analyzed by electrophoreses in 1% agarose/SYBR green gels.

#### Example 5: Production of transgenic grass producing *coIA* bacteriocin

This Example demonstrates genetic modification and production of the bacteriocyte regulatory peptide Coleopteracin A (*coIA*) in a transgenic grass for delivery to aphids.

Transgenic forage grass blue grama, *Bouteloua gracilis*, that produces Coleopteracin A, through expression of a chimeric gene stably-integrated into the plant genome, is produced by microprojectile bombardment, using a system described in Example 4.

The embryogenic, highly chlorophyllous 'TIANSJ98' cell line is subcultured every 20 days, transferring 1 ml of the cell suspension into 24 ml of fresh MPC medium.

Coleopteracin A (*coIA*)

atgacccgcacccatgctgttctggcgtgcgtggcgccgctgtatgtgtgcattagcgcgaccgcgggcaaaccggaagaattgcgaaactgagc  
gatgaagcgccgagcaacgatcaggcgatgtatgaaagcattcagcgcctatcgccgctttgtggatggcaaccgctataacggcggccagcagc  
agcagcagcagccgaaacagtgggaagtgcgcccggatctgagccgcatcagcgcggcaacaccaaagcgcaggtggaaattaacaaaa  
aaggcgataaccatgatattaacgcgggctggggcaaaaacattaacggcccggatagccataaagatacctggcatgtgggcggcagcgtgc  
gctgg (SEQ ID NO: 94)

A transformation plasmid is constructed for Coleopteracin A (*coIA*) expression. The plasmid contains the nucleic acid for *coIA* under the control of a double 35S Cauliflower Mosaic Virus promoter, and a leader sequence from Alfalfa Mosaic Virus (Aguado-Santacruz et al., *Theoretical and Applied Genetics* 104(5):763-771, 2002).

The highly chlorophyllous embryogenic cell line 'TIANSJ98' is used as the target for the microprojectile delivery experiments. The cells are distributed onto 2.0-cm diameter paper-filter disks (approximately 2 g FW cells). Bombardment mixtures are as follows: 50 µL of M10 tungsten particles (15 mg/ml), 10 µL of DNA (1 µg/ml), 50 µL of 2.5 M CaCl<sub>2</sub> and 20 µL of 0.1 M spermidine are mixed in sequential order, vortexed for 5 min and then briefly sonicated. The mixture is centrifuged at 10,000 rpm for 10 s. 60 µL of the supernatant are removed and the rest is dispensed into 5-µL aliquots for individual

shoots. Bombardments are performed using the Particle Inflow Gun (Finer et al., *Plant Cell Rep.* 11:323-328, 1992). The particle/DNA mixture is placed in the center of the syringe filter unit. Embryogenic cells are covered with a 500- $\mu$ m baffle, placed at a distance of 10 cm from the screen filter unit containing the particles, and bombarded once in the vacuum chamber at 60 mmHg. Two different osmotic media for pre- and post-bombardment treatments (0.4 and 1 M mannitol supplied in solidified MPC medium) and three bombardment pressures (60, 80 and 100 PSI) are tested. Pre-bombardment treatment is applied 24 hr before shooting. After discharge, the paper filters supporting the embryogenic cells are maintained for 3-days more on the same osmotic medium used in the pre-bombardment treatment. Thus, a total of nine treatments are evaluated with ten dishes bombarded per treatment. As a control, filters with suspension material are bombarded using particles without DNA.

*Selection of stable transformed clones and recovery of plants:*

After the 3-day post-bombardment osmotic treatment on MPC medium containing 0.4 or 1 M mannitol, but lacking antibiotic, the paper filter disks supporting the bombarded cells are transferred onto MPC medium containing 140 mg/l of kanamycin and incubated at  $30 \pm 1$  °C in white light provided by cold fluorescent lamps. The same procedure is followed for cells bombarded but not subjected to osmotic treatment. The kanamycin concentration is raised 2-months later to 150 or 160 mg/l. The cells are sub cultured every 3 weeks and maintained for 8 months in selection. After this period, kanamycin-resistant clones are transferred to regeneration medium containing full-strength MS medium, 3% sucrose, 2.5% phytigel (Sigma, St. Louis, Missouri) but no antibiotic. The regenerated shoots are transferred for rooting to 1/2 MS containing 3.0  $\mu$ M (0.56 mg/l)  $\alpha$ -naphthaleneacetic acid, 2.5  $\mu$ M (0.51 mg/l) indole-3-butyric acid and 2.5% phytigel, and incubated at  $30 \pm 1$  °C under continuous fluorescent light. Later, rooted plantlets are transferred to pots, hardened off, and grown to maturity in a greenhouse.

*PCR analysis for transformation verification:*

Total genomic DNA is prepared from kanamycin-resistant and untransformed control plants using the following protocol: Approximately 250 mg of cells are collected in 2-ml Eppendorf tubes and ground to a fine powder in liquid N<sub>2</sub> using a glass pestle attached to a homogenizer (Caframo, Stirrer type RZR). Powdered cells are re-suspended with 500  $\mu$ L of extraction buffer (7.0 M urea, 0.35 M NaCl, 0.05 M Tris-HCl, pH 8.0, 0.02 M EDTA, 1% sarcosine) for at least 45 min. The cell homogenate is extracted with 1 vol of phenol/ chloroform. The aqueous phase is separated by centrifugation and then precipitated using an equal volume of isopropyl alcohol. The precipitated DNA is washed once with 70% ethanol and resuspended in TE buffer (0.01 M Tris-HCl, 0.01 M EDTA, pH 8.0).

For PCR analysis, 100 to 150 ng, are used for genomic DNA amplifications, in 25- $\mu$ L reactions. Primers forward 5'-caacgatcaggcgatgtatg-3' (SEQ ID NO: 95) and reverse 5'-ttaattccacctgcgcttt-3' (SEQ ID NO: 96) are designed for amplifying an internal 165-bp fragment of the *co/A* gene inserted fragment. PCR reactions are carried out using a Perkin Elmer thermocycler for 30 cycles. Reaction temperatures are denaturation 95°C (2 min), annealing 56°C (30 s), and extension 72°C (30 s). The 25- $\mu$ L reaction

volumes contain: 1 × PCR buffer, 0.25 mM of dNTPs, 2 mM MgCl<sub>2</sub>, 0.2 μM of primers and 2.5 u of Taq. The amplification products are analyzed by electrophoreses in 1% agarose/SYBR green gels.

### Example 6: Production of transgenic grass producing *colA* bacteriocin

5 This example demonstrates genetic modification and production of *colA* bacteriocin in a transgenic grass for delivery to aphids.

Transgenic forage grass blue grama, *Bouteloua gracilis*, that produces *colA* bacteriocin, through expression of a chimeric gene stably-integrated into the plant genome, is produced by microprojectile bombardment, using a system based on the highly chlorophyllous and embryogenic cell line 'TIANSJ98' (Aguado- Santacruz et al., 2002. Theoretical and Applied Genetics, 104(5), 763–771).

10 'TIANSJ98' cell line establishment and maintenance: The embryogenic, highly chlorophyllous 'TIANSJ98' cell line is obtained from culturing shoot apice-derived green calli in liquid MPC medium as described in (Aguado- Santacruz et al., 2001. ex Steud. Plant Cell Rep 20:131–136). This cell line is subcultured every 20 days, transferring 1 ml of the cell suspension into 24 ml of fresh MPC medium. The reasons for utilizing the finely dispersed condition of the embryogenic calli are 1) to synchronize the physiological stage of the target cells, 2) to maximize the distribution of the totipotent material on the paper filters (optimizing the shoot cover of the bombarded plasmids), and 3) to facilitate the identification of independent transformation events (green spots) within the dispersed cell clusters under selection.

#### 20 *Microprojectile bombardment of embryogenic cells*

A transformation vector is used with the template fragment for *colA* bacteriocin expression.

*colA* bacteriocin

atgacccgcaccatgctgtttctggcgtgcgtggcggcgctgtatgtgtgcattagcgcgaccgcgggcaaaccggaagaatttgcgaa  
actgagcgatgaagcgccgagcaacgatcaggcgatgtatgaaagcattcagcgcctatcgccgctttgtggatggcaaccgctataacggcggcc  
25 agcagcagcagcagcagccgaaacagtgggaagtgcgcccggatctgagccgcgatcagcgcggcaacaccaaagcgcaggtggaaatta  
acaaaaaaggcgataaccatgatattaacgcgggctggggcaaaaacattaacggcccggatagccataaagatacctggcatgtgggcggca  
gcgtgcgctgg (SEQ ID NO: 94)

This plasmid contains the nucleic acid for *colA* bacteriocin under the control of a double 35S Cauliflower Mosaic Virus promoter, and a leader sequence from Alfalfa Mosaic Virus (Aguado- Santacruz  
30 et al., 2002. Theoretical and Applied Genetics, 104(5), 763–771).

The highly chlorophyllous embryogenic cell line 'TIANSJ98' is used as the target for the microprojectile delivery experiments. The cells are distributed onto 2.0-cm diameter paper-filter disks (approximately 2 g FW cells). Bombardment mixtures are as follows: 50 μL of M10 tungsten particles (15 mg/ml), 10 μL of DNA (1 μg/ml), 50 μL of 2.5 M CaCl<sub>2</sub> and 20 μL of 0.1 M spermidine are mixed in  
35 sequential order, vortexed for 5 min and then briefly sonicated. The mixture is centrifuged at 10,000 rpm for 10 s. 60 μL of the supernatant are removed and the rest is dispensed into 5-μL aliquots for individual shoots.

Bombardments are performed using the Particle Inflow Gun (Finer et al., 1992. Plant Cell Rep 11:323–328). The particle/DNA mixture is placed in the center of the syringe filter unit. Embryogenic  
40 cells are covered with a 500-μm baffle, placed at a distance of 10 cm from the screen filter unit containing

the particles, and bombarded once in the vacuum chamber at 60 mmHg. Two different osmotic media for pre- and post-bombardment treatments (0.4 and 1 M mannitol supplied in solidified MPC medium) and three bombardment pressures (60, 80 and 100 PSI) are tested. Pre-bombardment treatment is applied 24 hr before shooting.

5 After discharge, the paper filters supporting the embryogenic cells are maintained for 3-days more on the same osmotic medium used in the pre-bombardment treatment. Thus, a total of nine treatments are evaluated with ten dishes bombarded per treatment. As a control, filters with suspension material are bombarded using particles without DNA.

#### 10 *Selection of stable transformed clones and recovery of plants*

After the 3-day post-bombardment osmotic treatment on MPC medium with 0.4 or 1 M mannitol, but lacking antibiotic, the paper filter disks supporting the bombarded cells are transferred onto MPC medium containing 140 mg/l of kanamycin and incubated at  $30 \pm 1$  °C in white light provided by cold fluorescent lamps. The same procedure is followed for cells bombarded but not subjected to osmotic treatment. The kanamycin concentration is raised 2-months later to 150 or 160 mg/l. The cells are subcultured every 3 weeks and maintained for 8 months in selection. After this period, kanamycin-resistant clones are transferred to regeneration medium with full-strength MS medium, 3% sucrose, 2.5% phytigel (Sigma, St. Louis, Missouri) but no antibiotic. The regenerated shoots are transferred for rooting to 1/2 MS with 3.0  $\mu$ M (0.56 mg/l)  $\alpha$ -naphthaleneacetic acid, 2.5  $\mu$ M (0.51 mg/l) indole-3-butyric acid and 2.5% phytigel, and incubated at  $30 \pm 1$  °C under continuous fluorescent light. Later, rooted plantlets are transferred to pots, hardened off, and grown to maturity in a greenhouse.

#### *PCR analysis for transformation verification*

Total genomic DNA is prepared from kanamycin-resistant and untransformed control plants using the following protocol: Approximately 250 mg of cells are collected in 2-ml Eppendorf tubes and ground to a fine powder in liquid N<sub>2</sub> using a glass pestle attached to a homogenizer (Caframo, Stirrer type RZR). Powdered cells are re-suspended with 500  $\mu$ L of extraction buffer (7.0 M urea, 0.35 M NaCl, 0.05 M Tris-HCl, pH 8.0, 0.02 M EDTA, 1% sarcosine) for at least 45 min. The cell homogenate is extracted with 1 vol of phenol/ chloroform. The aqueous phase is separated by centrifugation and then precipitated using an equal volume of isopropyl alcohol. The precipitated DNA is washed once with 70% ethanol and resuspended in TE buffer (0.01 M Tris-HCl, 0.01 M EDTA, pH 8.0).

For PCR analysis, 100 to 150 ng, are used for genomic DNA amplifications, in 25- $\mu$ L reactions. Primers forward 5'-caacgatcaggcgatgtatg-3' and reverse 5'-ttaattccacctgcgcttt-3' are designed for amplifying an internal 165-bp fragment of the ColA gene inserted fragment. PCR reactions are carried out using a Perkin Elmer thermocycler for 30 cycles. Reaction temperatures are denaturation 95°C (2 min), annealing 56°C (30 s), and extension 72°C (30 s). The 25- $\mu$ L reaction volumes contain: 1  $\times$  PCR buffer, 0.25 mM of dNTPs, 2 mM MgCl<sub>2</sub>, 0.2  $\mu$ M of primers and 2.5 u of Taq. The amplification products are analyzed by electrophoreses in 1% agarose/SYBR green gels.

**Example 7: Production of a cDNA library from green peach aphid (*Myzus persicae*).**

This Example demonstrates the production of a cDNA library from green peach aphid larvae (*Myzus persicae*).

5 *Total RNA isolation from aphid larvae and cDNA library preparation:*

Total RNA from 0.9 g of whole first-instar larvae (4 to 5 days post-hatch; held at 16°C) is extracted using the following phenol/TRI REAGENT®-based method (MOLECULAR RESEARCH CENTER, Cincinnati, Ohio). Larvae are homogenized at room temperature in a 15 mL homogenizer with 10 mL of TRI REAGENT® until a homogenous suspension is obtained. Following 5 min. incubation at room temperature, the homogenate is dispensed into 1.5 mL microfuge tubes (1 mL per tube), 200 µL of chloroform is added, and the mixture is vigorously shaken for 15 seconds. After allowing the extraction to sit at room temperature for 10 min, the phases are separated by centrifugation at 12,000×g at 4° C. The upper phase (including about 0.6 mL) is carefully transferred into another sterile 1.5 mL tube, and an equal volume of room temperature isopropanol is added. After incubation at room temperature for 5 to 10 min, the mixture is centrifuged 8 min at 12,000×g (4°C or 25°C).

The supernatant is carefully removed and discarded, and the RNA pellet is washed twice by vortexing with 75% ethanol, with recovery by centrifugation for 5 min at 7,500×g (4°C or 25°C) after each wash. The ethanol is carefully removed, the pellet is allowed to air-dry for 3 to 5 min, and then is dissolved in nuclease-free sterile water. The RNA concentration is determined by measuring the absorbance (A) at 260 nm and 280 nm. The RNA is stored at -80°C, and the RNA quality is determined by running an aliquot through a 1% agarose gel.

The larval total RNA is converted into a cDNA library using random priming. The larval cDNA library is sequenced at ½ plate scale by GS FLX 454 Titanium™ series chemistry at EUROFINIS MWG Operon, which results in over 600,000 reads with an average read length of 348 bp. 350,000 reads are assembled into over 50,000 contigs. Both the unassembled reads and the contigs are converted into BLASTable databases using the publicly available program, FORMATDB (available from NCBI).

**Example 8: Production of *Bcr1* dsRNA**

This Example demonstrates the production and purification of a synthetic dsRNA from a cDNA library.

*Experimental design:*

*Bcr1* gene (ACYPI32128) is an essential gene for bacteriocyte regulation and function in insects. The cDNA described in the Example 6 is used for *Bcr1* dsRNA synthesis prepared by PCR using the primer pairs: Forward 5'-aaactgctgcatggcttct-3' (SEQ ID NO: 90) and reverse 5'-acaggccttcaggccttta-3' (SEQ ID NO: 91). For the target gene region, two separate PCR amplifications are performed. The first PCR amplification introduces a T7 promoter sequence at the 5' end ((TTAATACGACTCACTATAGGGAGA; SEQ ID NO: 92) of the amplified sense strands. The second reaction incorporates the T7 promoter sequence at the 5' ends of the antisense strands. The two PCR amplified fragments for each region of the target gene *Bcr1* are then mixed in equal amounts, and the

mixture is used as a transcription template for dsRNA production. Double-stranded RNA for insect bioassay is synthesized and purified using an AMBION® MEGASCRIP™ RNAi kit following the manufacturer's instructions (INVITROGEN) or HiScribe® T7 *In Vitro* Transcription Kit following the manufacturer's instructions (New England Biolabs, Ipswich, Mass.). The concentration of dsRNAs against *Bcr1* is measured using a NANODROP™ 8000 spectrophotometer (THERMO SCIENTIFIC, Wilmington, Del.) and the purified dsRNA molecules are prepared in TE buffer.

*Bcr1* dsRNA hairpin sequence of one strand:

AUGAAACUGCUGCAUGGCUUUCUGAUUUAUUAUGCUGACCAUGCAUCUGAGCAUUCAGUAUGCGU  
AUGGCGGCCCGUUUCUGACCAAUAUCUGUGCGAUCGCGUGUGCCAUAACUGUGCGGCGAUGA  
10 AUUUGUGUGCAGCUGCAUUCAGUAUAAAAGCCUGAAAGGCCUGUGGUUUCGCAUUGCCCGACC  
GGCAAAGCGAGCGUGGUGCUGCAUAACUUUCUGACCAGCCCGUUUUUUUUUUCGGGCUGGUCAG  
AAAGUUAUGCAGCACCACGCUCGCUUUGCCGGUCGGGCAAUGCGGAAACCACAGGCCUUUCAGG  
CUUUUAUACUGAAUGCAGCUGCACACAAUUAUCGCGCACAGUUUAUGGCACACGCGAUCGCA  
CAGAUUUUGGUCAGAAACGGGCGCCAUACGCAUACUGAAUGCUCAGAUGCAUGGUCAGCAUAA  
15 UAAUCAGAAAGCCAUGCAGCAGUUUCAU (SEQ ID NO: 93)

#### **Example 9: Treatment of aphids (*Myzus persicae*) with a solution of *Bcr1* dsRNA**

This Example demonstrates the ability to kill or decrease the fitness of aphids, *Myzus persicae*, by treating them with a dsRNA solution by targeting expression of the *Bcr1* gene (ACYPI32128), which is an essential gene for bacteriocyte regulation and function in insects.

Aphids are one of the most important agricultural insect pests. They cause direct feeding damage to crops and serve as vectors of plant viruses. In addition, aphid honeydew promotes the growth of sooty mold and attracts nuisance ants. The use of chemical treatments, unfortunately still widespread, leads to the selection of resistant individuals whose eradication becomes increasingly difficult.

*Therapeutic design:* dsRNA solutions are formulated with 0 (negative control), 0.5, 1, or 5 µg/mL of *Bcr1* dsRNA from Example 7 in 10 mL of TE buffer with 0.5% sucrose and essential amino acids.

#### *Experimental design:*

To prepare for the treatment, aphids are grown in a lab environment and medium. In a climate-controlled room (16 hr light photoperiod; 60±5% RH; 20±2°C), plants are grown in a mixture of vermiculite and perlite and are infested with aphids. To limit maternal effects or health differences between plants, 5–10 adults from different plants are distributed among 10 two-week-old plants, and allowed to multiply to high density for 5–7 days. For experiments, second and third instar aphids are collected from healthy plants and divided into treatments so that each treatment receives approximately the same number of individuals from each of the collection plants.

Wells of a 96-well plate are filled with 200 µl of artificial aphid diet (Febvay et al., *Canadian Journal of Zoology* 66(11):2449-2453, 1988) and the plate is covered with parafilm to make a feeding sachet. Artificial diet is either mixed with the solution of TE buffer (Tris HCl (1 mM) plus EDTA (0.1 mM) buffer, pH 7.2.) with 0.5% sucrose and essential amino acids only as a negative control, or mixed with

dsRNA solutions diluted in TE buffer containing varying concentrations of dsRNA. dsRNA solutions are mixed with artificial diet to obtain final concentrations between 0.5 to 5 µg/ml.

For each replicate treatment, 30–50 second and third instar aphids are placed individually in wells of a 96-well plate and a feeding sachet plate is inverted above them, allowing the insects to feed through the parafilm and keeping them restricted to individual wells. Experimental aphids are kept under the same environmental conditions as aphid colonies. After the aphids are fed for 24 hr, the feeding sachet is replaced with a new one containing sterile artificial diet and a new sterile sachet is provided every 24 h for 4 days. At the time that the sachet is replaced, aphids are also checked for mortality. An aphid is counted as dead if it had turned brown or is at the bottom of the well and does not move during the observation. If an aphid is on the parafilm of the feeding sachet but not moving, it is assumed to be feeding and alive.

The survival rates of aphids treated with *Bcr1* dsRNA are compared to the aphids treated with the negative control. The survival rate of aphids treated with *Bcr1* dsRNA is decreased as compared to the control treated aphids.

#### **Example 10: Topical-plant delivery of bacteriocyte specific dsRNA for crop protection**

This Example demonstrates the ability to deliver bacteriocyte specific dsRNA by topical application on tobacco plant leaves. dsRNA targets the expression of the *Bcr1* gene (ACYPI32128), which is an essential gene for bacteriocyte regulation and function in insects.

*Therapeutic design:* dsRNA-LDH spray solutions are formulated with *Bcr1* dsRNA produced in Example 7 at 1:0 (negative control), 1:1, 1:2, or 1:3 ratio of dsRNA:LDH at 125 µL/cm<sup>2</sup> of leave surface.

#### *Preparation and characterization of LDH nanosheets:*

Sheet-like clay nanoparticles, specifically positively charged LDH, are excellent nanocarrier systems for dsRNA, to deliver RNAi as stable spray formulations for crop protection (Mitter et al., *Nature Plants* 1-10, 2017). LDH conjugates strongly adhere to the leaf surface even after vigorous rinsing, and enhance the stability of dsRNA for a longer period under environmental conditions. The sustained release of dsRNA is facilitated through the formation of carbonic acid on the leaf surface from atmospheric CO<sub>2</sub> and humidity, which helps degrade the LDH nanosheets.

LDH nanosheets are prepared according to (Mitter et al., *Nature Plants* 1-10, 2017) by modified non-aqueous precipitation, followed by heat treatment, purification and dispersion in water to get an average particle size of 45 nm. The particles (five repeated LDH samples) are characterized by the Nanosizer Nano ZS instrument (Malvern Instruments) to obtain the Z average size and Pdl and imaged using JEOL JSM-2010 TEM22. The chemical composition and crystal structure are verified by powder XRD with five LDH samples and Fourier transform infrared spectroscopy with attenuated total reflection mode, using a Rigaku Miniflex X-Ray diffractometer and a Nicolet 6700 FT-IR (Thermo Electron Corporation), respectively.

#### *dsRNA loading on LDH:*

To define optimal and complete loading of the dsRNA hairpin construct obtained in Example 7 into LDH nanosheets, the ratio of *in vitro* transcribed dsRNA (500 ng) to LDH (dsRNA-LDH (w/w)) is

assayed at 1:1, 1:2, 1:3, 1:4, 1:5 and 1:10 multiple times. dsRNA and LDH are incubated in a total volume of 10  $\mu$ L at room temperature for 30 min with gentle orbital agitation. Complete dsRNA loading is assessed by retention of dsRNA–LDH complexes in the well of a 1% agarose gel. Appropriate loading ratios are consistent, irrespective of the scale-up volume required.

5 Release of dsRNA from dsRNA–LDH complexes and stability of the released dsRNA are tested as described in (Mitter et al., *Nature Plants* 1-10, 2017).

*Northern blot analysis for dsRNA detection within the leaves:*

To detect dsRNA uptake, *N. tabacum* plants (three replicates) are sprayed with either LDH, *Bcr1*-dsRNA or *Bcr1*-dsRNA–LDH at day 0. The ratios of the complex tested are: 1:1, 1:2, or 1:3 ratios of dsRNA-LDH. Plants are grown in UQ23 soil in 10 cm wide pots under glasshouse conditions (average temperature 25°C with natural light). The apex of these plants is covered using a masking tape at the time of the spray. New leaves that emerge 20 days after the spray are collected. Total RNA is extracted by TRIzol extraction and enriched for small RNAs (Mitter et al., *Am. Phytopathological Soc.* 16:936-944, 2003). Small RNAs (20  $\mu$ g) for each treatment are run on a 15% (wt/vol) denaturing urea polyacrylamide gel (PAGE). ZR Small-RNA ladder (Zymo Research) labelled with DIG at the 3' end is used as a marker. The blots are transferred by trans-blot SD semi-dry transfer unit (Bio-Rad) on a Hybond–N membrane (Roche). Blots are processed as per manufacture's recommendation using DIG-labelled *Bcr1* 24 nt probe (5-atgctgacatgcatctgagcatt), proprietary buffer set (Roche). Following hybridization, filters are detected using the CSPD chemiluminescent alkaline phosphatase substrate. The quantification analysis is performed using the NIH Image 1.6 software.

**Example 11: Solid phase synthesis of a PNA**

This Example demonstrates solid phase synthesis of a PNA.

25 *Therapeutic design:* complementary antisense PNA constructs against the bacteriocyte target gene *Bcr1*: *gaatgcagctgc*

*Experimental design:*

The PNA antisense is synthesized automatically (MilliGen 9050 peptide synthesizer) by the solid-phase method using standard Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) chemistry in a continuous flow mode.

PNA antisense purification is performed by reversed-phase high-performance liquid chromatography (RP-HPLC) with UV detection at 260 nm using a semi-prep column C18 (10  $\mu$ m, 300 $\times$ 7.7 mm, Xterra Waters, 300 Å), eluting with water containing 0.1% TFA (eluent A) and acetonitrile containing 0.1% TFA (eluent B); elution gradient: from 100% A to 50% B in 30 min, flow: 4 ml/min. The purity and identity of the purified PNA antisense is examined by ultra-performance liquid chromatography tandem mass-spectrometry (UPLC-MS; Waters Acquity equipped with ESI-Q analyzer) using an Acquity UPLC BEH C18; 2.1 $\times$ 50 MM, 1.7  $\mu$ m column. The expected mass peaks are observed for the correct amino and nucleic acid sequence.

**Example 12: Production of Cy3 labeled PNA**

This Example demonstrates joining the PNA described in Example 10 to Cy3 dye as a marker to tag the PNA through click chemistry.

- 5     *Therapeutic design:* PNA with dibenzylcyclooctyne (DBCO) modification and Cy3 dye with azide modification: *Cy3-gaatgcagctgc*.

*Experimental design:*

10     To prepare for the click reaction the PNA synthesized in Example 10 is labeled with DBCO (Glen Research, Sterling, VA). DBCO-sulfo-NHS ester is dissolved at a concentration of 5.2 mg per 60  $\mu$ L in water or anhydrous DMSO. This stock solution is used to conjugate the amino-modified PNA in sodium carbonate/bicarbonate conjugation buffer, pH $\approx$ 9.

15     For a 0.2  $\mu$ mol synthesis of PNA, PNA is dissolved in 500  $\mu$ L of conjugation buffer. Approx. a 6 fold excess (6  $\mu$ L) of DBCO-sulfo-NHS ester solution is added to the dissolved PNA. The mixture is vortexed and incubated at room temperature for 2-4 hours up to about overnight. The conjugated PNA is desalted on a desalting column (Glen Research, Sterling, VA) to remove salts and organics.

20     Cy3-azide modified dye is obtained from Sigma Aldrich (777315). For the click reaction, 1 mg of Cy3-azide is dissolved in 150  $\mu$ L of DMSO. The Cy3-azide solution is added to 10 OD of DBCO conjugated PNA in 100  $\mu$ L of water. The mixture is incubated at room temperature overnight. The ligated PNAs are desalted on a desalting column (Glen Research, Sterling, VA) to remove salts and organics.

**Example 13: Treatment of aphids (*Myzus persicae*) with a solution of Cy3-PNA**

25     This Example demonstrates the ability to kill or decrease the fitness of aphids, *Myzus persicae*, by treating them with a Cy3-PNA constructs that targets expression of the *Bcr1* gene (ACYPI32128), which is an essential gene for bacteriocyte regulation and function in insects.

30     Aphids are one of the most important agricultural insect pests. They cause direct feeding damage to crops and serve as vectors of plant viruses. In addition, aphid honeydew promotes the growth of sooty mold and attracts nuisance ants. The use of chemical treatments, unfortunately still widespread, leads to the selection of resistant individuals whose eradication becomes increasingly difficult.

*Therapeutic design:* Cy3-PNA solutions are formulated with 0, 0.5, 1, 5 or 10 mg/L of Cy3-PNA from Example 11 in 10 mL of 0.5% sucrose and essential amino acids.

*Experimental design:*

35     To prepare for the treatment, aphids are grown in a lab environment and medium. In a climate-controlled room (16 h light photoperiod; 60 $\pm$ 5% RH; 20 $\pm$ 2  $^{\circ}$ C), plants are grown in a mixture of vermiculite and perlite and are infested with aphids. To limit maternal effects or health differences between plants, 5–10 adults from different plants are distributed among 10 two-week-old plants, and allowed to multiply to high density for 5–7 days. For experiments, second and third instar aphids are collected from healthy

plants and divided into treatments so that each treatment receives approximately the same number of individuals from each of the collection plants.

Wells of a 96-well plate are filled with 200  $\mu$ l of artificial aphid diet (Febvay et al., *Canadian Journal of Zoology* 66(11):2449-2453, 1988) and the plate is covered with parafilm to make a feeding sachet. Artificial diet is either mixed with the solution of sterile water with 0.5% sucrose and essential amino acids only as a negative control, or mixed with Cy3-PNA solutions. Cy3-PNA solutions are mixed with artificial diet to obtain final concentrations between 0.5 to 10 mg/L.

For each replicate treatment, 30-50 second and third instar aphids are placed individually in wells of a 96-well plate and a feeding sachet plate is inverted above them, allowing the insects to feed through the parafilm and keeping them restricted to individual wells. Experimental aphids are kept under the same environmental conditions as aphid colonies. After the aphids are fed for 24 hr, the feeding sachet is replaced with a new one containing sterile artificial diet and a new sterile sachet is provided every 24 h for 4 days. At the time that the sachet is replaced, aphids are also checked for mortality. An aphid is counted as dead if it had turned brown or is at the bottom of the well and does not move during the observation. If an aphid is on the parafilm of the feeding sachet but not moving, it is assumed to be feeding and alive.

The survival rates of aphids treated with *Bcr1* specific Cy3-PNA are compared to the aphids treated with the negative control. The survival rate of aphids treated with *Bcr1* specific Cy3-PNA is decreased as compared to the control treated aphids.

#### **Example 14: Topical-plant delivery of Cy3-PNA for crop protection**

This Example demonstrates the ability to deliver Cy3-PNA by topical application on tobacco plant leaves. Cy3-PNA targets the expression of the *Bcr1* gene (ACYPI32128), which is an essential gene for bacteriocyte regulation and function in insects.

*Therapeutic design:* Cy3-PNA-LDH spray solutions are formulated with Cy3-PNA synthesized in Example 11 at 1:0 (negative control), 1:1, 1:2, or 1:3 ratio of Cy3-PNA:LDH at 125  $\mu$ L/cm<sup>2</sup> of leave surface.

#### *Preparation and characterization of LDH nanosheets:*

Sheet-like clay nanoparticles, specifically positively charged LDH, are excellent nanocarrier systems to deliver nucleic acids as stable spray formulations for crop protection (Mitter et al., *Nature Plants* 1-10, 2017). LDH conjugates strongly adhere to the leaf surface even after vigorous rinsing, and enhance the stability of PNAs for a longer period under environmental conditions. The sustained release of PNAs is facilitated through the formation of carbonic acid on the leaf surface from atmospheric CO<sub>2</sub> and humidity, which helps degrade the LDH nanosheets.

LDH nanosheets are prepared according to (Mitter et al., *Nature Plants* 1-10, 2017) by modified non-aqueous precipitation, followed by heat treatment, purification, and dispersion in water to get an average particle size of 45 nm. The particles (five repeated LDH samples) are characterized by the Nanosizer Nano ZS instrument (Malvern Instruments) to obtain the Z average size and Pdl and imaged using JEOL JSM-2010 TEM22. The chemical composition and crystal structure are verified by powder XRD with five LDH samples and Fourier transform infrared spectroscopy with attenuated total reflection

mode, using a Rigaku Miniflex X-Ray diffractometer and a Nicolet 6700 FT-IR (Thermo Electron Corporation), respectively.

*Cy3-PNA loading on LDH:*

5 To define optimal and complete loading of the Cy3-PNAs obtained in Example 11 into LDH nanosheets, the ratio of Cy3-PNA (0.5 µg) to LDH (Cy3-PNA-LDH (w/w)) is assayed at 1:1, 1:2, 1:3, 1:4, 1:5 and 1:10 multiple times. Cy3-PNA and LDH are incubated in a total volume of 10 µL at room temperature for 30 min with gentle orbital agitation. Complete Cy3-PNA loading is assessed by retention of Cy3-PNA-LDH complexes in a well of a 1% agarose gel. Appropriate loading ratios are consistent,  
10 irrespective of the scale-up volume required.

*Confocal imaging for Cy3-PNA detection within plant leaves:*

Surface-sterilized seeds of *Arabidopsis thaliana* Col-0 are vernalized, plated and vertically grown at 21°C (16 h/8 h day/night) for 11 days on solid MS media.

15 To detect uptake of Cy3-PNA, 11-day-old *A. thaliana* seedlings in replicates of three are treated with 3 µL droplets of Cy3 only, Cy3-PNA (1 µg) and Cy3-PNA-LDH (1:3). Droplets are applied onto individual leaves. After 48 h the leaves are rinsed twice for 2 min in 3 ml of water with vigorous pipetting and visualized below the epidermis as florescence remained on the surface in case of LDH treatments. Natural chlorophyll florescence is detected with 650-800 nm bandpath filter. Z-stacks are used to verify  
20 that the florescence observed in the spongy mesophyll and xylem are internalized. The leaves are viewed under a Zeiss LSM510 META (Carl Zeiss) confocal laser scanning microscope. Cy3 florescence is visualized by excitation with a HeNe laser at 543 nm and detected with a 560–615 nm bandpath filter.

**Example 15: Treatment of aphids with a solution that increases body temperature**

25 This Example demonstrates the ability to kill or decrease the fitness of aphids by treating them with prostaglandin. Prostaglandins are eicosanoids associated with immune function that rapidly induce fever in vertebrates and invertebrates. This example demonstrates that the effect of prostaglandin on aphids is mediated through the modulation of bacterial populations endogenous to the aphid that are sensitive to an increase in temperature generated by prostaglandin. One targeted bacterial strain is  
30 *Buchnera*.

*Therapeutic design:* The prostaglandin solution E2 (PGE2) was formulated with 0 (negative control), 10, 20 or 50 µg/µl of prostaglandin formulated in a solution of ethanol and sterile water with 0.5% sucrose and essential aminoacids.

35 *Experimental design:*

To prepare for the treatment, aphids are grown in a lab environment and medium. In a climate-controlled room (16 h light photoperiod; 60±5% RH; 20±2 °C), fava bean plants are grown in a mixture of vermiculite and perlite at 24°C with 16 h of light and 8 h of darkness. To limit maternal effects or health differences between plants, 5–10 adults from different plants are distributed among 10 two-week-old  
40 plants, and allowed to multiply to high density for 5–7 days. For experiments, second and third instar

aphids are collected from healthy plants and divided into treatments so that each treatment receives approximately the same number of individuals from each of the collection plants.

Prostaglandin solutions are made by dissolving prostaglandin (SIGMA-ALDRICH, P5640) in sterile water and ethanol (1:1) with 0.5% sucrose and essential aminoacids. Wells of a 96-well plate are filled with 200 µl of artificial aphid diet (Febvay, et al., Canadian Journal of Zoology, 1988, 66(11): 2449-2453) and the plate is covered with parafilm to make a feeding sachet. Artificial diet is either mixed with sterile water and with 0.5% sucrose and essential aminoacids as a negative control or a prostaglandin solution containing one of the concentrations of prostaglandin. Prostaglandin solutions are mixed with artificial diet to get final concentrations between 10 and 50 µg/µL.

For each replicate treatment, 30–50 second and third instar aphids are placed individually in wells of a 96-well plate and a feeding sachet plate is inverted above them, allowing the insects to feed through the parafilm and keeping them restricted to individual wells. Experimental aphids are kept under the same environmental conditions as aphid colonies. After the aphids are fed for 24 hr, the feeding sachet is replaced with a new one containing sterile artificial diet and a new sterile sachet is provided every 24 h for four days. At the time that the sachet is replaced, aphids are also checked for mortality. An aphid is counted as dead if it had turned brown or is at the bottom of the well and does not move during the observation. If an aphid is on the parafilm of the feeding sachet but not moving, it is assumed to be feeding and alive.

The status of *Buchnera* in aphid samples is assessed by PCR. Total DNA is isolated from control (non- prostaglandin treated) and prostaglandin treated individuals using an Insect DNA Kit (OMEGA, Bio-tek) according to the manufacturer's protocol. The primers for *Buchnera*, forward primer 5'-GTCTGGCTCATCACATCC-3' (SEQ ID NO: 97) and reverse primer 5'-TTCCGTCTGTATTATCTCCT-3' (SEQ ID NO: 98), are designed based on *23S-5S rRNA* sequences obtained from the *Buchnera* genome (Accession Number: GCA\_000009605.1) (Shigenobu, et al., *Nature* 200.407, 81–86) using Primer 5.0 software (Primer-E Ltd., Plymouth, UK). The PCR amplification cycles included an initial denaturation step at 95°C for 5min, 35 cycles of 95°C for 30s, 55°C for 30s, and 72°C for 60s, and a final extension step of 10min at 72°C. Amplification products from prostaglandin treated and control samples are analyzed on 1% agarose gels, stained with SYBR safe, and visualized using an imaging System. Prostaglandin treated aphids show a reduction of *Buchnera* specific genes.

The survival rates of aphids treated with prostaglandin solution are compared to the aphids treated with the negative control. The survival rate of aphids treated with prostaglandin solution is decreased compared to the control.

#### **Example 16: Treatment of aphids with dsRNA to stimulate an insect immune response to decrease fitness**

This Example demonstrates that the treatment of aphids with double-stranded RNA (dsRNA) resulted in the knock-down of immunoregulatory genes to induce an immune response and decrease aphid fitness. By inducing an immune response by inhibiting an immunoregulatory gene, specifically, Cact a negative regulator of the Toll pathway, which is the primary immunity pathway in aphids, to induce Toll pathway activation to express and secrete antimicrobial peptides, lysozymes, and prophenoloxidase

that ultimately affect bacterial populations endogenous to the aphid. This Example demonstrates that the effect of lowering the levels of Cact in aphids to upregulate the systemic immune responses leading to the dysregulation of bacterial populations endogenous to the aphid that are sensitive to the increase in Aphid immune responses generated by Toll pathway activation. One targeted bacterial strain is *Buchnera*.

5

*Therapeutic design:* 5<sup>th</sup> instar LSR-1 aphids are microinjected. The injection solutions will be either dd-water (negative control) or dsRNA diluted in dd-water at various concentrations (8 or 60 ng/aphid; see below).

10

*Experimental design:*

Aphids LSR-1 (which harbor only *Buchnera*), *A. pisum* will be grown on fava bean plants (*Vicia faba* from Johnny's Selected Seeds) in a climate-controlled incubator (16 h light/8 h dark photoperiod; 60±5% RH; 25±2 °C). Prior to being used for aphid rearing, fava bean plants will be grown in potting soil at 24°C with 16 h of light and 8 h of darkness. To limit maternal effects or health differences between plants, 5–10 adults from different plants will be distributed among 10 two-week-old plants, and allowed to multiply to high density for 5–7 days. For experiments, fifth instar aphids will be collected from healthy plants and divided into 2 different treatment groups: 1) water (negative control) or 2) dsRNA against ApGLNT1 (at the concentrations indicated herein).

*Microinjection Delivery Experimental Design*

Microinjection will be performed using NanoJet III Auto-Nanoliter Injector with an in-house pulled borosilicate needle (Drummond Scientific; Cat# 3-000-203-G/XL). Aphids will be transferred using a paint brush to a tubing system connected to vacuum which held the aphid in place during the microinjection.

The injection site will be at the ventral thorax of the aphid. The injection volume will be 20 nl for adult aphids at a rate of 2 nl/sec. Each treatment group will have approximately the same number of individuals injected from each of the collection plants.

After injection, aphids will be released into a petri dish onto a fava bean leaves that have stems in an Eppendorf tube filled with 1 ml water. Aphid survival will be monitored daily, and dead aphids will be removed when they are found. The number of offspring from each group will be counted and fecundity will be measured as the number of offspring (F1's) produced per aphid at each time point.

In select experiments, development will be measured in groups of offspring from each treatment group by noting the developmental stage of offspring each day (1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> instar). Development will be also measured by imaging aphids at 4 days post-collection and determining their area. The template for synthesizing dsRNAs will be the cDNA reverse-transcribed from the mRNA. RNA will be extracted from one 5<sup>th</sup> instar *A. pisum* (LSR-1 strain) and is quantified by Nanodrop (Thermo fisher scientific). ~100 µg of total RNA will be added as template into the reverse-transcription reaction using Superscript IV kit (Thermo Fisher Scientific) following manufacturer's protocol.

To amplify the templates for the dsRNAs, the cDNA will be diluted 100-fold and used in the following PCRs. The PCR reactions (25 µl final volume) contain 12.5 µL of Go Taq Green 2X mix (Promega), 0.2 µl of forward primer (Table 12), 0.2 µl of reverse primer (Table 12), and 12.1 µl of 100-fold diluted cDNA. PCR reactions are performed using following conditions: 1) 95°C for 2 minutes, 2) 95°C for 20 seconds, 3) 55°C for 15 seconds, 4) 72°C for 30 seconds, 5) repeat steps 2-4 35x, 6) 72°C for 5 minutes. The sizes of PCR amplified products are verified by electrophoresis on 1.5% agarose and the expected-size bands are cut and purified by QIAquick DNA purification kit (Qiagen). The dsRNAs will be synthesized *in vitro* using T7 MEGAscript kit (Ambion, Thermo Fisher Scientific; Cat# AM1334) following manufacturer's protocol. The concentration of dsRNA is measured by Nanodrop (Thermo Fisher Scientific).

**Table 12: Cact dsRNA sequences:**

| Target gene | Forward primer                                                  | Reverse primer                                                    | product size |
|-------------|-----------------------------------------------------------------|-------------------------------------------------------------------|--------------|
| Mpca ctus   | taatacgactcactatagggTACACC<br>CATTGTGTGCACCT (SEQ ID<br>NO: 99) | taatacgactcactatagggCCTACTG<br>TCCAAGGCAATTTT (SEQ ID<br>NO: 100) | 487          |

The status of *Buchnera* in aphid samples will be assessed by PCR. Aphids adults from the negative control and phage treated will be first surface-sterilized with 70% ethanol for 1 min, 10% bleach for 1 min and three washes of ultrapure water for 1 min. Total DNA will be extracted from each individual (whole body) using an Insect DNA Kit (OMEGA, Bio-tek) according to the manufacturer's protocol. The primers for *Buchnera*, forward primer 5'-GTCGGCTCATCACATCC-3' (SEQ ID NO: 97) and reverse primer 5'-TTCCGTCTGTATTATCTCCT-3' (SEQ ID NO: 98), are designed based on *23S-5S rRNA* sequences obtained from the *Buchnera* genome (Accession Number: GCA\_000009605.1) (Shigenobu, et al., *Nature* 200.407, 81–86) using Primer 5.0 software (Primer-E Ltd., Plymouth, UK). The PCR amplification cycles include an initial denaturation step at 95°C for 5min, 35 cycles of 95°C for 30s, 55°C for 30s, and 72°C for 60s, and a final extension step of 10min at 72°C. Amplification products from rifampicin treated and control samples will be analyzed on 1% agarose gels, stained with SYBR safe, and visualized using an imaging System. dsRNA treated aphids are expected to show a reduction of *Buchnera* specific genes.

The expression of Cact and survival rates of aphids treated with Cactus are compared to the aphids treated with the negative control. The expression of Cact and survival rate of aphids treated with dsRNA to Cact are expected to decrease as compared to the control treated aphids.

#### **Example 17: Aphids treated with a fungi solution that stimulated an insect immune response to decrease fitness**

This Example demonstrates the treatment of aphids with a solution of *Pichia pastoris*, a yeast strain that resulted in decreased aphid fitness. Although many of the innate immune system pathways present in insects are absent in aphids, the key players that recognize and induce immune responses to fungi remain intact. Specifically, the presence of fungi induces the cleavage of the aphid serine protease, Persephone, which then activates the Toll pathway. Next, Cactus is phosphorylated and Dorsal is released and translocated to the nucleus which activates the expression and secretion of antimicrobial peptides, lysozymes, and prophenoloxidase that ultimately affect bacterial populations endogenous to the aphid. This Example demonstrates that the effect of the solution of *Pichia pastoris* on aphids was mediated through the modulation of bacterial populations endogenous to the aphid that were sensitive to the increase in the Aphid immune response generated by the presence of the yeast. One targeted bacterial strain was *Buchnera*.

Aphids are agricultural insect pests causing direct feeding damage to the plant and serving as vectors of plant viruses. In addition, aphid honeydew promotes the growth of sooty mold and attracts nuisance ants. The use of chemical treatments, unfortunately still widespread, leads to the selection of resistant individuals whose eradication becomes increasingly difficult.

### Therapeutic Design

*P. pastoris* was delivered using two different methods: Fava bean leaf perfusion and air brush spraying onto fava bean leaves. For each experiment, there were two experimental groups: 1) treatment with water as a negative control; and 2) treatment with *P. pastoris* in water. Treatment methods and  
5 doses are described herein in the Experimental Design section.

### Leaf Perfusion Experimental Design

Aphids (LSR-1 strain, *Acyrtosiphon pisum*) were grown on fava bean plants (*Vicia faba* from Johnny's Selected Seeds) in a climate-controlled incubator (16 h light/8 h dark photoperiod; 60±5% RH; 25±2 °C). Prior to being used for aphid rearing, fava bean plants were grown in potting soil at 24°C  
10 with 16 h of light and 8 h of darkness. To limit maternal effects or health differences between plants, 5–10 adults from different plants were distributed among 10 two-week-old plants, and allowed to multiply to high density for 5–7 days. For experiments, first and second instar aphids were collected from healthy plants and divided into two different treatment groups: 1) those allowed to feed on leaves perfused with  
15 water; and 2) those allowed to feed on leaves perfused with a *P. pastoris* solution in water.

The “protease wild-type” *ade2* knockout strain (Strain 1) from the Pichia Pink System (Thermo Fisher Scientific) was used for experiments. *P. pastoris* was grown in YPD overnight at 30°C, and the following day, the culture was washed once with water and resuspended in water to a final OD<sub>600</sub> of 0.918 for the first leaf perfusion on day 0 or OD<sub>600</sub> of 5.58 for the second leaf perfusion on day 3 of the  
20 experiment. For each leaf perfusion, approximately 1 ml of the *P. pastoris* solution or water (negative control) was injected into a fava bean leaf. The stem of the leaf was then placed in a 1.5 ml Eppendorf that was sealed with parafilm. The leaf stem was then placed into a deep petri dish (Fisher Scientific, Cat# FB0875711) and aphids were applied to the leaves of the plant and allowed to feed. Leaves were changed on day 3 of the experiment and old leaves were replaced with new leaves perfused with *P.*  
25 *pastoris* (at the indicated OD<sub>600</sub>) or water.

For each treatment, 62-63 aphids were placed onto each leaf. Aphids were monitored daily for survival and dead aphids were removed when they were discovered. In addition, the developmental stage (1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> instar) was determined daily throughout the experiment.

After 6 days of treatment, DNA was extracted from multiple aphids from each treatment group. Briefly, the aphid body surface was sterilized by dipping the aphid into a 6% bleach solution for  
30 approximately 5 seconds. Aphids were then rinsed in sterile water and DNA was extracted from each individual aphid using a DNA extraction kit (Qiagen, DNeasy kit) according to manufacturer's instructions. DNA concentration was measured using a nanodrop nucleic acid quantification, and *Buchnera* and aphid DNA copy numbers were measured by qPCR. The primers used for *Buchnera* were Buch\_groES\_18F  
35 (CATGATCGTGTGCTTGTTAAG; SEQ ID NO: 101) and Buch\_groES\_98R (CTGTTCCCTCGAGTCGATTTCC; SEQ ID NO: 102) (Chong and Moran, 2016 PNAS). The primers used for aphid were ApEF1a 107F (CTGATTGTGCCGTGCTTATTG; SEQ ID NO: 103) and ApEF1a 246R (TATGGTGGTTTCAGTAGAGTCC; SEQ ID NO: 104) (Chong and Moran, 2016 PNAS). qPCR was performed using a qPCR amplification ramp of 1.6 degrees C/s and the following conditions: 1) 95°C for  
40 10 minutes, 2) 95°C for 15 seconds, 3) 55°C for 30 seconds, 4) repeat steps 2-3 40x, 5) 95°C for 15

seconds, 6) 55°C for 1 minute, 7) ramp change to 0.15 degrees C/s, 8) 95°C for 1 second. qPCR data was analyzed using analytic (Thermo Fisher Scientific, QuantStudio Design and Analysis) software.

*Treatment with a yeast solution delayed progression of aphid development*

5 LSR-1 first and second instar aphids were divided into two treatment groups as defined in *Leaf Perfusion Experimental Design*. Aphids were monitored daily and the number of aphids at each developmental stage was determined. By day 6 of the experiment, approximately 30% of aphids feeding on leaves perfused with water reached the fifth instar stage (Fig. 1). In contrast, only approximately 3% of aphids feeding on leaves perfused with *P. pastoris* reached the fifth instar stage by day 6 of the  
10 experiment (Fig. 1). The majority of aphids feeding on the leaves perfused with *P. pastoris* that survived to day 6 were at the 4<sup>th</sup> instar stage (~5%) (Fig. 1). These data indicated that *P. pastoris* treatment slowed aphid development.

*Treatment with a yeast solution increased aphid mortality*

15 Survival of aphids was also measured during the treatments. Approximately 55.5% of aphids feeding on leaves perfused with water survived to day 6 of the experiment (Fig. 2). In contrast, aphids feeding on leaves perfused with *P. pastoris* rapidly began to die at day 2 post-treatment and by day 6, only 11% of aphids were alive (Fig. 2). These data indicated that *P. pastoris* treatment delivered through leaf perfusion resulted in a significant ( $p < 0.0001$ ) increase in aphid mortality.

20 To test whether *P. pastoris* treatment specifically resulted in loss of *Buchnera* in aphids, and that this loss impacted aphid fitness, DNA was extracted from aphids in each treatment group after 6 days of treatment and qPCR was performed to determine the *Buchnera*/aphid copy numbers. Results were inconclusive due to the low number of aphids remaining in the *P. pastoris* treatment group at day 6 and the lack of extractable *Buchnera* DNA.

25 *Airbrush Spraying Experimental Design:*  
Aphids (LSR-1 strain, *Acyrtosiphon pisum*) were grown on fava bean plants as described herein. For experiments, first and second instar aphids were collected from healthy plants and divided into 2 different treatment groups: 1) aphids placed on leaves sprayed with water (negative control), and 2)  
30 aphids placed on leaves sprayed with *P. pastoris*.

The “protease wild-type” *ade2* knockout strain (Strain 1) from the Pichia Pink System (Thermo Fisher Scientific) was used for experiments. *P. pastoris* was grown in YPD overnight at 30°C and the following day, the culture was washed once with water and resuspended in water to a final OD600 of 0.918 for the first and second leaf spraying (on day 0 and 3) and OD600 of 5.58 for the final leaf spraying  
35 on day 6 of the experiment. For the treatments, fava bean leaves were cut and the stems were placed into a 1.5 ml Eppendorf tube with sterile water and sprayed on both sides using an airbrush with water or *P. pastoris* at the concentration indicated above. The leaves were then placed in a deep petri dish (Fisher Scientific, Cat# FB0875711) and aphids were applied to the leaves of the plant and allowed to feed. On day 3 and 6 of the experiment, old leaves were replaced with new, freshly sprayed leaves.

Each treatment group received approximately the same number of individuals from each of the collection plant. For each treatment, 30 aphids were placed onto each leaf. Each treatment was performed in duplicate for a total of 60 aphids/treatment group. Aphids were monitored daily for survival, and dead aphids were removed when they were discovered. In addition, the developmental stage (1st, 2nd, 3rd, 4th, and 5th instar) was determined daily throughout the experiment.

After 6 and 9 days of treatment, DNA was extracted from multiple aphids from each treatment group and qPCR for quantifying *Buchnera* levels was done as described herein.

#### *Treatment with a yeast solution did not greatly impact aphid development*

LSR-1 first and second instar aphids were divided into two treatment groups as defined in the *Airbrush Spraying Experimental Design* described herein. Aphids were monitored daily and the number of aphids at each developmental stage was determined. By 6 days post-treatment, aphids from both treatment groups began reaching the 5<sup>th</sup> instar stage (Fig. 3). By day 9 post-treatment, nearly all remaining aphids in each group reached maturity (Fig. 3). These data indicated that there was little difference in development between aphids treated with water or *P. pastoris*.

#### *A yeast solution treatment via spraying resulted in increased aphid mortality*

Survival of aphids was also measured during the leaf spraying experiments. Throughout the course of the experiment, only a few aphids feeding on leaves sprayed with water died, whereas a greater number of aphids died at each time point in the *P. pastoris* treatment group (Fig. 4). Specifically, on day 2 post-treatment 91% of water-treated aphids remained alive, while only 80% of *P. pastoris*-treated aphids remained alive. This trend continued through days 3, 6, and 7 of the experiment where 82%, 66%, and 60% of water-treated aphids were alive at each time point, respectively, in contrast to only 73%, 55%, and 49% of *P. pastoris*-treated aphids were alive at each time point, respectively (Fig. 4). These data indicated that *P. pastoris* treatment delivered by airbrush spraying resulted in higher mortality compared to treatment with water alone.

#### *A yeast solution treatment reduced Buchnera titers in aphids*

To test whether *P. pastoris* delivered through airbrush spraying, specifically resulted in loss of *Buchnera* in aphids, and that this loss impacted aphid fitness, DNA was extracted from aphids in each treatment group after 6 and 9 days of treatment and qPCR was performed to determine the *Buchnera*/aphid copy numbers. At 6 days post-treatment the mean *Buchnera*/aphid ratios were 47.5 in aphids feeding on leaves sprayed with water (Fig. 5). In contrast, mean *Buchnera*/aphid ratios were approximately 1.2-fold lower (~39) in aphids feeding on leaves sprayed with *P. pastoris* (Fig. 5). At 9 days post-treatment however, the ratios of *Buchnera*/aphid DNA copies were similar in both treatment groups (Fig. 5).

These data indicated that spraying *P. pastoris* onto fava bean leaves decreased endosymbiotic *Buchnera* in aphids after 6 days post-treatment. It was possible that *P. pastoris*-treated aphids that survived to 9 days post-treatment were able to overcome the detrimental effects of the increased immune response and could explain why these aphids had similar *Buchnera* titers to the water treated controls.

Nonetheless, examination of the median *Buchnera*/aphid copies on day 9, revealed a trend for decreased *Buchnera* in the *P. pastoris* treatment group.

*Confirmation of that P. pastoris upregulates the immune response in aphids*

Given that treatment with *P. pastoris* resulted in decreased aphid fitness and lower titers of the aphid endosymbiont, *Buchnera*, the next experiment is to confirm that *P. pastoris* is upregulating the immune response. To test this, RNA has been isolated from aphids in each of the experiments described herein (Leaf Perfusion Experimental Design and Airbrush Spraying Experimental Design) and the expression of genes involved in the fungal immune response pathway will be measured (Persephone, Cactus, and Dorsal) as described in Gerardo et al., 2010, Genome Biology, 11:R21, <https://doi.org/10.1186/gb-2010-11-2-r21>. It is expected to see upregulation of the genes involved in this pathway.

Together, the data described in these Examples demonstrated the ability to kill and decrease the development, longevity, and endogenous bacterial populations, e.g., fitness, of aphids by treating them with *P. pastoris* which likely resulted in activation of the aphid immune response.

**Example 18: Aphids treated with a solution of a peptide nucleic acid**

This Example demonstrates that the treatment of aphids with a peptide nucleic acid (PNA) to BCR-4 fused to a cell penetrating peptide (CPP) (Cermenati et al., 2011; Zhou et al., 2015), herein referred to as PNA to BCR-4 or BCR-4 PNA, resulted in knock-down of BCR-4 gene expression and reduction of aphid fitness.

BCR-4 is one of several cysteine-rich secreted peptides expressed in bacteriocytes of the pea aphid, *Acyrtosiphon pisum*. The obligate aphid bacterial symbiont, *Buchnera*, is housed inside of the bacteriocytes. BCR-4 has sequence homology to many of the nodule cysteine rich (NCRs) peptides involved in keeping plant bacterial symbiont numbers in check (Pan and Wang, 2017 Nature Plants and Durgo et al., 2015 Proteomics). Given the sequence similarity of BCR-4 and the NCRs, this Example demonstrates that BCR-4 played a similar role in maintaining endosymbiont homeostasis inside the bacteriocyte by disrupting BCR-4 through gene knockdown to dysregulate *Buchnera*, thereby negatively affecting aphid fitness.

*Therapeutic Design*

The BCR-4 PNA was delivered either by microinjection or through fava bean leaf perfusion. For microinjection experiments, injection solutions were either water (negative control) or BCR-4 PNA in water. For leaf perfusion studies, fava bean leaves were perfused with water (negative control) or with BCR-4 PNA in water. Each experimental delivery design is explained in detail below.

*Microinjection Delivery Experimental Design*

Microinjection was performed using NanoJet III Auto-Nanoliter Injector with an in-house pulled borosilicate needle (Drummond Scientific; Cat# 3-000-203-G/XL). Aphids were transferred using a paint brush to a tubing system connected to vacuum which held them in place during the injection. The

injection site was at the ventral thorax of the aphid. The injection solutions were water (negative control) or 321 ng/μl of BCR-4 PNA in water. The injection volume was 20 nl for adult (4th and 5th instar) aphids at a rate of 2 nl/sec. Each treatment group had approximately the same number of individuals injected from each of the collection plants. After injection, aphids were released into a petri dish onto fava bean leaves that had stems in an Eppendorf tube filled with 1 ml water. Survival and fecundity were monitored over the course of the experiment.

#### *Aphid Rearing and Maintenance:*

Aphids LSR-1 (which harbor only *Buchnera*), *A. pisum* were grown on fava bean plants (*Vicia faba* from Johnny's Selected Seeds) in a climate-controlled incubator (16 h light/8 h dark photoperiod; 60±5% RH; 25±2 °C). Prior to being used for aphid rearing, fava bean plants were grown in potting soil at 24°C with 16 h of light and 8 h of darkness. To limit maternal effects or health differences between plants, 5–10 adults from different plants were distributed among 10 two-week-old plants, and allowed to multiply to high density for 5–7 days. For experiments, fourth and fifth instar aphids were collected from healthy plants and divided into two treatment groups: 1) water (negative control) or 2) BCR-4 PNA in water.

#### *BCR-4 PNA synthesis*

BCR-4-CPP was synthesized by PNA bio and the sequence is YGRKKRRQRRR-CGTACAATAATCTCATGG; SEQ ID NOs: 105 and 106. The sequence of the CPP (Tat) is YGRKKRRQRRR; SEQ ID NO: 106. The PNA was dissolved in 80% Acetonitrile and 20% Water supplemented with 0.1% Trifluoroacetic acid (TFA). Once dissolved, the PNA was aliquoted [32.1 μg / 5 nmol per aliquot], air dried, and stored at -20°C. Working solutions of the PNA were made in water at 50 μM.

After 7 days of treatment, RNA was extracted from the remaining aphids in each treatment group. Briefly, the aphid body surface was sterilized by dipping the aphid into a 6% bleach solution for approximately 5 seconds. Aphids were then rinsed in sterile water and total RNA was extracted from each individual aphid using the RNA extraction kit (Qiagen miRNeasy kit) according to manufacturer's instructions. RNA concentrations were measured using a nanodrop nucleic acid quantification. BCR-4 relative expression was measured by RT-qPCR. The primers used were ApBCR-4F (CTCTGTCAACCACCATGAGATTA; SEQ ID NO: 107) and ApBCR-4R (TGCAGACTACAGCACAATACTT; SEQ ID NO: 108). The internal reference gene primers were for Actin (housekeeping gene). The forward sequence was GATCAGCAGCCACACACAAG; SEQ ID NO: 109 and the reverse sequence was TTTGAACCGGTTTACGACGA; SEQ ID NO: 110. RT-qPCR was performed using a qPCR amplification ramp of 1.6 degrees C/s and the following conditions: 1) 48°C for 30 min, 2) 95°C for 10 minutes, 3) 95°C for 15 seconds, 4) 60°C for 1 minute, 5) repeat steps 3-4 40x, 6) 95°C for 15 seconds, 7) 60°C for 1 minute, 8) ramp change to 0.15 degrees C/s, 9) 95°C for 1 second. RT-qPCR data was analyzed using analytic (Thermo Fisher Scientific, QuantStudio Design and Analysis) software.

*Microinjection with a PNA against the BCR-4 gene resulted in decreased expression of BCR-4*

To test whether a PNA against BCR-4 delivered by microinjection results in decreased BCR-4 gene expression in aphids, aphids were injected with water (control) or BCR-4 PNA. After 7 days of treatment, RNA was extracted from aphids in each treatment group and RT-qPCR was performed.

- 5 Aphids microinjected with water had approximately 2-fold higher expression of BCR-4 compared to aphids injected with BCR-4 PNA (Fig. 6), indicating that injection of BCR-4 PNA resulted in knockdown of BCR-4.

*Treatment with a PNA to BCR-4 increased aphid mortality*

- 10 Survival rate of aphids was also measured during the treatments. At most time points during the experiment, there were more control (water injected) aphids alive compared to aphids injected with the PNA to BCR-4 (Fig. 7). These data indicated that there was a slight decrease in survival upon injection with a PNA to BCR-4.

15 *Treatment with a PNA to BCR-4 reduced aphid fecundity*

- The number of offspring produced from aphids in each treatment group was also assessed during the experiment and fecundity was represented as the number of offspring produced per adult at each time point assessed. Overall, there was a trend for control (water-injected) aphids producing more offspring compared to those injected with the PNA to BCR-4. Specifically, at 3 and 7 days post-treatment, aphids in the water treated group produced an average of 5 offspring/adult whereas aphids in the BCR-4 PNA treatment group only produced 4 offspring/adult (Fig. 8). These data indicated that treatment with BCR-4 PNA resulted in a slight decrease in aphid fecundity.
- 20

*Leaf Perfusion Experimental Design*

- 25 Aphids (LSR-1 strain, *Acyrtosiphon pisum*) were grown on fava bean plants (*Vicia faba* from Johnny's Selected Seeds) in a climate-controlled incubator (16 h light/8 h dark photoperiod; 60±5% RH; 25±2 °C). Prior to being used for aphid rearing, fava bean plants were grown in potting soil at 24°C with 16 h of light and 8 h of darkness. To limit maternal effects or health differences between plants, 5–10 adults from different plants were distributed among 10 two-week-old plants and allowed to multiply to high density for 5–7 days. For experiments, first and second instar aphids were collected from healthy plants and divided into two different treatment groups: 1) those allowed to feed on leaves perfused with water, and, 2) those allowed to feed on leaves perfused with a BCR-4 PNA solution in water.
- 30

The BCR-4 PNA fused to a CPP was synthesized as described herein (see Microinjection Delivery Experimental Design BCR-4 PNA synthesis).

- 35 For each leaf perfusion, approximately 1 ml of water (negative control) or a 1 µM BCR-4 PNA solution was injected into a fava bean leaf. The stem of the leaf was then placed in a 1.5 ml Eppendorf that was sealed with parafilm. The leaf stem was then placed into a deep petri dish (Fisher Scientific, Cat# FB0875711) and aphids were applied to the leaves of the plant and allowed to feed. Old leaves were replaced with new, freshly injected leaves every 2-3 days throughout the experiment. For each treatment, 60 aphids were placed onto each leaf. Aphids were monitored daily for survival and dead
- 40

aphids were removed when they were discovered. In addition, the developmental stage (1st, 2nd, 3rd, 4th, and 5th instar) was determined daily throughout the experiment.

After 5 and 6 days of treatment, DNA was extracted from dead aphids from each treatment group. Briefly, the aphid body surface was sterilized by dipping the aphid into a 6% bleach solution for approximately 5 seconds. Aphids were then rinsed in sterile water and DNA was extracted from each individual aphid using a DNA extraction kit (Qiagen, DNeasy kit) according to manufacturer's instructions. DNA concentration was measured using a nanodrop nucleic acid quantification, and *Buchnera* and aphid DNA copy numbers were measured by qPCR. The primers used for *Buchnera* were Buch\_groES\_18F (CATGATCGTGTGCTTGTTAAG; SEQ ID NO: 101) and Buch\_groES\_98R (CTGTTCCCTCGAGTCGATTTCC; SEQ ID NO: 102) (Chong and Moran, 2016 PNAS). The primers used for aphid were ApEF1a 107F (CTGATTGTGCCGTGCTTATTG; SEQ ID NO: 103) and ApEF1a 246R (TATGGTGGTTCAGTAGAGTCC; SEQ ID NO: 104) (Chong and Moran, 2016 PNAS). qPCR was performed using a qPCR amplification ramp of 1.6 degrees C/s and the following conditions: 1) 95°C for 10 minutes, 2) 95°C for 15 seconds, 3) 55°C for 30 seconds, 4) repeat steps 2-3 40x, 5) 95°C for 15 seconds, 6) 55°C for 1 minute, 7) ramp change to 0.15 degrees C/s, 8) 95°C for 1 second. qPCR data was analyzed using analytic (Thermo Fisher Scientific, QuantStudio Design and Analysis) software.

At 7 post-treatment, RNA was extracted from live aphids and RT-pPCR was done as described above (see Microinjection Delivery Experimental Design) to quantify expression of BCR-4.

#### 20 *Leaf perfusion treatment with BCR-4 PNA delayed progression of aphid development*

LSR-1 first and second instar aphids were divided into two groups as defined in Leaf Perfusion Experimental Design. Aphids were monitored daily and the number of aphids at each developmental stage was determined. At several time points during the experiment, development was delayed in aphids treated with the BCR-4 PNA. For example, on day 2 post-treatment, approximately 19% of aphids feeding on leaves perfused with water were in the 3rd instar stage (Fig. 9). In contrast, only 8% of aphids feeding on leaves perfused with the PNA to BCR-4 were in the 3rd instar stage. These data indicated that treatment with a PNA to BCR-4 slowed aphid development.

#### *Leaf perfusion treatment with BCR-4 PNA increased aphid mortality*

Survival was also monitored throughout the course of treatment. By 7 days post-treatment, 53% of aphids in the water (control) treatment group remained alive (Fig. 10). In contrast, only 30% of aphids in the PNA BCR-4 treatment group were alive on day 7 (Fig. 10). These data showed that treatment with a PNA to BCR-4 resulted in increased aphid mortality.

#### 35 *Leaf perfusion treatment with BCR-4 PNA increased Buchnera titers in aphids*

To test whether BCR-4 PNA delivered through leaf perfusion, specifically resulted affecting *Buchnera* in aphids, and that this impacted aphid fitness, DNA was extracted from dead aphids in each treatment group after 5 and 6 days post-treatment, and qPCR was performed to determine *Buchnera* tiers in each treatment group. While aphids feeding on leaves perfused with water had a mean ratio of approximately 20 *Buchnera*/aphid copies, aphids feeding on leaves perfused with the PNA to BCR-4 had

substantially higher *Buchnera*/aphid copies (approximately 57), see Fig. 11. These data indicate that treatment with a PNA to BCR-4 leads to a misbalance in *Buchnera* titers.

*Treatment with a PNA against BCR-4 via leaf perfusion knocked down BCR-4*

To assess whether BCR-4 expression was reduced in aphids feeding on leaves perfused with a PNA to BCR-4, RNA was isolated from living aphids after 7 days of treatment and RT-qPCR was performed. The median transcript levels of BCR-4 in aphids treated with a PNA to BCR-4 were approximately 3-fold lower compared to aphids treated with water alone (Fig. 12), confirming that the PNA to BCR-4 knocked down BCR-4 expression.

Together, these data demonstrated the ability to kill and decrease the development, fecundity, and longevity, (e.g., fitness), of aphids by treating them with a PNA targeting a gene expressed in bacteriocytes (BCR-4) to control the population of endosymbiont *Buchnera*.

**Example 19: Aphids treated with a solution of dsRNA against bacteriocyte transporters**

This Example demonstrates that the treatment of aphids with double-stranded RNA (dsRNA) resulted in the knock-down of some essential genes, including glutamine transporter 1 gene (ApGLNT1), which has been identified in the bacteriocytes in the aphid, *Acyrtosiphon pisum*. The glutamine transporter is responsible for glutamine uptake from the aphid hemolymph into the bacteriocytes where the obligate endosymbiont, *Buchnera aphidicola*, is located. The combined biosynthetic capability of the holobiont (*A. pisum* and *Buchnera*) is sufficient for biosynthesis of all twenty protein coding amino acids, including amino acids that aphids alone cannot synthesize. Blocking the glutamine uptake by deactivating or silencing (e.g. RNAi by dsRNA) the glutamine transporter negatively affected amino acid and protein synthesis in the bacteriocytes and in the entire aphid, thereby negatively affecting their fitness.

*Therapeutic Design*

5<sup>th</sup> instar LSR-1 aphids were microinjected. The injection solutions were either dd-water (negative control) or dsRNA diluted in dd-water at various concentrations (8 or 60 ng/aphid; see below).

*Aphid Rearing and Maintenance*

Aphids LSR-1 (which harbor only *Buchnera*), *A. pisum* were grown on fava bean plants (*Vicia faba* from Johnny's Selected Seeds) in a climate-controlled incubator (16 h light/8 h dark photoperiod; 60±5% RH; 25±2 °C). Prior to being used for aphid rearing, fava bean plants were grown in potting soil at 24°C with 16 h of light and 8 h of darkness. To limit maternal effects or health differences between plants, 5–10 adults from different plants were distributed among 10 two-week-old plants, and allowed to multiply to high density for 5–7 days. For experiments, fifth instar aphids were collected from healthy plants and divided into 2 different treatment groups: 1) water (negative control) or 2) dsRNA against ApGLNT1 (at the concentrations indicated herein).

### Microinjection Delivery Experimental Design

Microinjection was performed using NanoJet III Auto-Nanoliter Injector with an in-house pulled borosilicate needle (Drummond Scientific; Cat# 3-000-203-G/XL). Aphids were transferred using a paint brush to a tubing system connected to vacuum which held the aphid in place during the microinjection.

5 The injection site was at the ventral thorax of the aphid. The injection volume was 20 nl for adult aphids at a rate of 2 nl/sec. Each treatment group had approximately the same number of individuals injected from each of the collection plants.

After injection, aphids were released into a petri dish onto a fava bean leaves that had stems in an Eppendorf tube filled with 1 ml water. Aphid survival was monitored daily, and dead aphids were  
10 removed when they were found. The number of offspring from each group was counted and fecundity was measured as the number of offspring (F1's) produced per aphid at each time point.

In select experiments, development was measured in groups of offspring from each treatment group by noting the developmental stage of offspring each day (1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> instar). Development was also measured by imaging aphids at 4 days post-collection and determining their area.

15 The template for synthesizing dsRNAs was the cDNA reverse-transcribed from the mRNA. RNA was extracted from one 5<sup>th</sup> instar *A. pisum* (LSR-1 strain) and was quantified by Nanodrop (Thermo fisher scientific). ~100 µg of total RNA was added as template into the reverse-transcription reaction using Superscript IV kit (Thermo Fisher Scientific) following manufacturer's protocol.

To amplify the templates for the dsRNAs, the cDNA was diluted 100-fold and used in the  
20 following PCRs. The PCR reactions (25 µl final volume) contain 12.5 µL of Go Taq Green 2X mix (Promega), 0.2 µl of forward primer (Table 12), 0.2 µl of reverse primer (Table 12), and 12.1 µl of 100-fold diluted cDNA. PCR reactions were performed using following conditions: 1) 95°C for 2 minutes, 2) 95°C for 20 seconds, 3) 55°C for 15 seconds, 4) 72°C for 30 seconds, 5) repeat steps 2-4 35x, 6) 72°C for 5 minutes. The sizes of PCR amplified products were verified by electrophoresis on 1.5% agarose and the  
25 expected-size bands were cut and purified by QIAquick DNA purification kit (Qiagen). The dsRNAs were synthesized *in vitro* using T7 MEGAscript kit (Ambion, Thermo Fisher Scientific; Cat# AM1334) following manufacturer's protocol. The concentration of dsRNA was measured by Nanodrop (Thermo Fisher Scientific).

30 **Table 12: Accession numbers and primers for dsRNA syntheses.**

| Gene        | gene ID     | Forward primer                                                   | Reverse primer                                                    |
|-------------|-------------|------------------------------------------------------------------|-------------------------------------------------------------------|
| ApGLNT<br>1 | ACYPI001018 | TAATACGACTCACTATAGGG<br>CAATTACAAAAGGACGGCAG<br>(SEQ ID NO: 111) | TAATACGACTCACTATAG<br>GGCCGCTCTAGGAACAC<br>CGTAT (SEQ ID NO: 112) |

At the indicated time point post-treatment, DNA and/or RNA was extracted from aphids in each treatment group. Briefly, the aphid body surface was sterilized by dipping the aphid into a 6% bleach solution for approximately 5 seconds. Aphids were then rinsed in sterile water and total DNA or RNA was  
35 extracted from each individual aphid using either the DNA or RNA extraction kit (Qiagen, DNeasy or miRNeasy kit, respectively) according to manufacturer's instructions. DNA and RNA concentrations were measured using a nanodrop nucleic acid quantification. *Buchnera* and aphid DNA copy numbers were

measured by qPCR. The primers used for *Buchnera* were Buch\_groES\_18F (CATGATCGTGTGCTTGTTAAG; SEQ ID NO: 101) and Buch\_groES\_98R (CTGTTCCCTCGAGTCGATTTCC; SEQ ID NO: 102) (Chong and Moran, 2016 PNAS). The primers used for aphid were ApEF1a 107F (CTGATTGTGCCGTGCTTATTG; SEQ ID NO: 103) and ApEF1a 246R (TATGGTGGTTCAGTAGAGTCC; SEQ ID NO: 104) (Chong and Moran, 2016 PNAS). qPCR was performed using a qPCR amplification ramp of 1.6 degrees C/s and the following conditions: 1) 95°C for 10 minutes, 2) 95°C for 15 seconds, 3) 55°C for 30 seconds, 4) repeat steps 2-3 40x, 5) 95°C for 15 seconds, 6) 55°C for 1 minute, 7) ramp change to 0.15 degrees C/s, 8) 95°C for 1 second. qPCR data was analyzed using analytic (Thermo Fisher Scientific, QuantStudio Design and Analysis) software.

ApGLNT1 relative expression was measured by RT-qPCR. The primers used for ApGLNT1 were ACYPI001018-fwd (CCTGAAATCGACGGGGTCC; SEQ ID NO: 113) and ACYPI001018-rev (AGATCGGCAACATCTGTTCGT; SEQ ID NO: 114) (both designed by NCBI pick primers). The internal reference gene was Actin (housekeeping gene). The primers used for Actin were Actin-F (GATCAGCAGCCACACACAAG; SEQ ID NO: 109) and Actin-R (TTTGAACCGGTTTACGACGA; SEQ ID NO: 110) RT-qPCR was performed using a qPCR amplification ramp of 1.6 degrees C/s and the following conditions: 1) 48°C for 30 min, 2) 95°C for 10 minutes, 3) 95°C for 15 seconds, 4) 60°C for 1 minute, 5) repeat steps 3-4 40x, 6) 95°C for 15 seconds, 7) 60°C for 1 minute, 8) ramp change to 0.15 degrees C/s, 9) 95°C for 1 second. RT-qPCR data was analyzed using analytic (Thermo Fisher Scientific, QuantStudio Design and Analysis) software.

#### *Microinjection with dsRNA knocked-down the bacteriocyte transporter gene expression in aphids*

The preliminary experiment assessed whether injecting dsRNA into aphids would result in decreased gene expression. Adult aphids were injected with 8 ng dsRNA or water (as a negative control). On two days post-injection, RNA was extracted from the remaining aphids in each treatment group and RT-qPCR was performed to quantify expression of ApGLNT1. Aphids microinjected with the negative control solution (water) had high relative expression of the ApGLNT1 gene. In contrast, aphid adults microinjected with the dsRNA of ApGLNT1 had a drastic, and significant, reduction of ApGLNT1 gene expression (Fig. 13), indicating that dsRNA microinjection treatment decreased expression of ApGLNT1 gene.

#### *Microinjection of dsRNA-ApGLNT1 resulted in increased aphid mortality*

To assess the effect of dsRNA-ApGLNT1 microinjection on insect fitness, LSR-1 fifth instar aphids were injected with 60 ng dsRNA and survival was monitored for 5 days. At 3 days post-injection, approximately 72% of water-injected aphids were alive (Fig. 14). In contrast, only 52% of dsRNA-injected aphids were alive (Fig. 14). At days 4 and 5 post-injection, there were significantly more aphids alive in the water-injected group compared to the dsRNA-injected group. Approximately 62% and 51% of water-injected aphids were alive on days 4 and 5, respectively, whereas only 30% and 12.5% of dsRNA-injected aphids were alive on day 4 and 5, respectively (Fig. 14). These data indicated that dsRNA against ApGLNT1 significantly ( $p=0.0004$ ) increased aphid mortality compared to water-injected controls.

### *Microinjection of dsRNA-ApGLNT1 resulted in decreased Buchnera titers*

To test whether the decrease in survival and fitness was due to decreased number of endosymbionts, DNA was extracted from living aphids at 5 days post-injection and qPCR was performed to quantify the amount of *Buchnera* present in the aphid. Aphids microinjected with water had mean ratios of approximately 11 *Buchnera*/aphid copies (Fig. 15). In contrast, *Buchnera*/aphid copies was approximately 1.28 times lower in aphids injected with the dsRNA against the ApGLNT1 (Fig. 15). These data showed that microinjection of dsRNA-ApGLNT1 resulted in decreased *Buchnera* titers which led to decreased aphid fitness.

### *Development was delayed in offspring of aphids microinjected with dsRNA-ApGLNT1*

On day 3 post-injection, 40 offspring (first instars) from each treatment group were transferred to their own artificial feeding system (a petri dish containing a fava bean leaf stem put into a 1.5 ml Eppendorf tube sealed with parafilm) and development was monitored over time. Overall, development was delayed in offspring taken from adults injected with dsRNA-ApGLNT1. By day 4 post offspring transfer, approximately 22.5% of offspring from water-injected aphids began reaching the 5<sup>th</sup> instar stage (Fig. 4A). In contrast, on day 4, only 7.5% of offspring from adults injected with dsRNA-ApGLNT1 reached the 5<sup>th</sup> instar stage (Fig. 16A). Additionally, aphid areas were measured on day 4 by imaging each aphid in each group. While the average size of offspring from water-injected adults was 0.55 mm<sup>2</sup>, offspring dsRNA-ApGLNT1-injected adults were significantly smaller ( $p=0.009$ ) and averaged 0.4 mm<sup>2</sup> (Fig. 16B). These data indicated that treatment of adults with dsRNA-ApGLNT1 resulted in offspring with severely delayed development.

### **Example 20: Production of transgenic plants expressing dsRNAs that target multiple pathways to destabilize insect-symbiont homeostasis through treatment of aphids with the transgenic plants**

This Example demonstrates the ability to genetically modify *Nicotiana tabacum* to produce dsRNA for delivery to aphids to affect insect-symbiont homeostasis. Genetic constructs that stably express dsRNA in *Nicotiana tabacum* are delivered to the plant using transgenic *Agrobacterium tumefaciens* that will carry the plasmid to the plant.

### *Experimental design*

Several genes will be targeted for knockdown in multiple pathways that are critical for the symbiotic relationship between the aphids and their obligate endosymbiotic bacteria, *Buchnera*. Specifically, the glutamine transporter of the bacteriocytes (GlnT1), ultrabithorax (Ubx), beta alanine synthase (bAS), and cactus (Cact) are targeted. GlnT1 is a glutamine transporter that is used for the import of glutamine into the bacteriocytes, and the downstream products of glutamine are essential for the synthesis of essential amino acids by *Buchnera*. Ubx is a gene involved in both the general development of the aphids, as well as the formation of the bacteriocytes which will house *Buchnera*. bAS is an aphid gene required to synthesize beta Alanine which is a precursor for the synthesis of vitamin B5 by *Buchnera*. Cact is the negative regulator of the Toll pathway, which is the primary immunity pathway in

aphids. Lowering the levels of Cact could upregulate the systemic immune responses leading to the dysregulation of the *Buchnera* levels.

*Generating a plasmid containing dsRNA expression cassette:*

5 A shuttle vector between *E. coli* and *A. tumefaciens* will be used to carry the dsRNA expression cassette, which includes the cauliflower mosaic virus 35S promoter (pCaMV 35S) upstream of the dsRNA expressing sequence. The dsRNA expressing sequence includes the sense sequence followed by the antisense sequence of a region of the target aphid gene connected by a small hairpin loop sequence (Fig. 17).

10 The dsRNA expression cassette will then be placed in a shuttle vector for *E. coli* and *A. tumefaciens* (Fig. 18). The plasmid will also contain both Kanamycin and Gentamycin resistance cassettes which can be used as selection markers. A transcription terminator will be placed after the dsRNA expression transcript to eliminate runaway transcription.

The dsRNA expressing sequences for various genes (GlnT1, Ubx, bAS, and Cact) can be introduced into the vector via Gibson assembly. First, the sense and the antisense amplicons with overhangs that match the neighboring regions in the plasmid will be generated using the appropriate primers in a PCR. Specifically, the left overhang of the sense strand will have region of homology (~30bp) to the pCaMV 35S, and the right overhang of the antisense strand will have region of homology (~30bp) to the 35S terminator region. The right overhang of the sense strand and the left overhang of the antisense strands will have overlapping pieces of each other with the inclusion of a small hairpin region (acacgt, SEQ ID NO: 115). The primer sequences to produce the amplicons are shown in Table 13.

**Table 13: List of primers to generate amplicons for the Gibson assembly to generate hairpin dsRNA sequences against *A. pisum* genes. Each target gene's region will be amplified as either sense or antisense with flanking regions that are homologous to the plasmid backbone for Gibson assembly.**

| Primer name           | Primer sequence                                                      |
|-----------------------|----------------------------------------------------------------------|
| ApGlnT1 sense fwd     | ctacaaatctatctctcctaggCAATTACAAAAGGACGGCAG (SEQ ID NO: 116)          |
| ApGlnT1 sense rev     | ttacaaactgggaagaacctggagacgtgCCGCTCTAGGAACACCGTAT (SEQ ID NO: 117)   |
| ApGlnT1 antisense fwd | ttacaaactgggaagaacctggaacacgtcCCGCTCTAGGAACACCGTAT (SEQ ID NO: 118)  |
| ApGlnT1 antisense rev | agaaactagagcttgatcgatcgttaattaaCAATTACAAAAGGACGGCAG (SEQ ID NO: 119) |
| ApUbx sense fwd       | ctacaaatctatctctcctaggTTTTACCGTCACAGGCATCA (SEQ ID NO: 120)          |
| ApUbx sense rev       | acgagtgcgaagtccctagccagacgtgTCGTGCTCGTTACCAAATGT (SEQ ID NO: 121)    |

|                        |                                                                        |
|------------------------|------------------------------------------------------------------------|
| ApUbx antisense fwd    | acgagtgcgaagtccttagccaacacgctTCGTGCTCGTTACCAAATGT (SEQ ID NO: 122)     |
| ApUbx antisense rev    | agaaactagagcttgatcgatcgtaattaaTTTTACCGTCACAGGCATCA (SEQ ID NO: 123)    |
| ApbAs sense fwd        | ctacaaatctatctctcctaggGGTGTCCACATCGAGACGTT (SEQ ID NO: 124)            |
| ApbAs sense rev        | aaaaaccacatacctcgagtggggacgtgtCATGACTCTGGCAGTTGAAGTT (SEQ ID NO: 125)  |
| ApbAs antisense fwd    | aaaaaccacatacctcgagtggggacacgtcCATGACTCTGGCAGTTGAAGTT (SEQ ID NO: 126) |
| ApbAs antisense rev    | agaaactagagcttgatcgatcgtaattaaGGTGTCCACATCGAGACGTT (SEQ ID NO: 127)    |
| Apcactus sense fwd     | ctacaaatctatctctcctaggGTCGTGTCGTGTCGTGTCGTAGT (SEQ ID NO: 128)         |
| Apcactus sense rev     | accaaaattgccttgacagtgaggacgtgtGCACGCACGGAAAACATTTA (SEQ ID NO: 129)    |
| Apcactus antisense fwd | accaaaattgccttgacagtgaggacacgtcGCACGCACGGAAAACATTTA (SEQ ID NO: 130)   |
| Apcactus antisense rev | agaaactagagcttgatcgatcgtaattaaGTCGTGTCGTGTCGTGTCGTAGT (SEQ ID NO: 131) |

**Table 14: List of primers to generate amplicons for the Gibson assembly to generate hairpin dsRNA sequences against *Myzus persicae* genes. Each target gene's region will be amplified as either sense or antisense with flanking regions that are homologous to the plasmid backbone for Gibson assembly.**

| Primer name           | Primer sequence                                                                |
|-----------------------|--------------------------------------------------------------------------------|
| MpGlnT1 sense fwd     | ctacaaatctatctctcctaggTTGGAAGGGATTGGTGTGTAATGCC (SEQ ID NO: 132)               |
| MpGlnT1 sense rev     | ttacaaactgggaagaacctggagacgtgTTCCAGGTTCTTCCCAGTTTGTAAGTAGATCG (SEQ ID NO: 133) |
| MpGlnT1 antisense fwd | ttacaaactgggaagaacctggaacacgtcTCCAGGTTCTTCCCAGTTTGTAAGTAGATCG (SEQ ID NO: 134) |
| MpGlnT1 antisense rev | agaaactagagcttgatcgatcgtaattaaTTGGAAGGGATTGGTGTGTAATGCC (SEQ ID NO: 135)       |
| MpUbx sense fwd       | ctacaaatctatctctcctaggTCGTGTGGAGCAAGTACAGCG (SEQ ID NO: 136)                   |
| MpUbx sense rev       | acgagtgcgaagtccttagccagacgtgtTGGCTAGGGACTTCAGCACTCG (SEQ ID NO: 137)           |
| MpUbx antisense fwd   | acgagtgcgaagtccttagccaacacgtcTGGCTAGGGACTTCAGCACTCG (SEQ ID NO: 138)           |

|                        |                                                                              |
|------------------------|------------------------------------------------------------------------------|
| MpUbx antisense rev    | agaaactagagcttgatcgatcgtaattaaTCGTGTGGAGCAAGTACAGCGG (SEQ ID NO: 139)        |
| MpbAs sense fwd        | ctacaaatctatctctcctaggGAGGAACTCCAAGTCCAGAGTCATG (SEQ ID NO: 140)             |
| MpbAs sense rev        | aaaaaccacatacctcgagtgaggacgtgtCCCACTCGAGGTATGTGGTTTTTCCTATG (SEQ ID NO: 141) |
| MpbAs antisense fwd    | aaaaaccacatacctcgagtgaggacacgtcCCCACTCGAGGTATGTGGTTTTTCC (SEQ ID NO: 142)    |
| MpbAs antisense rev    | agaaactagagcttgatcgatcgtaattaaGAGGAACTCCAAGTCCAGAGTCATG (SEQ ID NO: 143)     |
| Mpcactus sense fwd     | ctacaaatctatctctcctaggTACACCCATTGTGTGCACCTGAGTAC (SEQ ID NO: 144)            |
| Mpcactus sense rev     | accaaattgccttgacagtgaggacgtgtCCACTGTCCAAGGCAATTTTGGTTG (SEQ ID NO: 145)      |
| Mpcactus antisense fwd | accaaattgccttgacagtgaggacacgtcCCACTGTCCAAGGCAATTTTGGTTG (SEQ ID NO: 146)     |
| Mpcactus antisense rev | agaaactagagcttgatcgatcgtaattaaTACACCCATTGTGTGCACCTGAGTAC (SEQ ID NO: 147)    |

Once the sense and the antisense amplicons are generated, the plasmid will be double digested to generate overhangs at the end of the pCaMV 35S and the start of the 35S terminator. The digested plasmid, sense, and the antisense amplicons will be assembled together in a Gibson assembly using Gibson assembly kits (such as SGI Gibson Assembly kit). *E. coli* (DH5 $\alpha$ , electrocompetent cells) will then be transformed with the plasmid and grown on LB plates containing 50mg/ml kanamycin. Resistant colonies will be used to harvest the plasmid containing the dsRNA expression cassette. The plasmids will be named pGlnT1dsRNA, pUbxdsRNA, pbASdsRNA, and pCactdsRNA for each of the four different inserts. These plasmids will be used to transform *A. tumefaciens* via electroporation. Upon selection on LB medium containing Kanamycin (50mg/ml) and Gentamycin (50mg/ml), resistant colonies will be isolated and maintained on selection plates.

#### *Transformed A. tumefaciens infiltrated into N. tabacum:*

Transformed *A. tumefaciens* were grown overnight in LB medium with the selection antibiotics until the OD 600 was 0.6. The cells were pelleted down, resuspended in infiltration medium (10mM MES, 150  $\mu$ M acetosyringone, and 10 mM MgCl<sub>2</sub>, pH 5.5), and adjusted to an OD<sub>600</sub> of 0.6. The cells were incubated at room temperature for 2-4 hours. The cell suspension was infiltrated into healthy *N. tabacum* leaves. The infiltration process was achieved by placing the blunt open end of a 1 ml syringe on the underside of the leaf and forcing the cell suspension into the leaf. The infiltrated areas were easily distinguished from the untreated areas and were clearly demarcated using a marker. The plants were then covered to create a high moisture environment for 24h.

The leaves expressed green fluorescent protein and were visualized under an epi fluorescence microscope after 1-2 days, see Fig. 19. The genetic material that was transferred to the plant by the *A. tumefaciens* contained a green fluorescence (GFP) expression cassette that was driven by an ubiquitin

promoter that was active in *N. tabacum*. The expression of GFP will be used a proxy for the expression of dsRNA.

*Generating a stable clone of N. tabacum expressing dsRNA*

5           The infiltrated leaves that show expression of GFP will be isolated, and the regions of the leaf near the midrib that express GFP very strongly will be cutout. The leaf disks will then be thoroughly sterilized using a sterilization solution (2% hypochlorite, 0.01% tween 20) for 10 min by agitation. The sterile leaf disks will be placed on petri dishes containing shooting medium (2.15 g/l Murashige and Skoog salts (without IAA, kinetin or sucrose), 0.8% (w/v) agar, 3.0% (w/v) sucrose, 0.1 mg/l indole butyric acid, 0.8 mg/l 6-benzylaminopurine, 0.1 mg/l Carbenicillin, 0.2 mg/l Ticarcillin/Clavulanic acid and suitable selection for the binary vector carrying your FP fusion construct) at 25°C, 16h:8h light:dark cycle till shoots appear. The new shoots will then be transplanted into plates containing rooting medium (2.15 g/l Murashige and Skoog salts, 0.8% (w/v) agar, 3.0%(w/v) sucrose, 0.5mg/l indole butyric acid, 0.1 mg/l carbenicillin, 0.2 mg/l Ticarcillin/Clavulanic acid) at 25°C, 16h:8h light:dark cycle till roots appear. These new plants will be transferred to phytatrays to develop larger roots so that they can be transferred to soil. The expression of the GFP will be tested in all new clones to ensure stable expression.

*Treating aphids with dsRNA by rearing them on N. tabacum expressing dsRNA*

20           Aphids will be grown on 10-week-old *N. tabacum* plants in a climate controlled incubator (16h:8h light:dark cycle, 60% humidity, 25°C). To limit maternal effects or health differences between plants, 5–10 adults from different plants will be distributed among 10 two-week-old plants and allowed to multiply to high density for 5–7 days. For experiments, first and second instar aphids will be collected from healthy plants and divided into two different treatment groups: 1) those allowed to feed on leaves that express the dsRNA, and 2) those allowed to feed on control leaves that do no express the dsRNA.

25           For each feeding experiment, leaves will be taken from the plant and placed in a 1.5 ml Eppendorf sealed with parafilm. The leaf stem will be placed into a deep petri dish (Fisher Scientific, Cat# FB0875711) and aphids will be applied to the leaves of the plant and allowed to feed. Old leaves will be replaced with new, freshly injected leaves every 2-3 days throughout the experiment. For each treatment, 60 aphids will be placed onto each leaf. Aphids will be monitored daily for survival and dead aphids will be removed when they were discovered. In addition, the developmental stage (1st, 2nd, 3rd, 4th, and 5th instar) will be determined daily throughout the experiment.

35           After 5 and 6 days of treatment, DNA will be extracted from dead aphids from each treatment group. Briefly, the aphid body surface will be sterilized by dipping the aphid into a 6% bleach solution for approximately 5 seconds. Aphids will then be rinsed in sterile water and DNA will be extracted from each individual aphid using a DNA extraction kit (Qiagen, DNeasy kit) according to manufacturer's instructions. DNA concentration will be measured using a nanodrop nucleic acid quantification, and *Buchnera* and aphid DNA copy numbers will be measured by qPCR. The primers that will be used for *Buchnera* are Buch\_groES\_18F (CATGATCGTGTGCTTGTTAAG; SEQ ID NO: 101) and Buch\_groES\_98R (CTGTTCCCTCGAGTCGATTTCC; SEQ ID NO: 102) (Chong and Moran, 2016 PNAS). The primers that will be used for aphid are ApEF1a 107F (CTGATTGTGCCGTGCTTATTG; SEQ ID NO: 103) and ApEF1a

246R (TATGGTGGTTCAGTAGAGTCC; SEQ ID NO: 104) (Chong and Moran, 2016 PNAS). qPCR will be performed using a qPCR amplification ramp of 1.6 degrees C/s and the following conditions: 1) 95°C for 10 minutes, 2) 95°C for 15 seconds, 3) 55°C for 30 seconds, 4) repeat steps 2-3 40x, 5) 95°C for 15 seconds, 6) 55°C for 1 minute, 7) ramp change to 0.15 degrees C/s, 8) 95°C for 1 second. qPCR data will be analyzed using analytic (Thermo Fisher Scientific, QuantStudio Design and Analysis) software.

At 7 post-treatment, RNA will be extracted from live aphids and RT-pPCR will be performed to quantify expression of BCR-4. Briefly, the aphid body surface will be sterilized by dipping the aphid into a 6% bleach solution for approximately 5 seconds. Aphids will then be rinsed in sterile water and total RNA will be extracted from each individual aphid using the RNA extraction kit (Qiagen miRNeasy kit) according to manufacturer's instructions. RNA concentrations will be measured using a nanodrop nucleic acid quantification. BCR-4 relative expression will be measured by RT-qPCR. The primers used will be ApBCR-4F (CTCTGTCAACCACCATGAGATTA; SEQ ID NO: 107) and ApBCR-4R (TGCAGACTACAGCACAATACTT; SEQ ID NO: 108). The internal reference gene primers were for Actin (housekeeping gene). The forward sequence is GATCAGCAGCCACACACAAG; SEQ ID NO: 109 and the reverse sequence is TTTGAACCGGTTTACGACGA; SEQ ID NO: 110. RT-qPCR will be performed using a qPCR amplification ramp of 1.6 degrees C/s and the following conditions: 1) 48°C for 30 min, 2) 95°C for 10 minutes, 3) 95°C for 15 seconds, 4) 60°C for 1 minute, 5) repeat steps 3-4 40x, 6) 95°C for 15 seconds, 7) 60°C for 1 minute, 8) ramp change to 0.15 degrees C/s, 9) 95°C for 1 second. RT-qPCR data was analyzed using analytic (Thermo Fisher Scientific, QuantStudio Design and Analysis) software.

As shown in previous Examples, the aphids fed on the leaves expressing the dsRNA against the aphid target genes are expected to have lower survival, develop slower, contain fewer *Buchnera*, and have reduced target gene expression compared to the aphids reared on the control leaves.

Together this data described herein demonstrate the ability to kill and decrease the development and longevity (e.g., fitness) of aphids by treating them with plants expressing dsRNA targeting the essential gene(s) (e.g. glutamine transporter ApGLNT1) of bacteriocytes in the aphids.

### OTHER EMBODIMENTS

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, the descriptions and examples should not be construed as limiting the scope of the invention. The disclosures of all patent and scientific literature cited herein are expressly incorporated in their entirety by reference.

## CLAIMS

1. A method for decreasing the fitness of an insect, the method comprising delivering to the insect an effective amount of a polynucleotide comprising a dsRNA that decreases expression of an insect bacteriocyte regulatory gene or an insect immunoregulatory gene in the insect relative to an insect that has not been administered the dsRNA.
2. The method of claim 1, wherein the gene encodes a protein selected from the group consisting of a protein from the bacteriocyte-specific cysteine rich proteins BCR family, a protein from the secreted proteins SP family, BicD (Protein bicaudal D), Cact (cactus), DIF (Dorsal related immunity factor), Toll (Toll Interacting Protein), and imd (immune deficiency protein).
3. The method of claim 1 or 2, wherein the gene encodes a protein having at least 90% amino acid sequence identity, at least 95% amino acid sequence identity, or 100% amino acid sequence identity to a protein listed in Table 5, Table 8, or Table 9.
4. The method of any one of claims 1-3, wherein the dsRNA is complementary to 10 to 30 nucleotides of the gene in the insect.
5. The method of any one of claims 1-4, wherein the method is effective to inhibit expression of the gene in the insect.
6. The method of any one of claims 1-5, wherein the method is effective to decrease the level, diversity, or metabolism of one or more microorganisms resident in the insect relative to an insect that has not been delivered the agent.
7. The method of claim 6, wherein the one or more microorganisms is a *Buchnera* spp.
8. The method of any one of claims 1-7, wherein the method is effective to decrease the fitness of the insect relative to an insect that has not been delivered the agent.
9. The method of any one of claims 1-8, wherein the polynucleotide is delivered in a composition formulated for delivery to insects.
10. The method of any one of claims 1-9, wherein the delivery comprises delivering the polynucleotide to at least one habitat where the insect pest grows, lives, reproduces, feeds, or infests.
11. The method of any one of claims 1-10, wherein the delivery comprises spraying the polynucleotide on an agricultural crop.

12. The method of any one of claims 1-11, wherein the polynucleotide is delivered as an insect comestible composition for ingestion by the insect.

13. The method of any one of claims 1-12, wherein the polynucleotide is formulated with an agriculturally acceptable carrier as a liquid, a solid, an aerosol, a paste, a gel, or a gas composition.

14. The method of any one of claims 1-13, wherein the insect is an aphid.

15. A composition comprising a polynucleotide comprising a dsRNA formulated for delivery to an insect, wherein the dsRNA is complementary to 15 to 30 nucleotides of an insect bacteriocyte regulatory gene or an insect immunoregulatory gene.

16. The composition of claim 15, wherein the gene encodes a protein selected from the group consisting of a protein from the bacteriocyte-specific cysteine rich proteins BCR family, a protein from the secreted proteins SP family, BicD (Protein bicaudal D), Cact (cactus), DIF (Dorsal related immunity factor), Toll (Toll Interacting Protein), and imd (immune deficiency protein).

17. The composition of claim 15 or 16, wherein the gene encodes a protein having at least 90% amino acid sequence identity, 95% amino acid sequence identity, or 100% amino acid sequence identity to a protein listed in Table 5, Table 8, or Table 9.

18. A plant comprising a topical application of the composition of any one of claims 15-17.

19. A transgenic plant cell having in its genome a recombinant DNA construct, wherein the recombinant DNA construct comprises a heterologous promoter operably linked to a DNA encoding a RNA comprising at least one double-stranded RNA region, at least one strand of which comprises a nucleotide sequence that is complementary to 15 to 30 nucleotides of an insect bacteriocyte regulatory gene or an insect immunoregulatory gene.

20. The transgenic plant of claim 19, wherein the gene encodes a protein selected from the group consisting of a protein from the bacteriocyte-specific cysteine rich proteins BCR family, a protein from the secreted proteins SP family, BicD (Protein bicaudal D), Cact (cactus), DIF (Dorsal related immunity factor), Toll (Toll Interacting Protein), and imd (immune deficiency protein).

21. The transgenic plant of claim 19 or 20, wherein the gene encodes a protein having at least 90% amino acid sequence identity, 95% amino acid sequence identity, or 100% amino acid sequence identity to a protein listed in Table 5, Table 8, or Table 9.

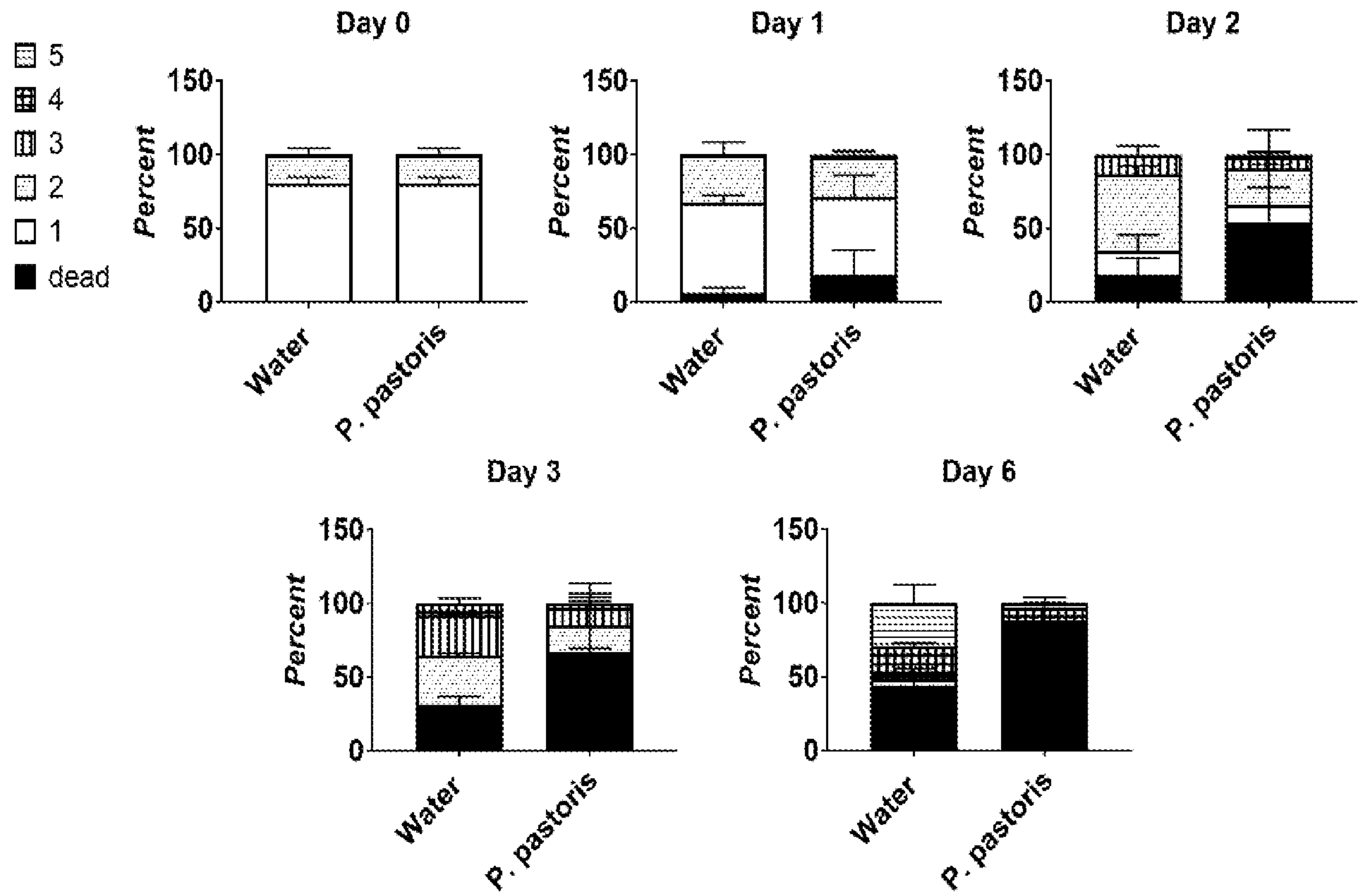


FIG. 1

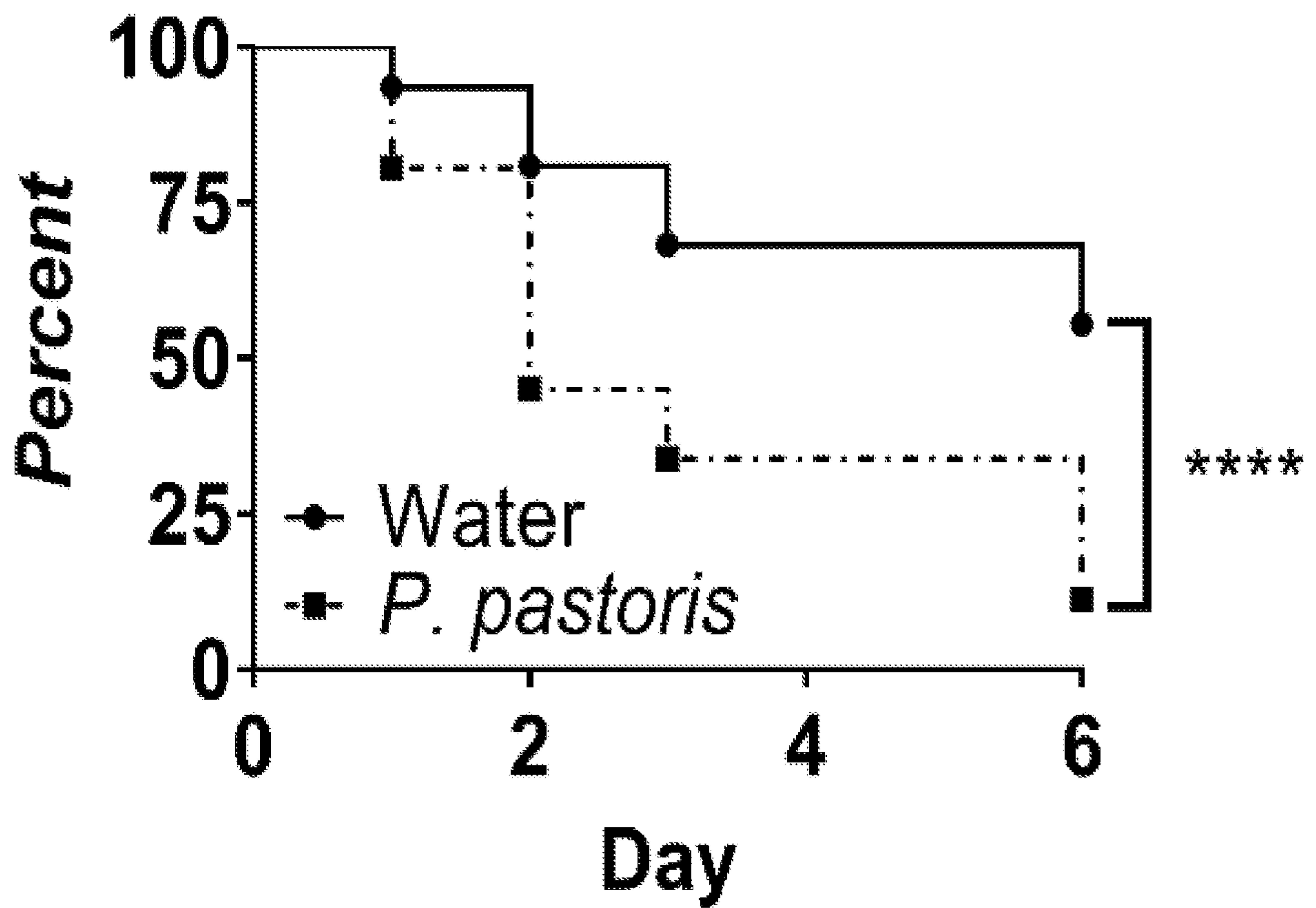


FIG. 2

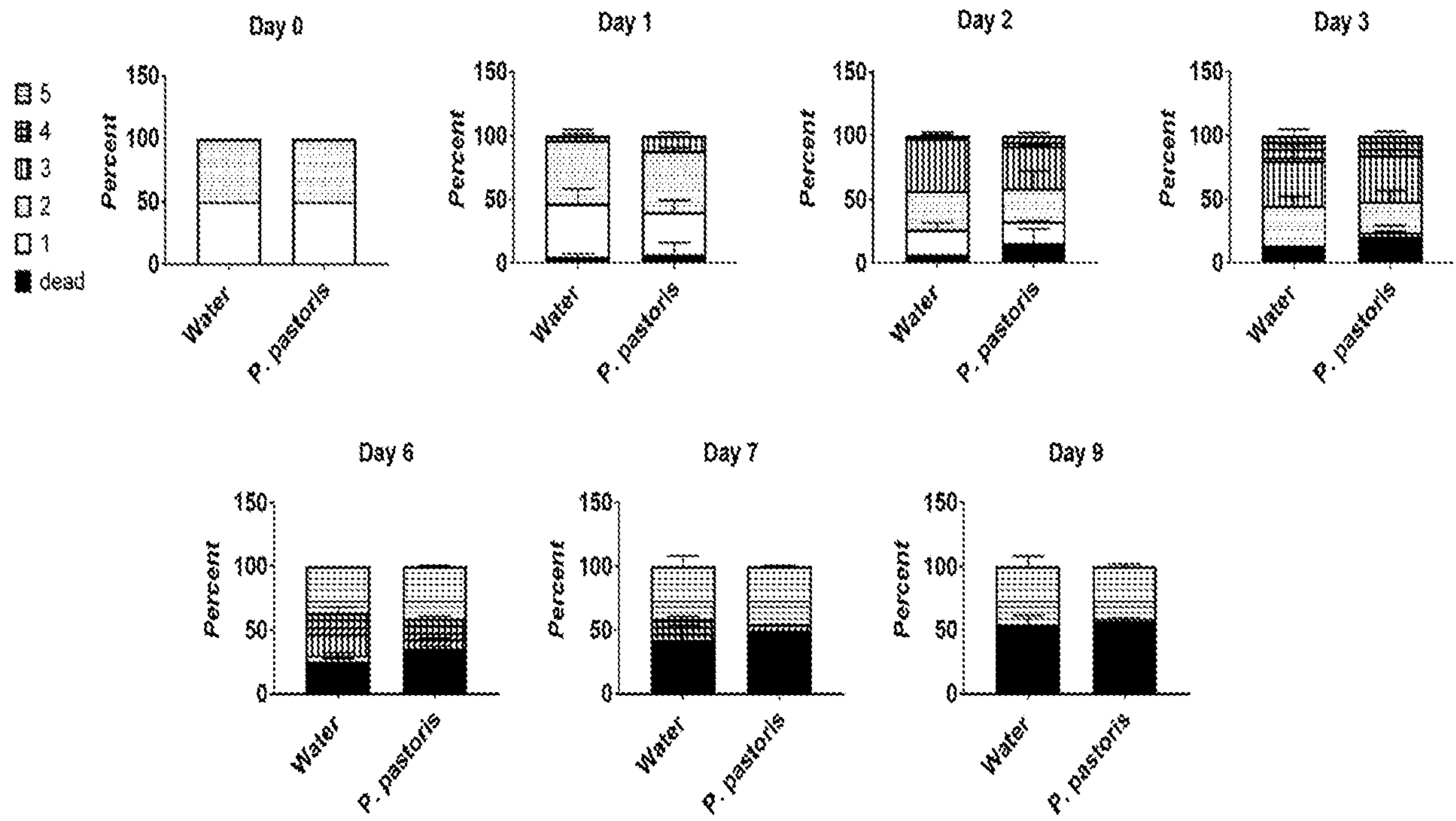


FIG. 3

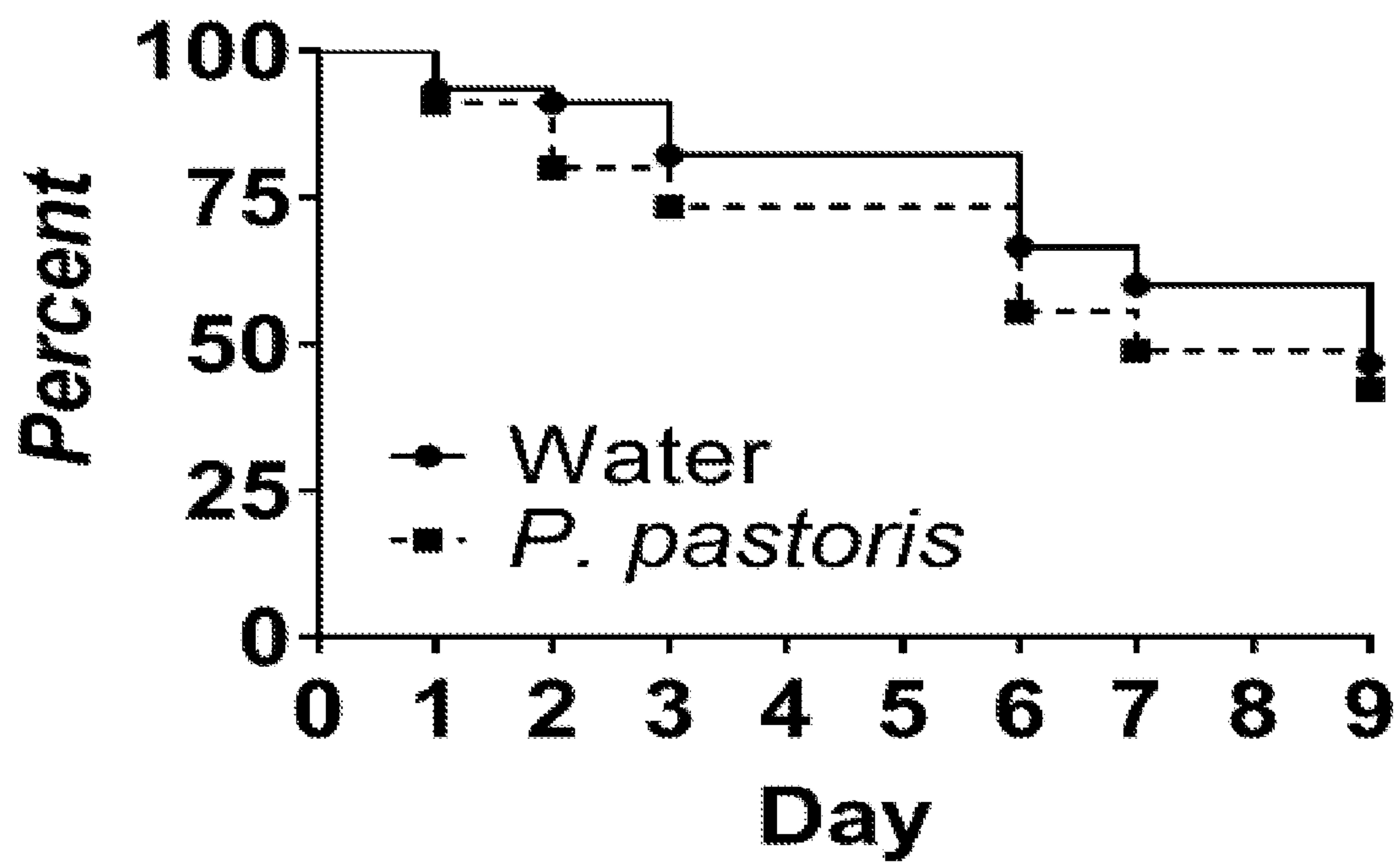


FIG. 4

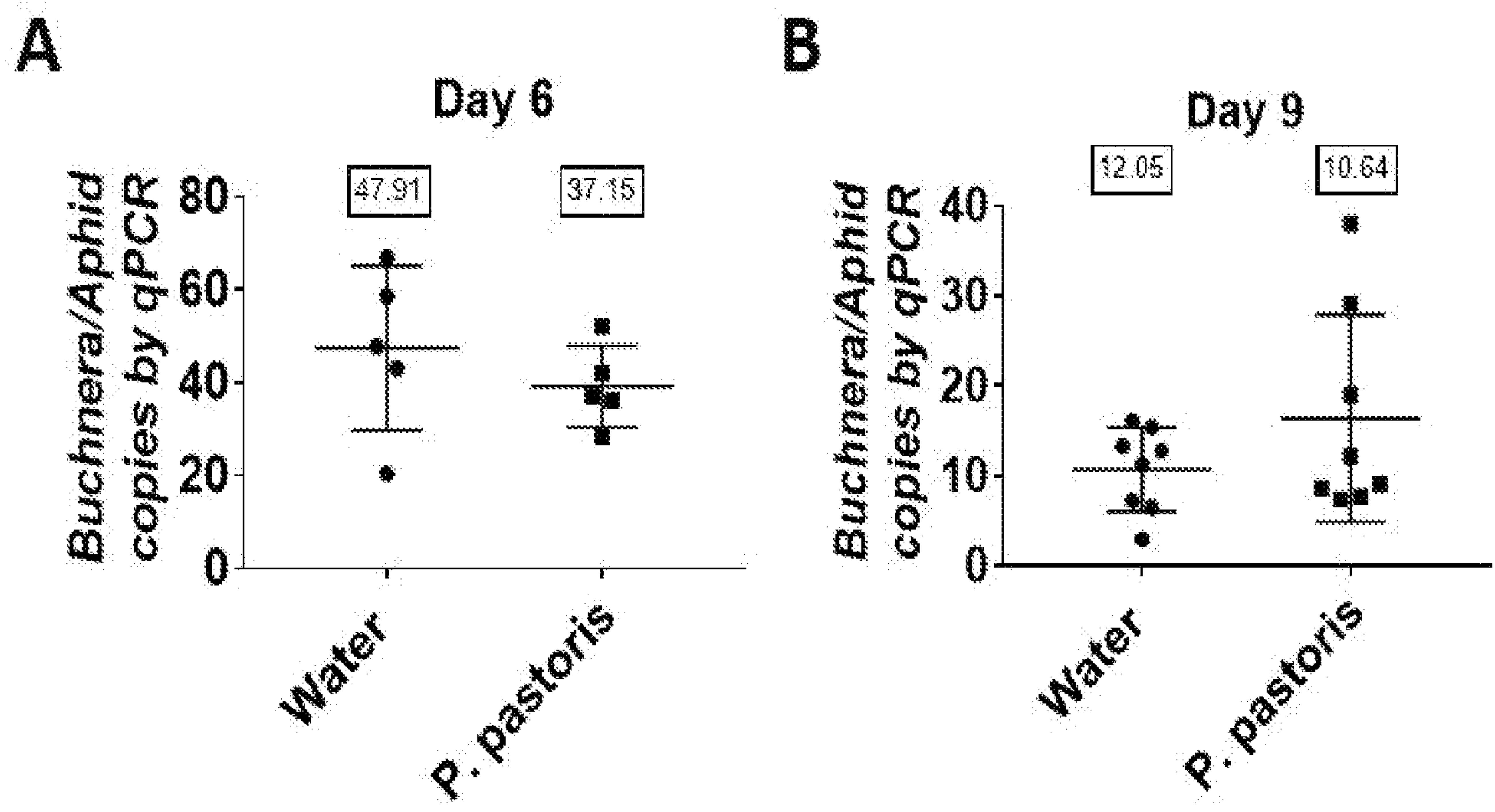


FIG. 5A-5B

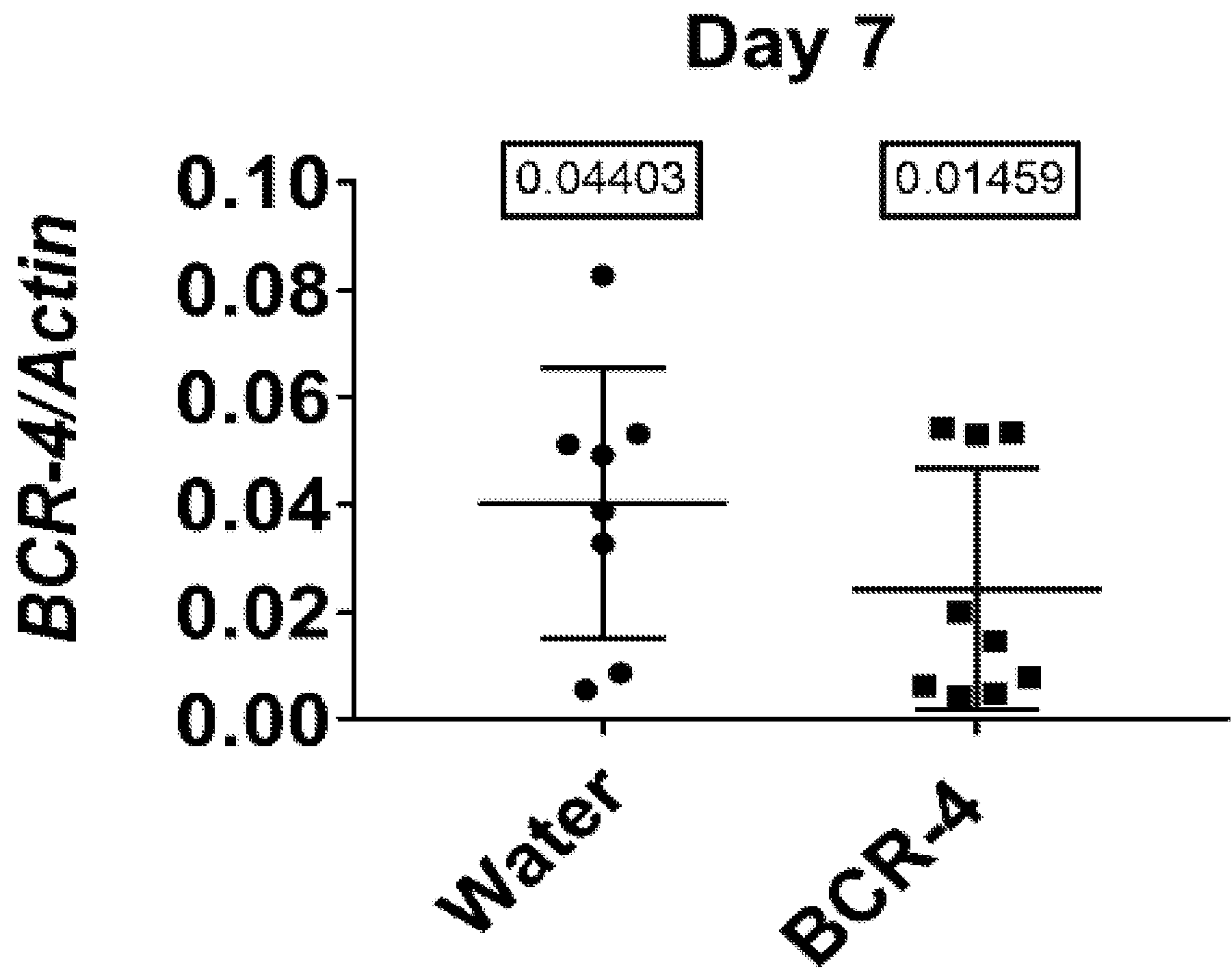


FIG. 6

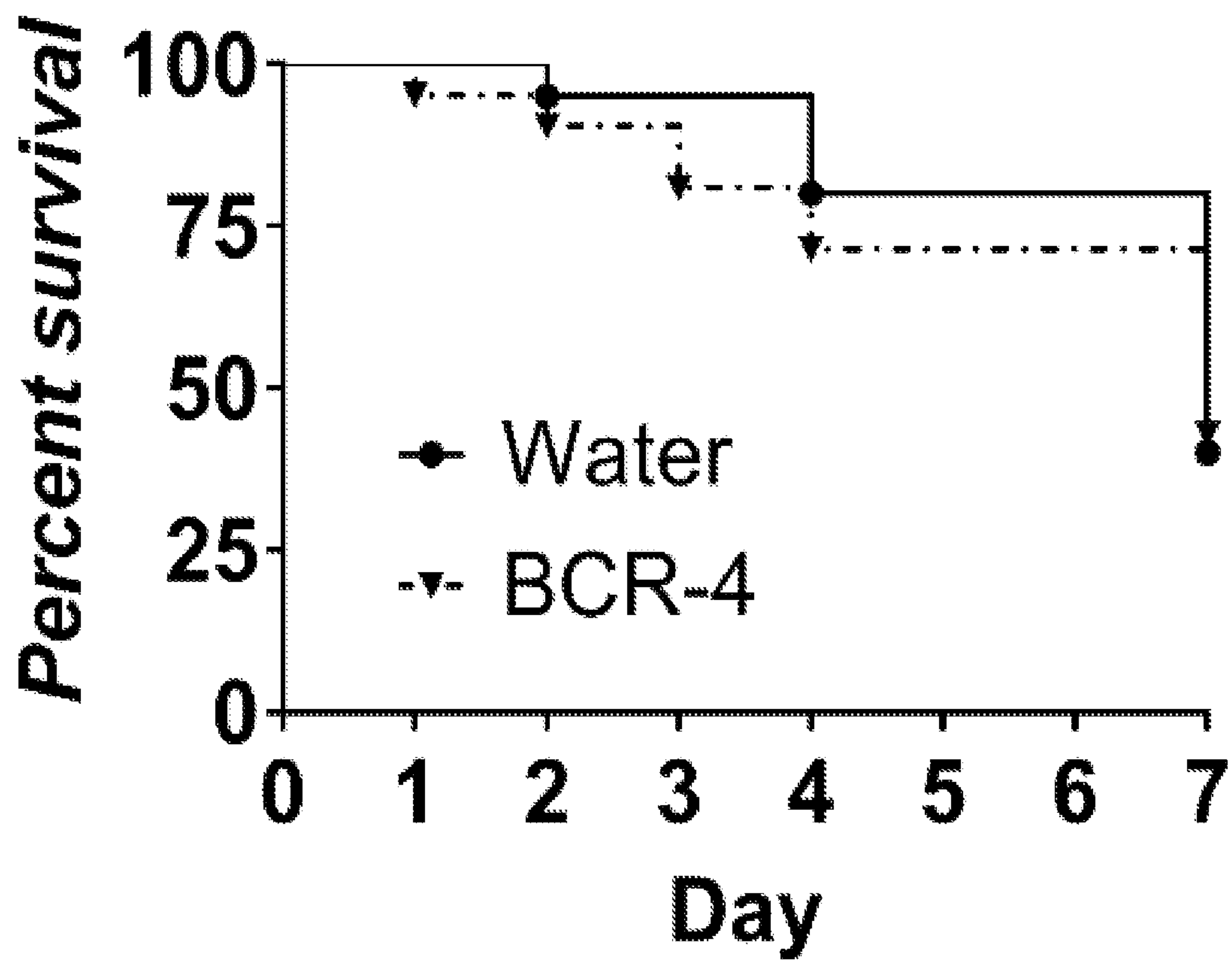


FIG. 7

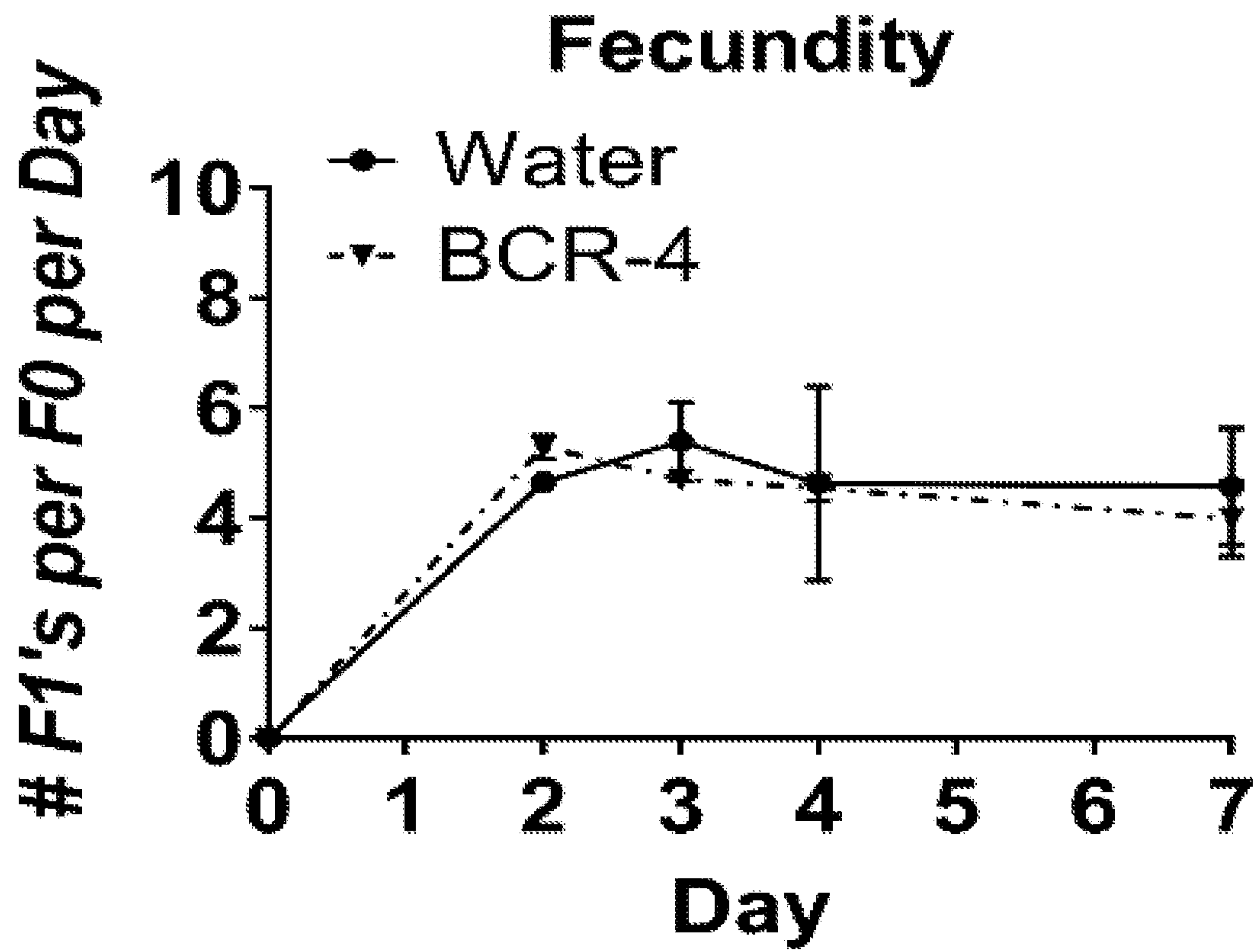


FIG. 8

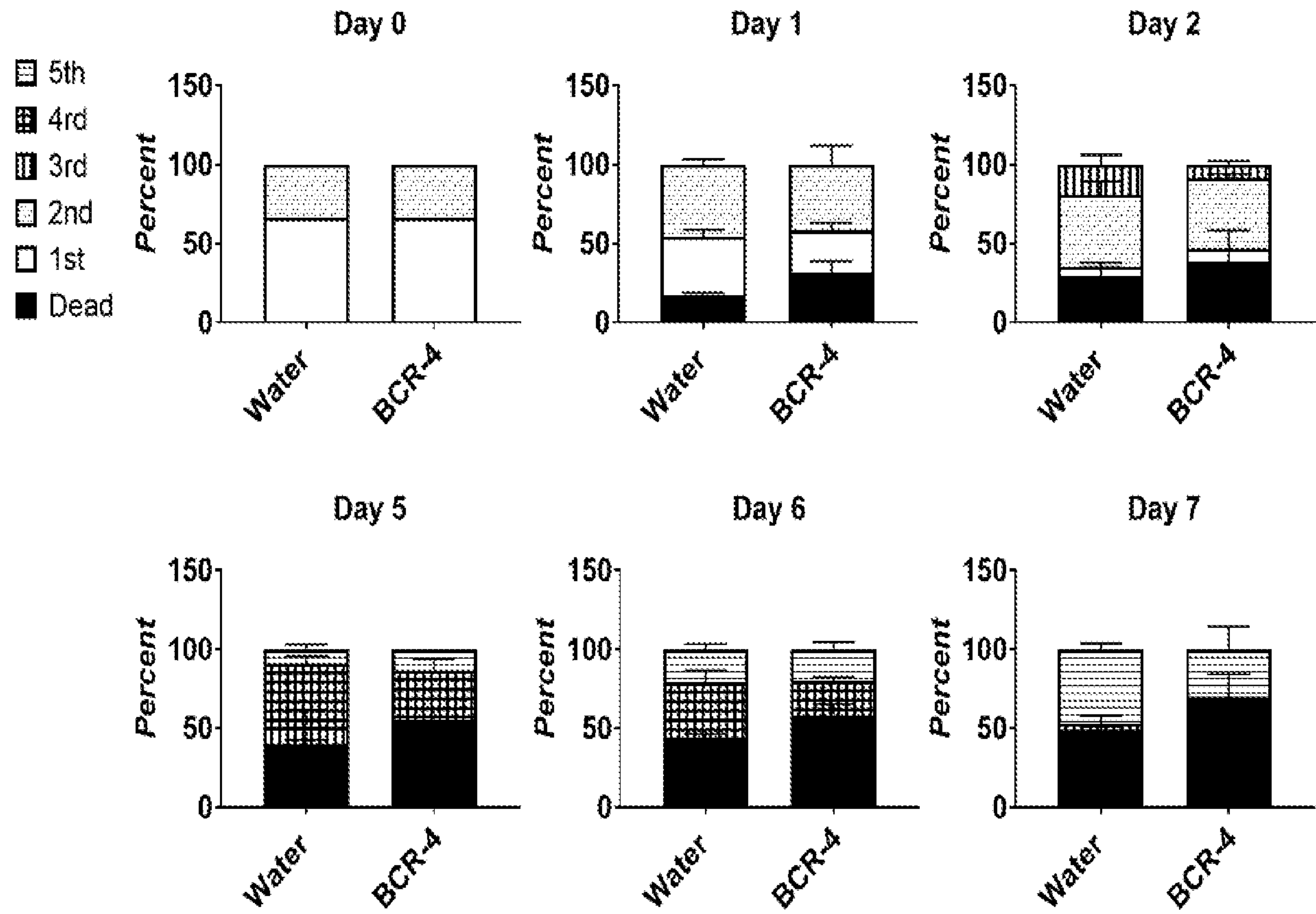


FIG. 9

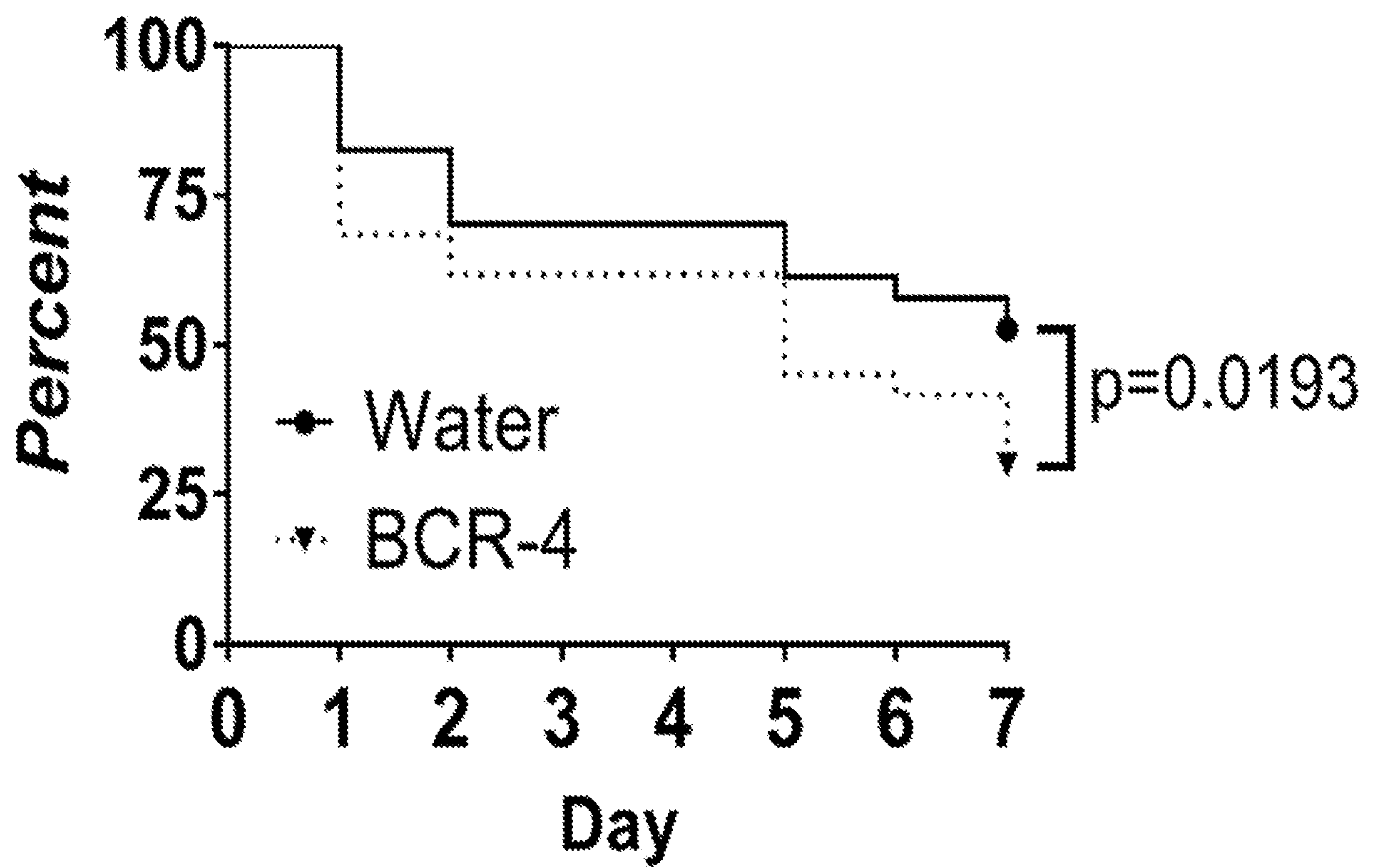


FIG. 10

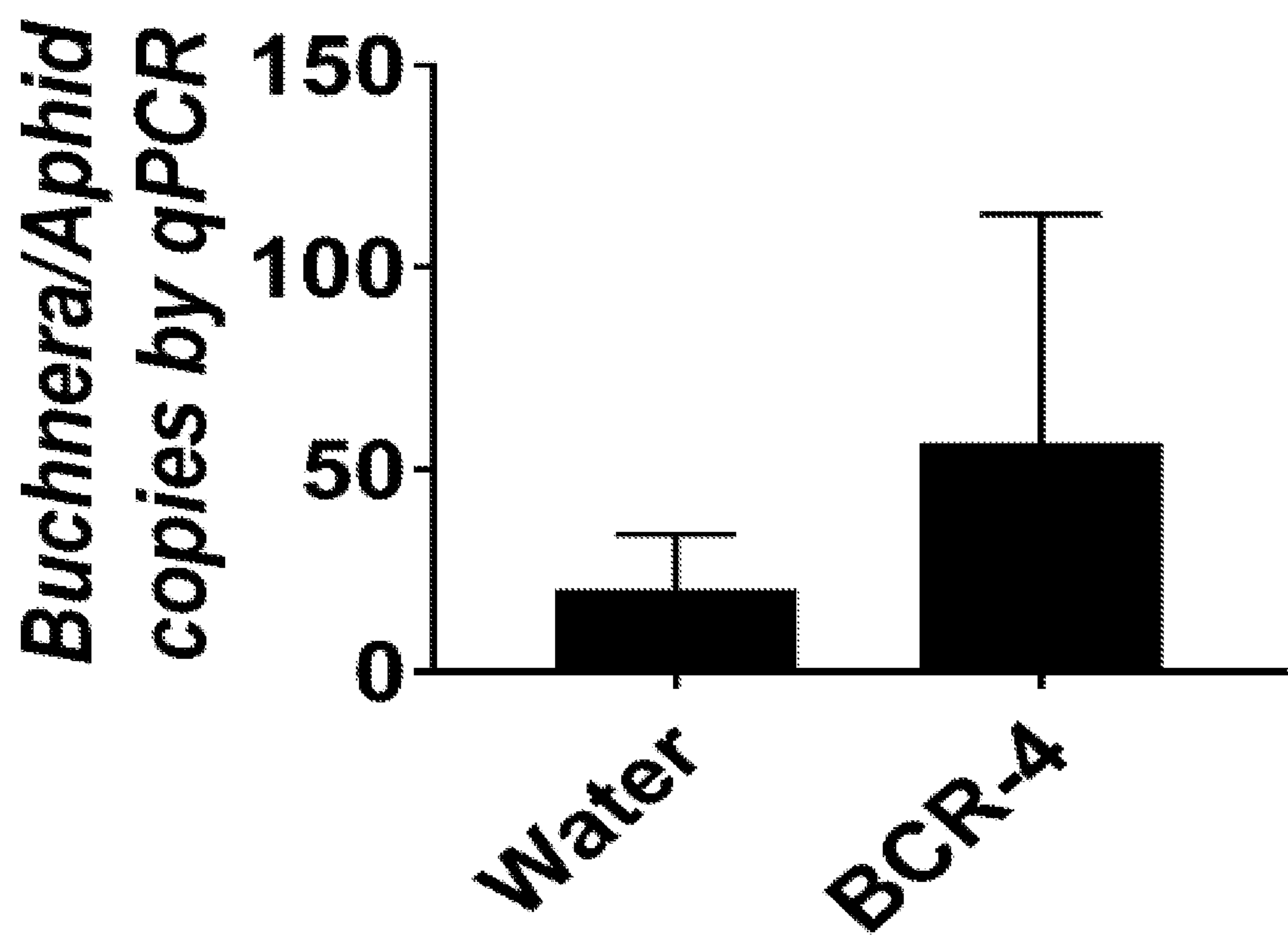


FIG. 11

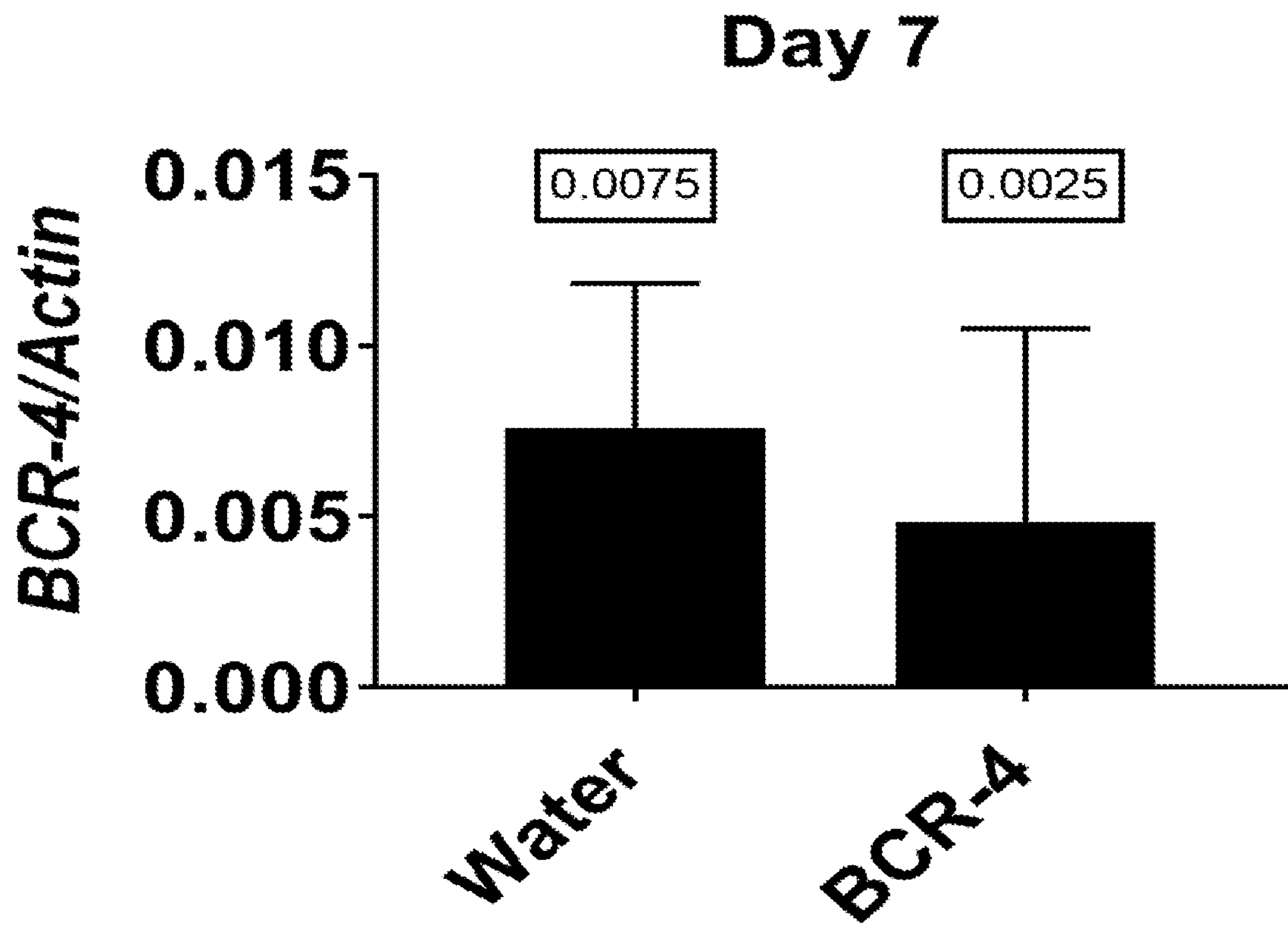


FIG. 12

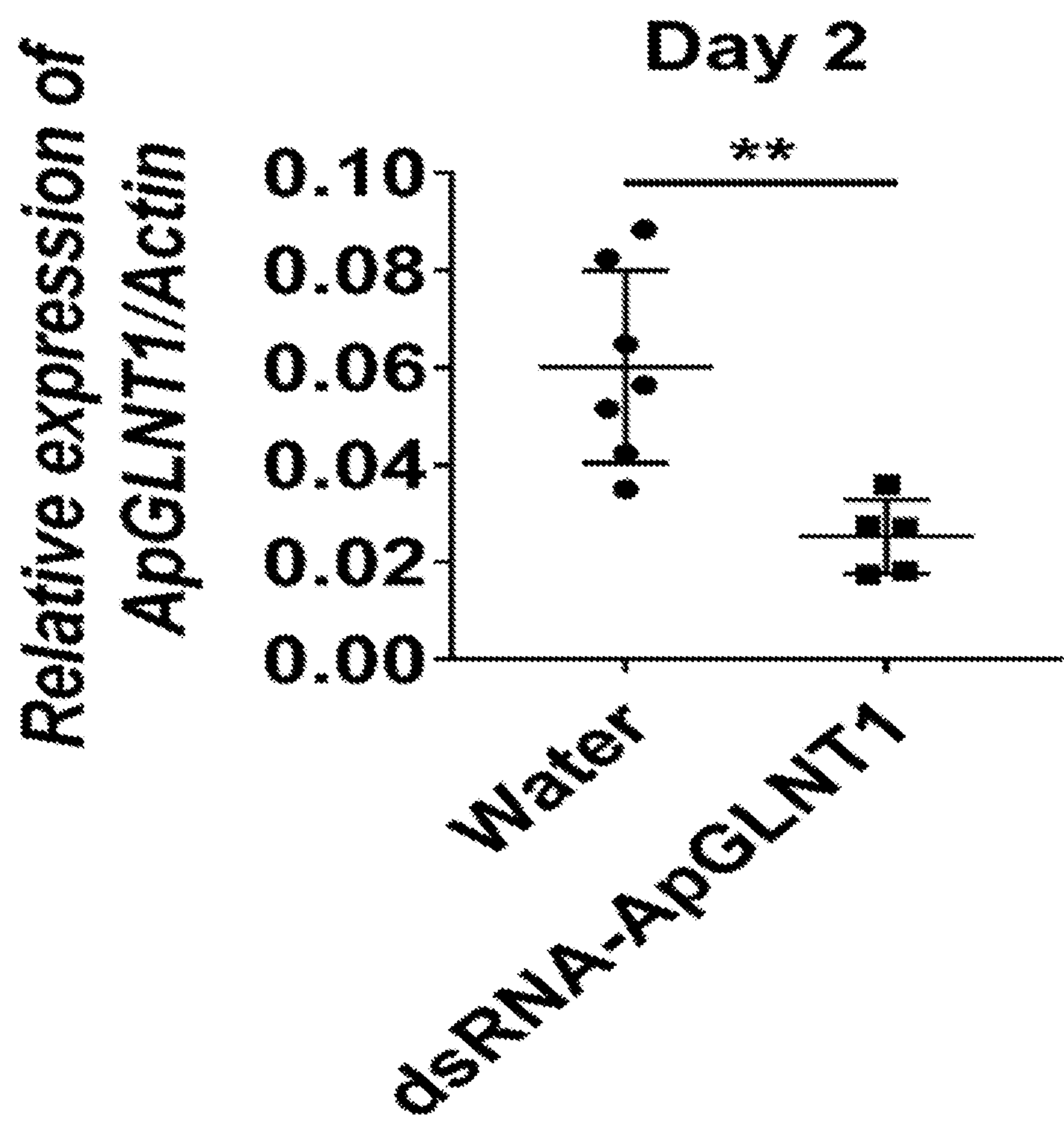


FIG. 13

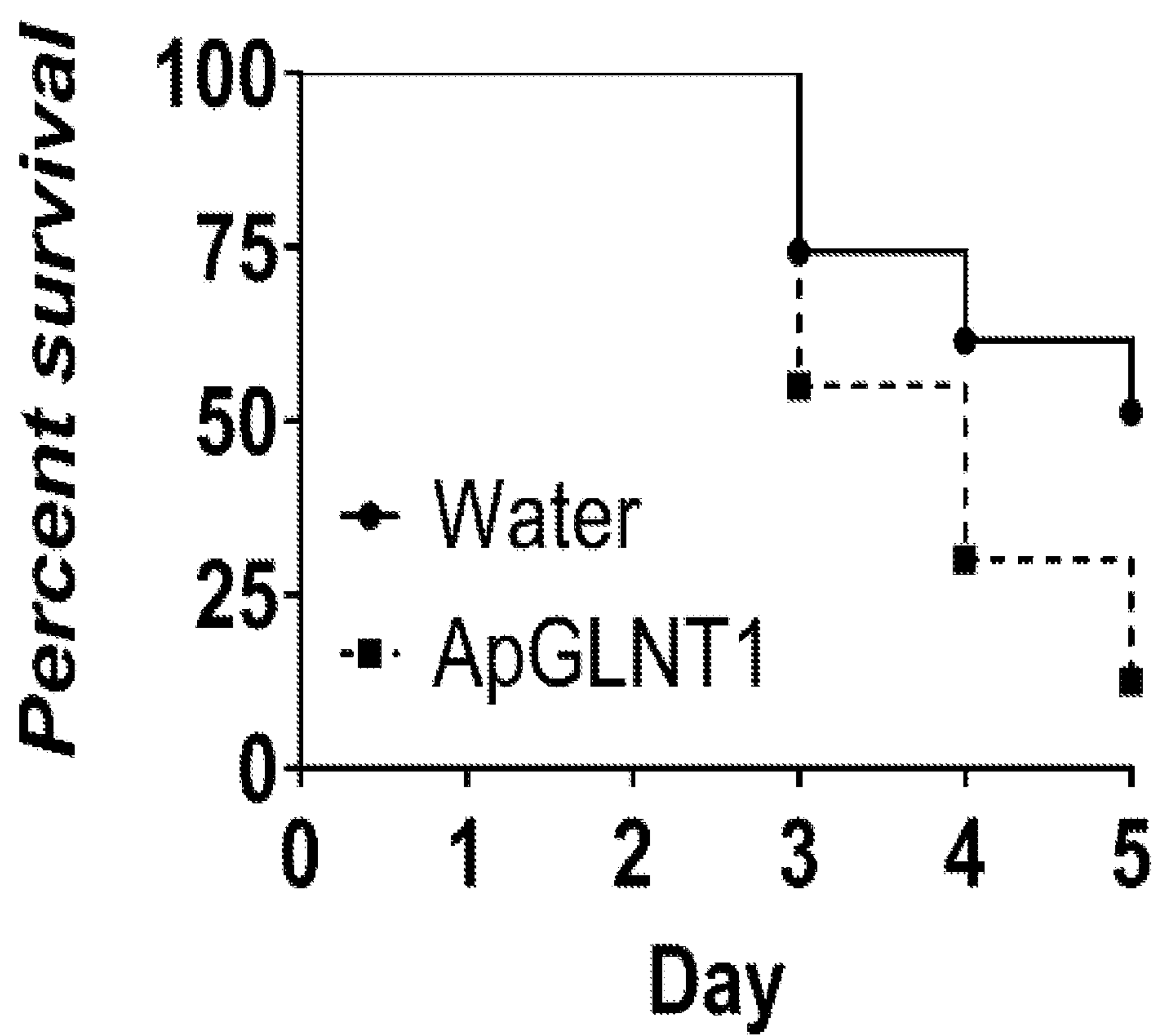


FIG. 14

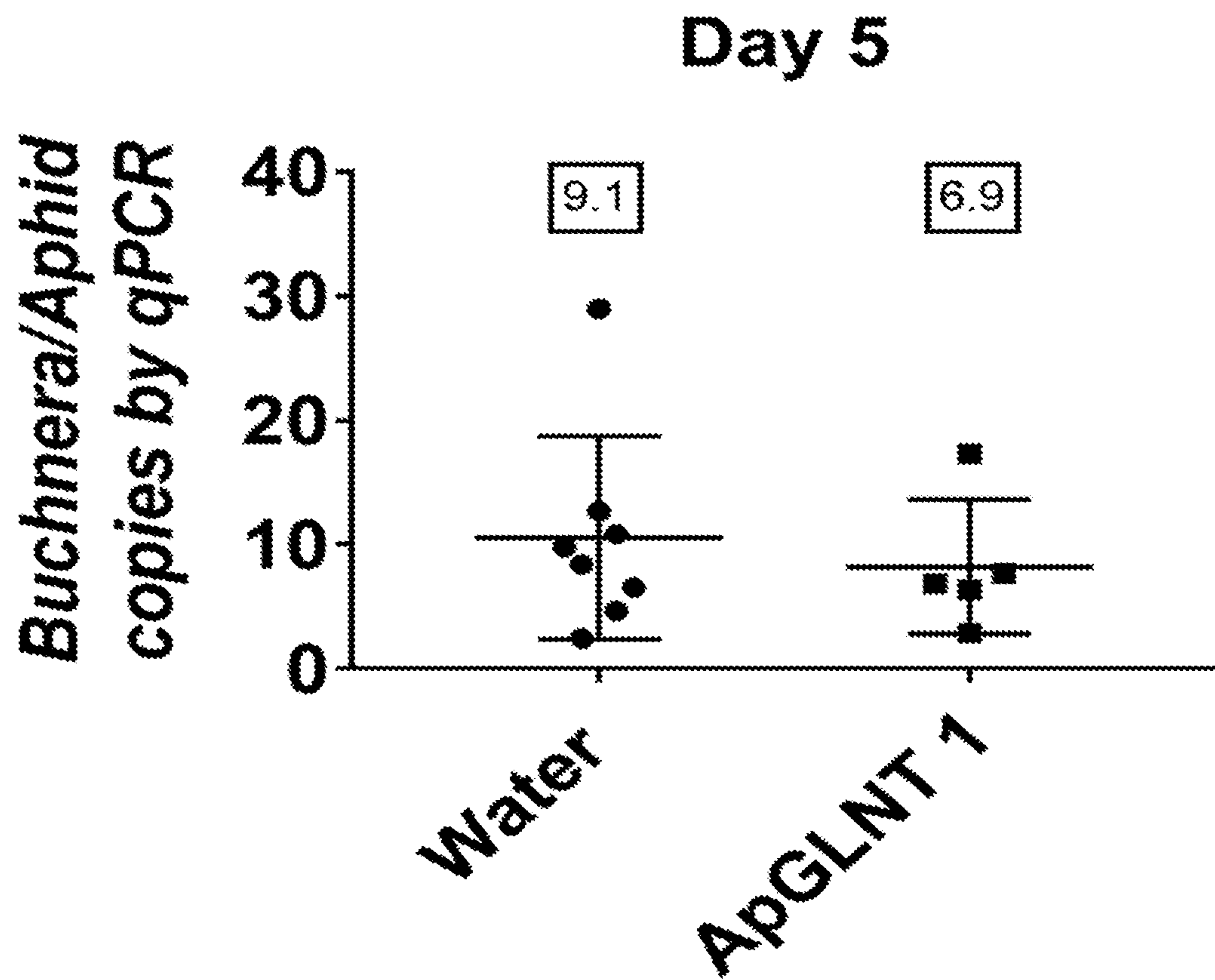


FIG. 15

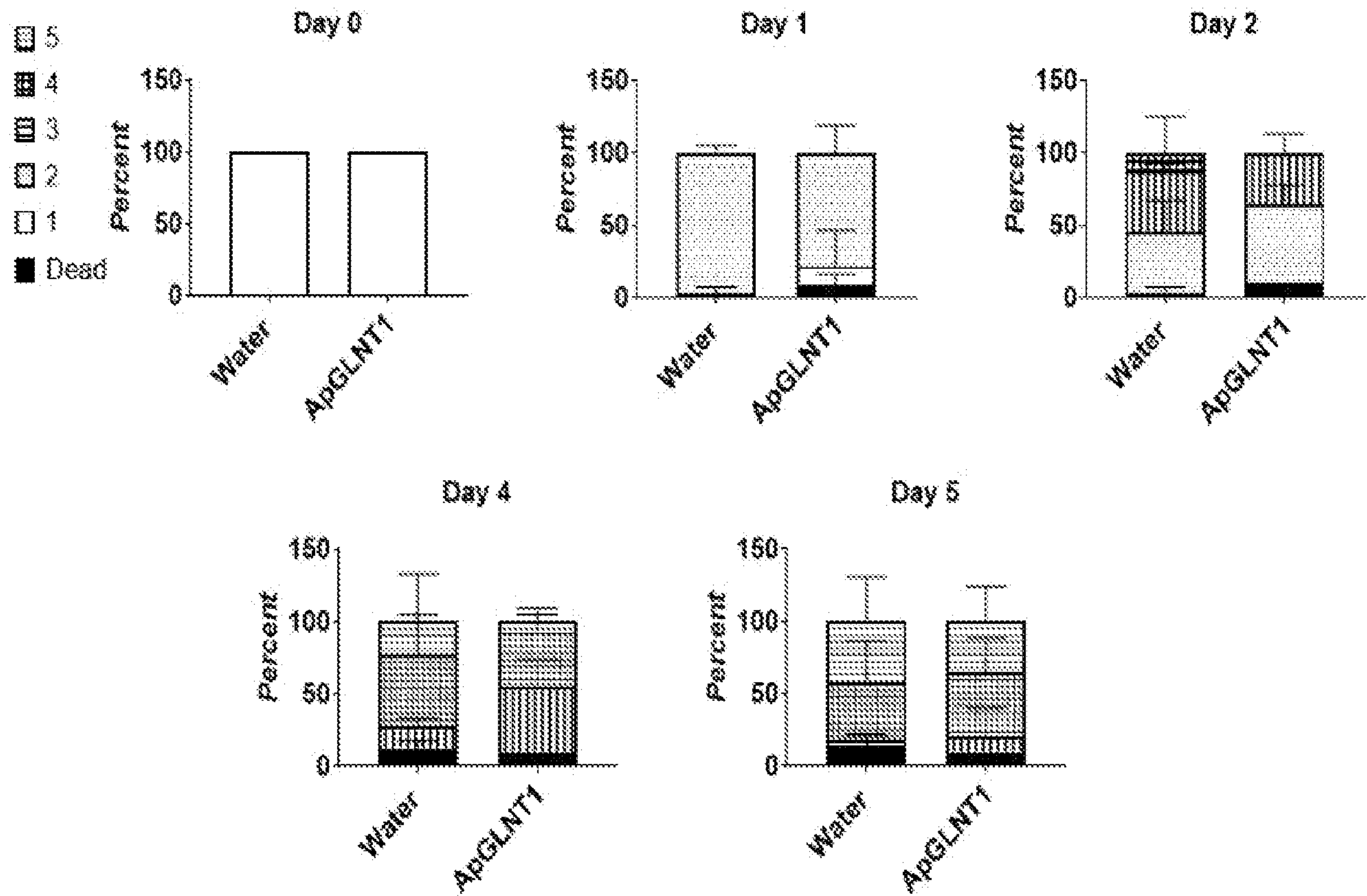


FIG. 16A

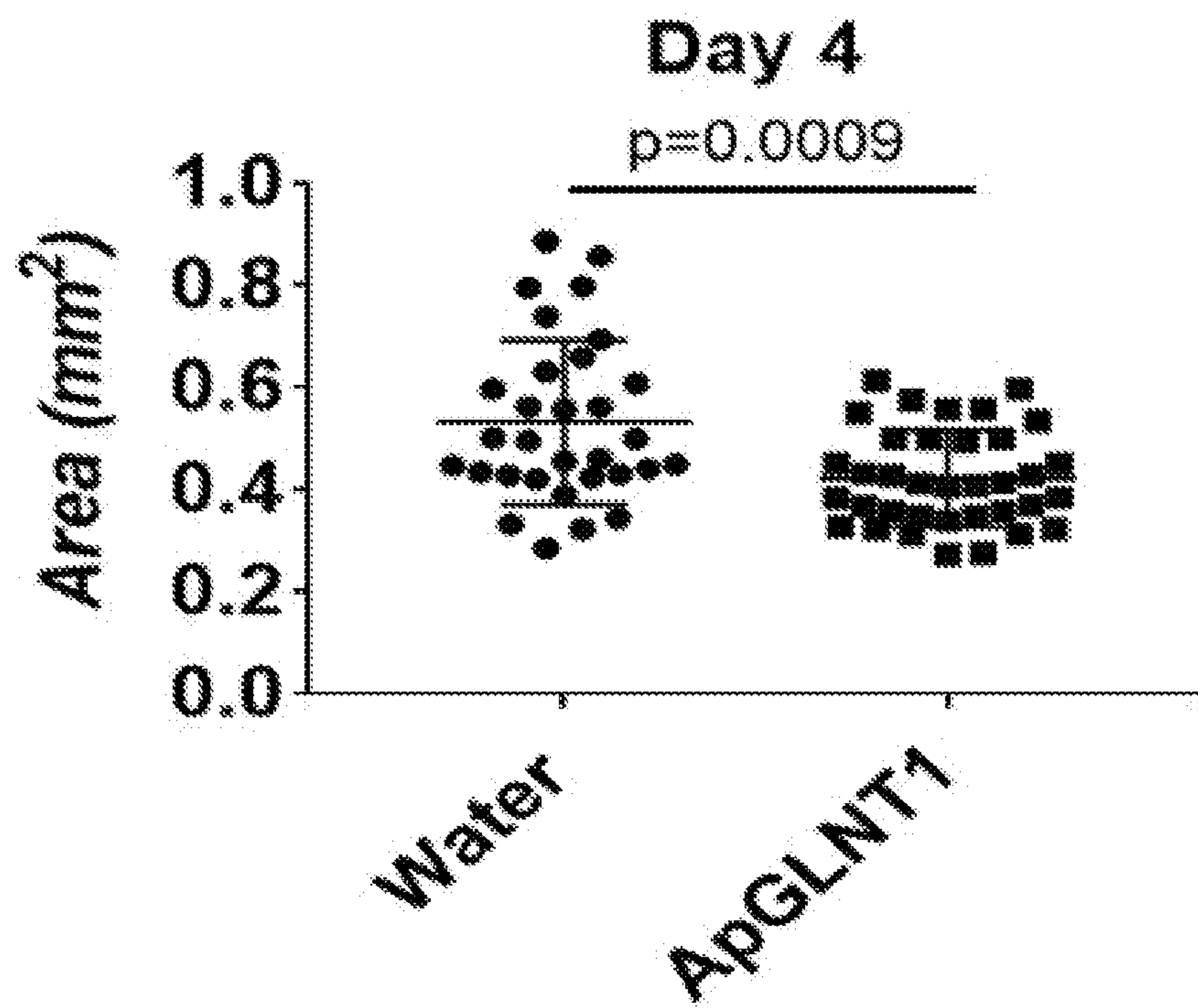


FIG. 16B

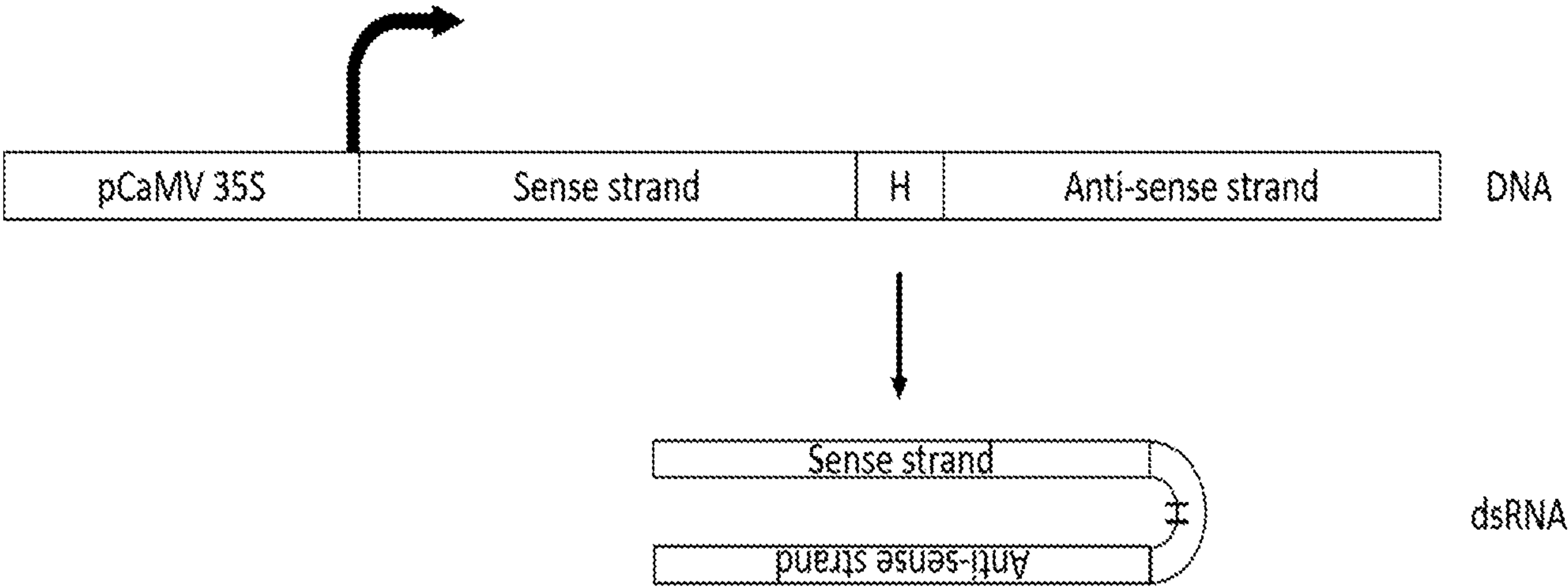


FIG. 17



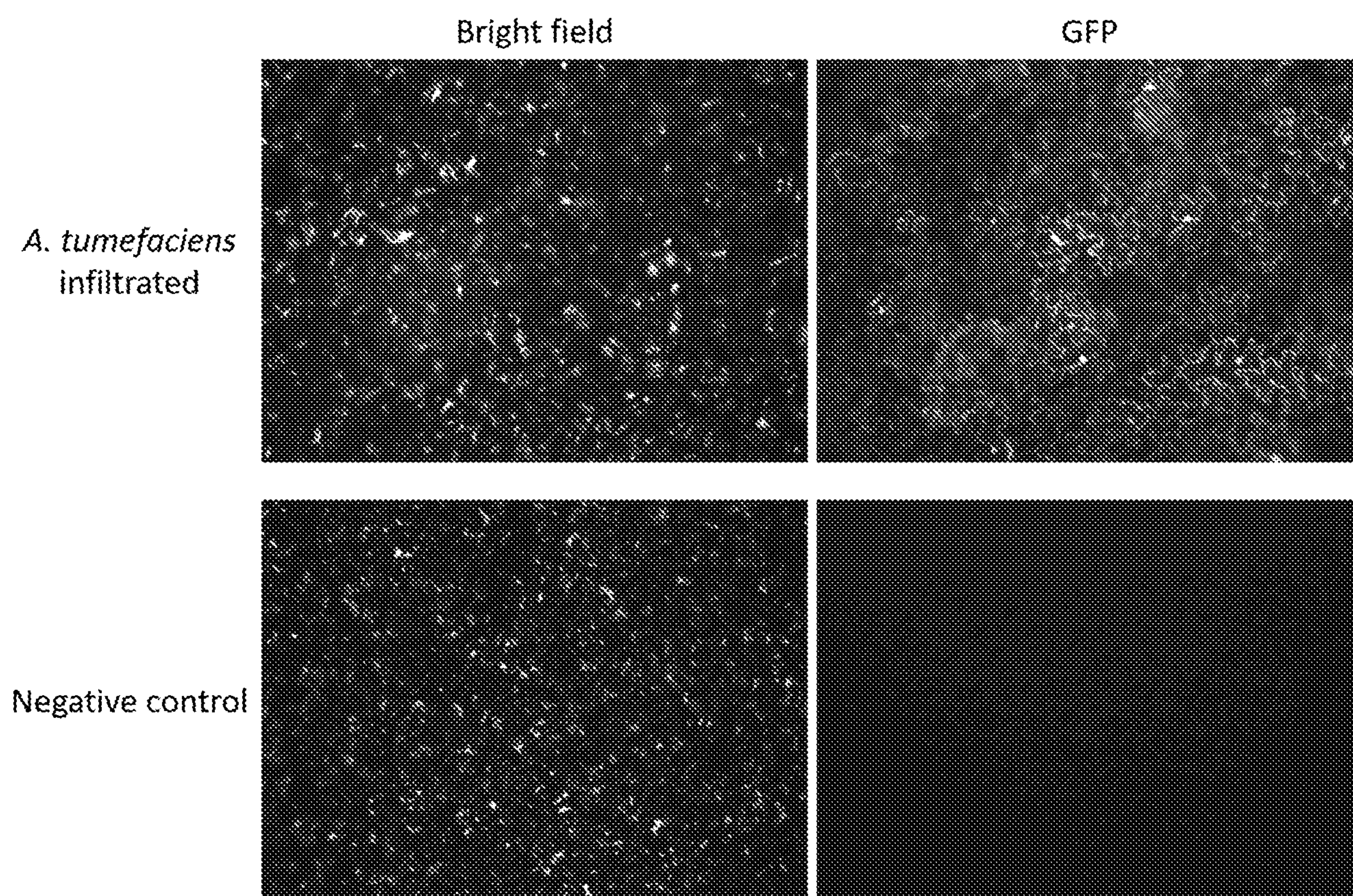


FIG. 19