ALPHA (1,2) FUCOSYLTRANSFERASE SYNGENES FOR USE IN THE PRODUCTION OF FUCOSYLATED OLIGOSACCHARIDES

FIELD OF THE INVENTION

The invention provides compositions and methods for producing purified oligosaccharides, in particular certain fucosylated oligosaccharides that are typically found in human milk.

BACKGROUND OF THE INVENTION

Human milk contains a diverse and abundant set of neutral and acidic oligosaccharides. More than 130 different complex oligosaccharides have been identified in human milk, and their structural diversity and abundance is unique to humans. Although these molecules may not be utilized directly by infants for nutrition, they nevertheless serve critical roles in the establishment of a healthy gut microbiome, in the prevention of disease, and in immune function. Prior to the invention described herein, the ability to produce human milk oligosaccharides (HMOS) inexpensively was problematic. For example, their production through chemical synthesis was limited by stereo-specificity issues, precursor availability, product impurities, and high overall cost. As such, there is a pressing need for new strategies to inexpensively manufacture large quantities of HMOS.

SUMMARY

The disclosure features an efficient and economical method for producing fucosylated oligosaccharides. Such production of a fucosylated oligosaccharide is accomplished using an isolated nucleic acid comprising a sequence encoding a lactose-utilizing α (1,2) fucosyltransferase gene product (e.g., polypeptide or protein), which is operably linked to one or more heterologous control sequences that direct the production of the recombinant fucosyltransferase gene product in a host production bacterium such as *Escherichia coli* (*E. coli*).

In an aspect, the present disclosure refers to a method for producing a fucosylated oligosaccharide in a bacterium comprising providing bacterium comprising an exogenous lactose-utilizing α(1,2) fucosyltransferase enzyme, wherein the amino acid sequence of said enzyme comprises

(i) the amino acid sequence of SEQ ID NO: 11;

(ii) the amino acid sequence of SEQ ID NO: 12;

(iii)the amino acid sequence of SEQ ID NO: 13;

(iv)the amino acid sequence of SEQ ID NO: 14;

(v) the amino acid sequence of SEQ ID NO: 17; or

(vi)the amino acid sequence of SEQ ID NO: 19.

Also disclosed herein are novel α (1,2) fucosyltransferases (also referred to herein as α(1,2) FTs) that utilize lactose and catalyzes the transfer of an L-fucose sugar from a GDP-fucose donor substrate to an acceptor substrate in an alpha-1,2-linkage. In a preferred example, the acceptor substrate is an oligosaccharide. The α(1,2) fucosyltransferases identified and described herein are useful for expressing in host bacterium for the production of human milk oligosaccharides (HMOS), such as fucosylated oligosaccharides. Exemplary

fucosylated oligosaccharides produced by the methods described herein include 2'-fucosyllactose (2'FL), lactodifucotetraose (LDFT), lacto-N-fucopentaose I (LNF I), or lacto-N-difucohexaose I (LDFH I). The "α(1,2) fucosyltransferases" disclosed herein encompasses the amino acid sequences of the α(1,2) fucosyltransferases and the nucleic acid sequences that encode the α(1,2) fucosyltransferases, as well as variants and fragments thereof that exhibit α(1,2) fucosyltransferase activity. Also within the disclosure is a nucleic acid construct comprising an isolated nucleic acid encoding a lactose-accepting α (1,2) fucosyltransferase enzyme, said nucleic acid being optionally operably linked to one or more heterologous control sequences that direct the production of the enzyme in a host bacteria production strain.

In another aspect, there is provided an isolated bacterium comprising an isolated nucleic acid encoding a lactose-accepting α (1,2) fucosyltransferase enzyme, wherein the amino acid sequence of said enzyme encoded by said nucleic acid comprises

(i) the amino acid sequence of SEQ ID NO: 11;

(ii) the amino acid sequence of SEQ ID NO: 12;

(iii)the amino acid sequence of SEQ ID NO: 13;

(iv)the amino acid sequence of SEQ ID NO: 14;

(v) the amino acid sequence of SEQ ID NO: 17; or

(vi)the amino acid sequence of SEQ ID NO: 19; and

wherein the isolated bacterium further comprising a α(1,3) fucosyltransferase enzyme and/or a α(1,4) fucosyltransferase enzyme.

The amino acid sequence of the lactose-accepting α(1,2) fucosyltransferases described herein is at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% identity to *Helicobacter pylori* 26695 alpha-(1,2) fucosyltransferase (futC or SEQ ID NO: 1). Preferably, the lactose-accepting α(1,2) fucosyltransferases described herein is at least 22% identical to *H. pylori* FutC, or SEQ ID NO: 1.

The amino acid sequence of the lactose-accepting α(1,2) fucosyltransferases described herein is at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% identity to *Bacteroides vulgatus* alpha-(1,2) fucosyltransferase (FutN or SEQ ID NO: 3). Preferably, the lactose-accepting α(1,2)

fucosyltransferases described herein is at least 25% identical to *B. vlugatos* FutN, or SEQ ID NO: 3.

Alternatively, the exogenous α (1,2) fucosyltransferase preferably comprises at least at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% identity to any one of the novel α (1,2) fucosyltransferases disclosed herein, for example, to the amino acid sequences in Table 1.

Exemplary α(1,2) fucosyltransferases include, but are not limited to, *Prevotella melaninogenica* FutO, *Clostridium bolteae* FutP, *Clostridium bolteae* +13 FutP, *Lachnospiraceae sp.* FutQ, *Methanosphaerula palustris* FutR, *Tannerella sp.* FutS, *Bacteroides caccae* FutU, *Butyrivibrio* FutV, *Prevotella sp.* FutW, *Parabacteroides johnsonii* FutX, *Akkermansia muciniphilia* FutY, *Salmonella enterica* FutZ, *Bacteroides sp.* FutZA.

For example, the α(1,2) fucosyltransferases comprise the amino acid sequences comprising any one of the following: *Prevotella melaninogenica* FutO (SEQ ID NO: 10), *Clostridium bolteae* FutP (SEQ ID NO: 11), *Clostridium bolteae* +13 FutP (SEQ ID NO: 292), *Lachnospiraceae sp.* FutQ (SEQ ID NO: 12), *Methanosphaerula palustris* FutR (SEQ ID NO: 13), *Tannerella sp.* FutS (SEQ ID NO: 14), *Bacteroides caccae* FutU (SEQ ID NO: 15), *Butyrivibrio* FutV (SEQ ID NO: 16), *Prevotella sp.* FutW (SEQ ID NO: 17), *Parabacteroides johnsonii* FutX (SEQ ID NO: 18), *Akkermansia muciniphilia* FutY (SEQ ID NO: 19), *Salmonella enterica* FutZ (SEQ ID NO: 20), and *Bacteroides sp.* FutZA (SEQ ID NO: 21), or a functional variant or fragment thereof. Other exemplary α(1,2) fucosyltransferases include any of the enzymes listed in Table 1, or functional variants or fragments thereof.

The present disclosure features a method for producing a fucosylated oligosaccharide in a bacterium by providing bacterium that express at least one exogenous lactose-utilizing α(1,2) fucosyltransferase. The amino acid sequence of the exogenous lactose-utilizing α(1,2) fucosyltransferase is preferably at least 22% identical to *H. pylori* FutC or at least 25% identical to *B. vulgatus* FutN. The bacterium also expresses one or more exogenous lactose-utilizing α(1,3) fucosyltransferase enzymes and/or one or more exogenous lactose-utilizing α(1,4) fucosyltransferase enzymes. The combination of fucosyltransferases expressed in the production bacterium is dependent upon the desired fucosylated oligosaccharide product. The method disclosed herein further includes retrieving the fucosylated oligosaccharide from said bacterium or from a culture supernatant of said bacterium.

Examples of suitable α(1,3) fucosyltransferase enzymes include, but are not limited to *Helicobacter pylori* 26695 *futA* gene (GenBank Accession Number HV532291 (GI:365791177), incorporated herein by reference), *H. hepaticus* Hh0072, *H.pylori* 11639 FucT, and *H.pylori* UA948 FucTa (e.g., GenBank Accession Number AF194963 (GI:28436396), incorporated herein by reference)(Rasko, D. A., Wang, G., Palcic, M. M. & Taylor, D. E. J Biol Chem 275, 4988-4994 (2000)). Examples of suitable a(1,4) fucosyltransferase enzymes include, but are not limited to *H.pylori* UA948 FucTa (which has has relaxed acceptor specificity and is able to generate both α(1,3)- and α(1,4)-fucosyl linkages). An example of an enzyme possessing only α(1,4) fucosyltransferase activity is given by the FucT III enzyme from *Helicobacter pylori* strain DMS6709 (e.g., GenBank Accession Number AY450598.1 (GI:40646733), incorporated herein by reference) (S. Rabbani, V. Miksa, B. Wipf, B. Ernst, *Glycobiology* **15**, 1076-83 (2005).)

The disclosure also features a nucleic acid construct or a vector comprising a nucleic acid enconding at least one α (1,2) fucosyltransferase or variant, or fragment thereof, as described herein. The vector can further include one or more regulatory elements, e.g., a heterologous promoter. By "heterologous" is meant that the control sequence and protein-encoding sequence originate from different bacterial strains. The regulatory elements can be operably linked to a gene encoding a protein, a gene construct encoding a fusion protein gene, or a series of genes linked in an operon in order to express the fusion protein. The disclosure comprises an isolated recombinant cell, e.g., a bacterial cell containing an aforementioned nucleic acid molecule or vector. The nucleic acid is optionally integrated into the genome of the host bacterium. In some embodiments, the nucleic acid construct also further comprises one or more α(1,3) fucosyltransferases and/or α(1,4) fucosyltransferases. Alternatively, the α (1,2) fucosyltransferase also exhibits α(1,3) fucosyltransferase and/or α(1,4) fucosyltransferase activity.

The bacterium utilized in the production methods described herein is genetically engineered to increase the efficiency and yield of fucosylated oligosaccharide products. For example, the host production bacterium is characterized as having a reduced level of β-galactosidase activity, a defective colanic acid synthesis pathway, an inactivated ATP-dependent intracellular protease, an inactivated *lacA*, or a combination thereof. In one embodiment, the bacterium is characterized as having a reduced level of β-galactosidase activity, a defective colanic acid synthesis pathway, an inactivated ATP-dependent intracellular protease, and an inactivated *lacA*.

As used herein, an "inactivated" or "inactivation of a" gene, encoded gene product (*i.e.*, polypeptide), or pathway refers to reducing or eliminating the expression (*i.e.*, transcription or translation), protein level (*i.e.*, translation, rate of degradation), or enzymatic activity of the gene, gene product, or pathway. In the instance where a pathway is inactivated, preferably one enzyme or polypeptide in the pathway exhibits reduced or negligible activity. For example, the enzyme in the pathway is altered, deleted or mutated such that the product of the pathway is produced at low levels compared to a wild-type bacterium or an intact pathway. Alternatively, the product of the pathway is not produced. Inactivation of a gene is achieved by deletion or mutation of the gene or regulatory elements of the gene such that the gene is no longer transcribed or translated. Inactivation of a polypeptide can be achieved by deletion or mutation of the gene that encodes the gene product or mutation of the polypeptide to disrupt its activity. Inactivating mutations include

additions, deletions or substitutions of one or more nucleotides or amino acids of a nucleic acid or amino acid sequence that results in the reduction or elimination of the expression or activity of the gene or polypeptide. In other embodiments, inactivation of a polypeptide is achieved through the addition of exogenous sequences (*i.e.*, tags) to the N or C-terminus of the polypeptide such that the activity of the polypeptide is reduced or eliminated (*i.e.*, by steric hindrance).

A host bacterium suitable for the production systems described herein exhibits an enhanced or increased cytoplasmic or intracellular pool of lactose and/or GDP-fucose. For example, the bacterium is *E. coli* and endogenous *E. coli* metabolic pathways and genes are manipulated in ways that result in the generation of increased cytoplasmic concentrations of lactose and/or GDP-fucose, as compared to levels found in wild type *E. coli*. Preferably, the bacterium accumulates an increased intracellular lactose pool and an increased intracellular GDP-fucose pool. For example, the bacteria contain at least 10%, 20%, 50%, or 2X, 5X, 10X or more of the levels of intracellular lactose and/or intracellular GDP-fucose compared to a corresponding wild type bacteria that lacks the genetic modifications described herein.

Increased intracellular concentration of lactose in the host bacterium compared to wild-type bacterium is achieved by manipulation of genes and pathways involved in lactose import, export and catabolism. In particular, described herein are methods of increasing intracellular lactose levels in *E. coli* genetically engineered to produce a human milk oligosaccharide by simultaneous deletion of the endogenous β-galactosidase gene (*lacZ*) and the lactose operon repressor gene (*lacI*). During construction of this deletion, the *lacIq* promoter is placed immediately upstream of (contiguous with) the lactose permease gene, *lacY*, i.e., the sequence of the *lacIq* promoter is directly upstream and adjacent to the start of the sequence encoding the *lacY* gene, such that the *lacY* gene is under transcriptional regulation by the lacIq promoter. The modified strain maintains its ability to transport lactose from the culture medium (via LacY), but is deleted for the wild-type chromosomal copy of the *lacZ* (encoding β-galactosidase) gene responsible for lactose catabolism. Thus, an intracellular lactose pool is created when the modified strain is cultured in the presence of exogenous lactose.

Another method for increasing the intracellular concentration of lactose in *E. coli* involves inactivation of the *lacA* gene. A inactivating mutation, null mutation, or deletion of *lacA* prevents the formation of intracellular acetyl-lactose, which not only removes this molecule as a contaminant from subsequent purifications, but also eliminates *E.coli*'s ability

to export excess lactose from its cytoplasm (Danchin A. Cells need safety valves. Bioessays 2009, Jul;31(7):769-73.), thus greatly facilitating purposeful manipulations of the *E.coli* intracellular lactose pool.

The disclosure also provides methods for increasing intracellular levels of GDP-fucose in a bacterium by manipulating the organism's endogenous colanic acid biosynthesis pathway. This increase is achieved through a number of genetic modifications of endogenous *E. coli* genes involved either directly in colanic acid precursor biosynthesis, or in overall control of the colanic acid synthetic regulon. Particularly preferred is inactivation of the genes or encoded polypeptides that act in the colanic acid synthesis pathway after the production of GDP-fucose (the donor substrate) and before the generation of colanic acid. Exemplary colanic acid synthesis genes include, but are not limited to: a *wcaJ* gene, (e.g., GenBank Accession Number (amino acid) BAA15900 (GI:1736749), incorporated herein by reference), a *wcaA* gene (e.g., GenBank Accession Number (amino acid) BAA15912.1 (GI:1736762), incorporated herein by reference), a *wcaC* gene (e.g., GenBank Accession Number (amino acid) BAE76574.1 (GI:85675203), incorporated herein by reference), a *wcaE* gene (e.g., GenBank Accession Number (amino acid) BAE76572.1 (GI:85675201), incorporated herein by reference), a *wcaI* gene (e.g., GenBank Accession Number (amino acid) BAA15906.1 (GI:1736756), incorporated herein by reference), a *wcaL* gene (e.g., GenBank Accession Number (amino acid) BAA15898.1 (GI:1736747), incorporated herein by reference), a *wcaB* gene (e.g., GenBank Accession Number (amino acid) BAA15911.1 (GI:1736761), incorporated herein by reference), a *wcaF* gene (e.g., GenBank Accession Number (amino acid) BAA15910.1 (GI:1736760), incorporated herein by reference), a *wzxE* gene (e.g., GenBank Accession Number (amino acid) BAE77506.1 (GI:85676256), incorporated herein by reference), a *wzxC* gene, (e.g., GenBank Accession Number (amino acid) BAA15899 (GI:1736748), incorporated herein by reference), a *wcaD* gene, (e.g., GenBank Accession Number (amino acid) BAE76573 (GI:85675202), incorporated herein by reference), a *wza* gene (e.g., GenBank Accession Number (amino acid) BAE76576 (GI:85675205), incorporated herein by reference), a *wzb* gene (e.g., GenBank Accession Number (amino acid) BAE76575 (GI:85675204), incorporated herein by reference), and a *wzc* gene (e.g., GenBank Accession Number (amino acid) BAA15913 (GI:1736763), incorporated herein by reference).

Preferably, a host bacterium, such as *E. coli*, is genetically engineered to produce a human milk oligosaccharide by the inactivation of the *wcaJ* gene, which encoding the UDP-glucose lipid carrier transferase. The inactivation of the *wcaJ* gene can be by deletion of the

gene, a null mutation, or inactivating mutation of the *wcaJ* gene, such that the activity of the encoded wcaJ is reduced or eliminated compared to wild-type *E. coli*. In a *wcaJ* null background, GDP-fucose accumulates in the *E. coli* cytoplasm.

Over-expression of a positive regulator protein, RcsA (e.g., GenBank Accession Number M58003 (GI:1103316), incorporated herein by reference), in the colanic acid synthesis pathway results in an increase in intracellular GDP-fucose levels. Over-expression of an additional positive regulator of colanic acid biosynthesis, namely RcsB (e.g., GenBank Accession Number E04821 (GI:2173017), incorporated herein by reference), is also utilized, either instead of or in addition to over-expression of RcsA, to increase intracellular GDP-fucose levels.

Alternatively, colanic acid biosynthesis is increased following the introduction of a mutation into the *E. coli lon* gene (e.g., GenBank Accession Number L20572 (GI:304907), incorporated herein by reference). Lon is an adenosine-5'-triphosphate (ATP)-dependant intracellular protease that is responsible for degrading RcsA, mentioned above as a positive transcriptional regulator of colanic acid biosynthesis in *E. coli*. In a *lon* null background, RcsA is stabilized, RcsA levels increase, the genes responsible for GDP-fucose synthesis in *E. coli* are up-regulated, and intracellular GDP-fucose concentrations are enhanced. Mutations in *lon* suitable for use with the methods presented herein include null mutations or insertions that disrupt the expression or function of *lon*.

A functional lactose permease gene is also present in the bacterium. The lactose permease gene is an endogenous lactose permease gene or an exogenous lactose permease gene. For example, the lactose permease gene comprises an *E. coli lacY* gene (e.g., GenBank Accession Number V00295 (GI:41897), incorporated herein by reference). Many bacteria possess the inherent ability to transport lactose from the growth medium into the cell, by utilizing a transport protein that is either a homolog of the *E. coli* lactose permease (e.g., as found in *Bacillus licheniformis*), or a transporter that is a member of the ubiquitous PTS sugar transport family (e.g., as found in *Lactobacillus casei* and *Lactobacillus rhamnosus*). For bacteria lacking an inherent ability to transport extracellular lactose into the cell cytoplasm, this ability is conferred by an exogenous lactose transporter gene (e.g., *E. coli lacY*) provided on recombinant DNA constructs, and supplied either on a plasmid expression vector or as exogenous genes integrated into the host chromosome.

As described herein, in some embodiments, the host bacterium preferably has a reduced level of β-galactosidase activity. In the embodiment in which the bacterium is

characterized by the deletion of the endogenous β-galactosidase gene, an exogenous β-galactosidase gene is introduced to the bacterium. For example, a plasmid expressing an exogenous β-galactosidase gene is introduced to the bacterium, or recombined or integrated into the host genome. For example, the exogenous β-galactosidase gene is inserted into a gene that is inactivated in the host bacterium, such as the *lon* gene.

The exogenous b-galactosidase gene is a functional b-galactosidase gene characterized by a reduced or low leve of b-galactosidase activity compared to β-galactosidase activity in wild-type bacteria lacking any genetic manipulation. Exemplary β-galactosidase genes include *E. coli lacZ* and β-galactosidase genes from any of a number of other organisms (e.g., the *lac4* gene of *Kluyveromyces lactis* (e.g., GenBank Accession Number M84410 (GI:173304), incorporated herein by reference) that catalyzes the hydrolysis of b-galactosides into monosaccharides. The level of β-galactosidase activity in wild-type *E. coli* bacteria is, for example, 6,000 units. Thus, the reduced β-galactosidase activity level encompassed by engineered host bacterium of the present disclosure includes less than 6,000 units, less than 5,000 units, less than 4,000 units, less than 3,000 units, less than 2,000 units, less than 1,000 units, less than 900 units, less than 800 units, less than 700 units, less than 600 units, less than 500 units, less than 400 units, less than 300 units, less than 200 units, less than 100 units, or less than 50 units. Low, functional levels of β-galactosidase include β-galactosidase activity levels of between 0.05 and 1,000 units, e.g., between 0.05 and 750 units, between 0.05 and 500 units, between 0.05 and 400 units, between 0.05 and 300 units, between 0.05 and 200 units, between 0.05 and 100 units, between 0.05 and 50 units, between 0.05 and 10 units, between 0.05 and 5 units, between 0.05 and 4 units, between 0.05 and 3 units, or between 0.05 and 2 units of β-galactosidase activity. For unit definition and assays for determining β-galactosidase activity, see Miller JH, Laboratory CSH. Experiments in molecular genetics. Cold Spring Harbor Laboratory Cold Spring Harbor, NY; 1972; (incorporated herein by reference). This low level of cytoplasmic β-galactosidase activity is not high enough to significantly diminish the intracellular lactose pool. The low level of β-galactosidase activity is very useful for the facile removal of undesired residual lactose at the end of fermentations.

Optionally, the bacterium has an inactivated *thyA* gene. Preferably, a mutation in a *thyA* gene in the host bacterium allows for the maintenance of plasmids that carry *thyA* as a selectable marker gene. Exemplary alternative selectable markers include antibiotic resistance genes such as BLA (beta-lactamase), or *proBA* genes (to complement a *proAB*

host strain proline auxotropy) or *purA* (to complement a *purA* host strain adenine auxotrophy).

The *E. coli* bacterium comprises the genotype Δ*ampC*::P_{trp}^{B}*cI*, Δ(*lacI-lacZ*)::*FRT*, P_{lacIq}*lacY*$^{+}$, Δ*wcaJ*::FRT, *thyA*::Tn*10*, Δ*lon*:(*npt3*, *lacZ*$^{+}$), Δ*lacA*, and also comprises any one of the exogenous α(1,2) fucosyltransferases described herein.

The bacterium comprising these characteristics is cultured in the presence of lactose. In some cases, the method further comprises culturing the bacterium in the presence of tryptophan and in the absence of thymidine. The fucosylated oligosaccharide is retrieved from the bacterium (*i.e.*, a cell lysate) or from a culture supernatant of the bacterium.

The disclosure provides a purified fucosylated oligosaccharide produced by the methods described herein. The fucosylated oligosaccharide is purified for use in therapeutic or nutritional products, or the bacterium is used directly in such products. The fucosylated oligosaccharide produced by the engineered bacterium is 2'-fucosyllactose (2'-FL) or lactodifucotetraose (LDFT). The new alpha 1,2-fucosyltransferases are also useful to synthesize HMOS of larger molecular weight bearing alpha 1,2 fucose moieties, e.g., lacto-N-fucopentaose (LNF I) and lacto-N-difucohexaose (LDFH I). For example, to produce LDFT, the host bacterium is engineered to express an exogenous α (1,2) fucosyltransferase that also possesses α (1,3) fucosyltransferase activity, or an exogenous α (1,2) fucosyltransferase and an exogenous α (1,3) fucosyltransferase. For the production of LNF I and LDFH I, the host bacterium is engineered to express an exogenous α (1,2) fucosyltransferase that also possesses α (1,3) fucosyltransferase activity and/or α (1,4) fucosyltransferase activity, or an exogenous α (1,2) fucosyltransferase, an exogenous α (1,3) fucosyltransferas, and an exogenous α (1,4) fucosyltransferase.

A purified fucosylated oligosaccharide produced by the methods described above is also within the disclosure. The purified oligosaccharide (2'-FL) obtained at the end of the process is a white/slightly off-white, crystalline, sweet powder. For example, an engineered bacterium, bacterial culture supernatant, or bacterial cell lysate according to the disclosure comprises 2'-FL, LDFT, LNF I or LDFH I produced by the methods described herein, and does not substantially comprise a other fucosylated oligosaccharides prior to purification of the fucosylated oligosaccharide products from the cell, culture supernatant, or lysate. As a general matter, the fucosylated oligosaccharide produced by the methods contains a negligible amount of 3-FL in a 2'-FL-containing cell, cell lysate or culture, or supernatant, e.g., less than 1% of the level of 2'-FL or 0.5% of the level of 2'-FL. Moreover, the

fucosylated oligosaccharide produced by the methods described herein also have a minimal amount of contaminating lactose, which can often be co-purified with the fucosylated oligosaccharide product, such as 2'FL. This reduction in contaminating lactose results from the reduced level of β-galactosidase activity present in the engineered host bacterium.

A purified oligosaccharide, e.g., 2'-FL, LDFT, LNF I, or LDFH I, is one that is at least 90%, 95%, 98%, 99%, or 100% (w/w) of the desired oligosaccharide by weight. Purity is assessed by any known method, e.g., thin layer chromatography or other chromatographic techniques known in the art. The disclosure includes a method of purifying a fucosylated oligosaccharide produced by the genetically engineered bacterium described above, which method comprises separating the desired fucosylated oligosaccharide (e.g., 2'-FL) from contaminants in a bacterial cell lysate or bacterial cell culture supernatant of the bacterium.

The oligosaccharides are purified and used in a number of products for consumption by humans as well as animals, such as companion animals (dogs, cats) as well as livestock (bovine, equine, ovine, caprine, or porcine animals, as well as poultry). For example, a pharmaceutical composition comprises purified 2'-FL and a pharmaceutically-acceptable excipient that is suitable for oral administration. Large quantities of 2'-FL are produced in bacterial hosts, e.g., an *E. coli* bacterium comprising an exogenous α (1,2) fucosyltransferase gene.

A method of producing a pharmaceutical composition comprising a purified human milk oligosaccharide (HMOS) is carried out by culturing the bacterium described above, purifying the HMOS produced by the bacterium, and combining the HMOS with an excipient or carrier to yield a dietary supplement for oral administration. These compositions are useful in methods of preventing or treating enteric and/or respiratory diseases in infants and adults. Accordingly, the compositions are administered to a subject suffering from or at risk of developing such a disease.

The disclosure also provides methods of identifying an α (1,2) fucosyltransferase gene capable of synthesizing fucosylated oligosaccharides in a host bacterium, *i.e.*, 2'-fucosyllactose (2'-FL) in *E. coli*. The method of identifying novel lactose-utilizing, α(1,2)fucosyltransferase enzyme comprises the following steps:

1) performing a computational search of sequence databases to define a broad group of simple sequence homologs of any known, lactose-utilizing α(1,2)fucosyltransferase;

2) using the list from step (1), deriving a search profile containing common sequence and/or structural motifs shared by the members of the list;

3) searching sequence databases, using a derived search profile based on the common sequence or structural motif from step (2) as query, and identifying a candidate sequences, wherein a sequence homology to a reference lactose-utilizing α(1,2)fucosyltransferase is a predetermined percentage threshold;

4) compiling a list of candidate organisms, said organisms being characterized as expressing α(1,2)fucosyl- glycans in a naturally-occurring state;

5) selecting candidate sequences that are derived from candidate organisms to generate a list of candidate lactose-utilizing enzymes;

6) expressing the candidate lactose- utilizing enzyme in a host organism; and

7) testing for lactose- utilizing α(1,2)fucosyltransferase activity, wherein detection of the desired fucosylated oligosaccharide product in said organism indicates that the candidate sequence comprises a novel lactose- utilizing α(1,2)fucosyltransferase. In another embodiment, the search profile is generated from a multiple sequence alignment of the amino acid sequences of more than one enzyme with known α(1,2)fucosyltransferase activity. The database search can then be designed to refine and iteratively search for novel α(1,2)fucosyltransferases with significant sequence similarlity to the multiple sequence alignment query.

The disclosure provides a method of treating, preventing, or reducing the risk of infection in a subject comprising administering to said subject a composition comprising a purified recombinant human milk oligosaccharide, wherein the HMOS binds to a pathogen and wherein the subject is infected with or at risk of infection with the pathogen. The infection is caused by a Norwalk-like virus or *Campylobacter jejuni.* The subject is preferably a mammal in need of such treatment. The mammal is, e.g., any mammal, e.g., a human, a primate, a mouse, a rat, a dog, a cat, a cow, a horse, or a pig. In a preferred embodiment, the mammal is a human. For example, the compositions are formulated into animal feed (e.g., pellets, kibble, mash) or animal food supplements for companion animals, e.g., dogs or cats, as well as livestock or animals grown for food consumption, e.g., cattle, sheep, pigs, chickens, and goats. Preferably, the purified HMOS is formulated into a powder (e.g., infant formula powder or adult nutritional supplement powder, each of which is mixed with a liquid such as water or juice prior to consumption) or in the form of tablets, capsules or pastes or is incorporated as a component in dairy products such as milk, cream, cheese, yogurt or kefir, or as a component in any beverage, or combined in a preparation containing live microbial cultures intended to serve as probiotics, or in prebiotic preparations to enhance the growth of beneficial microorganisms either *in vitro* or *in vivo*.

Polynucleotides, polypeptides, and oligosaccharides of the disclosure are purified and/or isolated. Purified defines a degree of sterility that is safe for administration to a human subject, e.g., lacking infectious or toxic agents. Specifically, as used herein, an "isolated" or "purified" nucleic acid molecule, polynucleotide, polypeptide, protein or oligosaccharide, is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized. For example, purified HMOS compositions are at least 60% by weight (dry weight) the compound of interest. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight the compound of interest. Purity is measured by any appropriate standard method, for example, by column chromatography, thin layer chromatography, or high-performance liquid chromatography (HPLC) analysis. For example, a "purified protein" refers to a protein that has been separated from other proteins, lipids, and nucleic acids with which it is naturally associated. Preferably, the protein constitutes at least 10, 20, 50, 70, 80, 90, 95, 99-100% by dry weight of the purified preparation.

Similarly, by "substantially pure" is meant an oligosaccharide that has been separated from the components that naturally accompany it. Typically, the oligosaccharide is substantially pure when it is at least 60%, 70%, 80%, 90%, 95%, or even 99%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated.

By "isolated nucleic acid" is meant a nucleic acid that is free of the genes which, in the naturally-occurring genome of the organism from which the DNA of the disclosure is derived, flank the gene. The term covers, for example: (a) a DNA which is part of a naturally occurring genomic DNA molecule, but is not flanked by both of the nucleic acid sequences that flank that part of the molecule in the genome of the organism in which it naturally occurs; (b) a nucleic acid incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner, such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA; (c) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment; and (d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion protein. Isolated nucleic acid molecules according to the present disclosure further include molecules produced synthetically, as well as any nucleic acids that have been altered chemically and/or that have modified backbones.

A "heterologous promoter" is a promoter which is different from the promoter to which a gene or nucleic acid sequence is operably linked in nature.

The term "overexpress" or "overexpression" refers to a situation in which more factor is expressed by a genetically-altered cell than would be, under the same conditions, by a wild type cell. Similarly, if an unaltered cell does not express a factor that it is genetically altered to produce, the term "express" (as distinguished from "overexpress") is used indicating the wild type cell did not express the factor at all prior to genetic manipulation.

The terms "treating" and "treatment" as used herein refer to the administration of an agent or formulation to a clinically symptomatic individual afflicted with an adverse condition, disorder, or disease, so as to effect a reduction in severity and/or frequency of symptoms, eliminate the symptoms and/or their underlying cause, and/or facilitate improvement or remediation of damage. The terms "preventing" and "prevention" refer to the administration of an agent or composition to a clinically asymptomatic individual who is susceptible to a particular adverse condition, disorder, or disease, and thus relates to the prevention of the occurrence of symptoms and/or their underlying cause.

By the terms "effective amount" and "therapeutically effective amount" of a formulation or formulation component is meant a nontoxic but sufficient amount of the formulation or component to provide the desired effect.

The transitional term "comprising," which is synonymous with "including," "containing," or "characterized by," is inclusive or open-ended and does not exclude additional, unrecited elements or method steps. By contrast, the transitional phrase "consisting of" excludes any element, step, or ingredient not specified in the claim. The transitional phrase "consisting essentially of" limits the scope of a claim to the specified materials or steps "and those that do not materially affect the basic and novel characteristic(s)" of the claimed disclosure.

The host organism used to express the lactose-accepting fucosyltransferase gene is typically the enterobacterium *Escherichia coli* K12 (*E. coli*). *E. coli* K-12 is not considered a human or animal pathogen nor is it toxicogenic. E. coli K-12 is a standard production strain of bacteria and is noted for its safety due to its poor ability to colonize the colon and establish infections (see, e.g., epa.gov/oppt/biotech/pubs/fra/fra004.htm). However, a variety of bacterial species may be used in the oligosaccharide biosynthesis methods, e.g., *Erwinia herbicola* (*Pantoea agglomerans*), *Citrobacter freundii*, *Pantoea citrea*, *Pectobacterium carotovorum*, *or Xanthomonas campestris.* Bacteria of the genus *Bacillus* may also be used, including *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus coagulans*, *Bacillus thermophilus*,

Bacillus laterosporus, *Bacillus megaterium*, *Bacillus mycoides*, *Bacillus pumilus*, *Bacillus lentus*, *Bacillus cereus*, *and Bacillus circulans*. Similarly, bacteria of the genera *Lactobacillus* and *Lactococcus* may be modified using the methods of this disclosure, including but not limited to *Lactobacillus acidophilus, Lactobacillus salivarius*, *Lactobacillus plantarum*, *Lactobacillus helveticus*, *Lactobacillus delbrueckii*, *Lactobacillus rhamnosus*, *Lactobacillus bulgaricus*, *Lactobacillus crispatus*, *Lactobacillus gasseri*, *Lactobacillus casei*, *Lactobacillus reuteri*, *Lactobacillus jensenii*, *and Lactococcus lactis. Streptococcus thermophiles and Proprionibacterium freudenreichii* are also suitable bacterial species for the disclosure described herein. Also included as part of this disclosure are strains, modified as described here, from the genera *Enterococcus* (e.g., *Enterococcus faecium* and *Enterococcus thermophiles*), *Bifidobacterium* (e.g., *Bifidobacterium longum*, *Bifidobacterium infantis*, and *Bifidobacterium bifidum*), *Sporolactobacillus* spp., *Micromomospora* spp., *Micrococcus* spp., *Rhodococcus* spp., and *Pseudomonas* (e.g., *Pseudomonas fluorescens* and *Pseudomonas aeruginosa*). Bacteria comprising the characteristics described herein are cultured in the presence of lactose, and a fucosylated oligosaccharide is retrieved, either from the bacterium itself or from a culture supernatant of the bacterium. The fucosylated oligosaccharide is purified for use in therapeutic or nutritional products, or the bacteria are used directly in such products. A suitable production host bacterial strain is one that is not the same bacterial strain as the source bacterial strain from which the fucosyltransferase-encoding nucleic acid sequence was identified.

Other features and advantages of the disclosure will be apparent from the following description of the preferred embodiments thereof, and from the claims. Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, suitable methods and materials are described below. All published foreign patents and patent applications cited herein are incorporated herein by reference. Genbank and NCBI submissions indicated by accession number cited herein are incorporated herein by reference. All other published references, documents, manuscripts and scientific literature cited herein are incorporated herein by reference. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic illustration showing the synthetic pathway of the major neutral fucosyl-oligosaccharides found in human milk.

FIG. 2 is a schematic demonstrating metabolic pathways and the changes introduced into them to engineer 2'-fucosyllactose (2'-FL) synthesis in *Escherichia coli* (*E. coli*). Specifically, the lactose synthesis pathway and the GDP-fucose synthesis pathway are illustrated. In the GDP-fucose synthesis pathway: *manA* = phosphomannose isomerase (PMI), *manB* = phosphomannomutase (PMM), *manC* = mannose-1-phosphate guanylyltransferase (GMP), *gmd* = GDP-mannose-4,6-dehydratase, *fcl* = GDP-fucose synthase (GFS), and Δ*wcaJ* = mutated UDP-glucose lipid carrier transferase.

FIG. 3A and FIG. 3B show the sequence identity and a multiple sequence alignment of 4 previously known lactose-utilizing α(1,2)-fucosyltransferase protein sequences. FIG. 3A is a table showing the sequence identity between the 4 known lactose-utilizing α(1,2)-fucosyltransferases: *H. pylori* futC (SEQ ID NO: 1), *H. mustelae* FutL (SEQ ID NO: 2), *Bacteroides vulgatus* futN (SEQ ID NO: 3), and *E. coli 0126* wbgL (SEQ ID NO: 4). FIG. 3B shows multiple sequence alignment of the 4 known α(1,2)-fucosyltransferases. The ovals highlight regions of particularly high sequence conservation between the four enzymes in the alignment.

FIG. 4 shows the sequence alignment of the 12 identified α(1,2)-fucosyltransferase syngenes identified, along with the 4 previously known lactose-utilizing α(1,2)-fucosyltransferase protein sequences. The 4 known lactose-utilizing α(1,2)-fucosyltransferases are boxed and include *H. pylori* futC (SEQ ID NO: 1), *H. mustelae* FutL (SEQ ID NO: 2), *Bacteroides vulgatus* futN (SEQ ID NO: 3), and *E. coli 0126* wbgL (SEQ ID NO: 4). The 12 identified α(1,2)-fucosyltransferase are as follows: *Prevotella melaninogenica* FutO (SEQ ID NO: 10), *Clostridium bolteae* +13 FutP (SEQ ID NO: 292), *Lachnospiraceae sp.* FutQ (SEQ ID NO: 12), *Methanosphaerula palustris* FutR (SEQ ID NO: 13), *Tannerella sp.* FutS (SEQ ID NO: 14), *Bacteroides caccae* FutU (SEQ ID NO: 15), *Butyrivibrio* FutV (SEQ ID NO: 16), *Prevotella sp.* FutW (SEQ ID NO: 17), *Parabacteroides johnsonii* FutX (SEQ ID NO: 18), *Akkermansia muciniphilia* FutY (SEQ ID NO: 19), *Salmonella enterica* FutZ (SEQ ID NO: 20), *Bacteroides sp.* FutZA (SEQ ID NO: 21). The sequence for *Clostridium bolteae* FutP (without the 13 additional amino acids in the N-terminus) (SEQ ID NO: 11) is also shown in the alignment.

FIG. 5A and FIG. 5B are two pictures of gels showing the construction of the syngenes for each of the 12 novel α(1,2)-fucosyltransferases. FIG. 5A shows post-Gibson assembly PCR. FIG. 5B shows gel-purified RI/Xho1 syngene fragments.

FIG. 6A and FIG. 6B are two photographs showing thin layer chromatograms of fucosylated oligosaccharide products produced in *E. coli* cultures using the 12 novel α(1,2)-fucosyltransferase syngenes. FIG. 6A shows fucosylated oligosaccharide products from 2μl of culture supernatant. FIG. 6B shows fucosylated oligosaccharide products from 0.2 OD_{600} cell equivalents of whole cell heat extracts.

FIG. 7 is a graph showing the growth curve of the host bacterium expressing plasmids containing the α(1,2) fucosyltransferase genes WbgL, FutN, FutO, FutQ, and FutX after tryptophan induction in the presence of lactose in the culture medium (i.e. lac + trp).

FIG. 8 is a photograph of a SDS-PAGE gel showing the proteins produced from host bacterium expressing α(1,2) fucosyltransferase genes WbgL, FutN, FutO, FutQ, and FutX after induction.

FIG. 9A and FIG. 9B are two photographs of thin layer chromatograms showing the production of fucosylated oligosaccharide products from in *E. coli* cultures expressing select α(1,2)-fucosyltransferase syngenes *WbgL*, *FutN*, *FutO*, *FutQ*, and *FutX* at 7 hours or 24 hours after induction. FIG. 9A shows fucosylated oligosaccharide products from 2μl of culture supernatant. FIG. 9B shows fucosylated oligosaccharide products from 0.2 OD_{600} cell equivalents of whole cell heat extracts.

FIG. 10A and FIG. 10B are two photographs of thin layer chromatograms showing the fucosylated oligosaccharide products after two different 1.5L fermentation runs from *E. coli* expressing FutN: FIG. 10A) 36B and FIG. 10B) 37A. The culture yield for run 36B was 33g/L while the yield for run 37A was 36.3 g/L.

FIG. 11 is a plasmid map of pG217 carrying the *B. vulgatus FutN* gene.

FIG. 12 is a schematic diagram showing the insertion of the *LacIq* promoter, the functional *lacY* gene, and the deletion of *lacA*.

FIG. 13 is a schematic diagram showing the deletion of the endogenous *wcaJ* gene using FRT recombination.

FIG. 14 is a schematic diagram of the *E. coli W3110* chromosome, showing the insertion of a DNA fragment carrying kanamycin resistance gene (derived from transposon Tn5) and wild-type *lacZ* into the *lon* gene.

DETAILED DESCRIPTION OF THE INVENTION

While some studies suggest that human milk glycans could be used as antimicrobial anti-adhesion agents, the difficulty and expense of producing adequate quantities of these agents of a quality suitable for human consumption has limited their full-scale testing and perceived utility. What has been needed is a suitable method for producing the appropriate glycans in sufficient quantities at reasonable cost. Prior to the invention described herein, there were attempts to use several distinct synthetic approaches for glycan synthesis. Some chemical approaches can synthesize oligosaccharides (Flowers, H. M. Methods Enzymol 50, 93-121 (1978); Seeberger, P. H. Chem Commun (Camb) 1115-1121 (2003)), but reactants for these methods are expensive and potentially toxic (Koeller, K. M. & Wong, C. H. Chem Rev 100, 4465-4494 (2000)). Enzymes expressed from engineered organisms (Albermann, C., Piepersberg, W. & Wehmeier, U. F. Carbohydr Res 334, 97-103 (2001); Bettler, E., Samain, E., Chazalet, V., Bosso, C., et al. Glycoconj J 16, 205-212 (1999); Johnson, K. F. Glycoconj J 16, 141-146 (1999); Palcic, M. M. Curr Opin Biotechnol 10, 616-624 (1999); Wymer, N. & Toone, E. J. Curr Opin Chem Biol 4, 110-119 (2000)) provide a precise and efficient synthesis (Palcic, M. M. Curr Opin Biotechnol 10, 616-624 (1999)); Crout, D. H. & Vic, G. Curr Opin Chem Biol 2, 98-111 (1998)), but the high cost of the reactants, especially the sugar nucleotides, limits their utility for low-cost, large-scale production. Microbes have been genetically engineered to express the glycosyltransferases needed to synthesize oligosaccharides from the bacteria's innate pool of nucleotide sugars (Endo, T., Koizumi, S., Tabata, K., Kakita, S. & Ozaki, A. Carbohydr Res 330, 439-443 (2001); Endo, T., Koizumi, S., Tabata, K. & Ozaki, A. Appl Microbiol Biotechnol 53, 257-261 (2000); Endo, T. & Koizumi, S. Curr Opin Struct Biol 10, 536-541 (2000); Endo, T., Koizumi, S., Tabata, K., Kakita, S. & Ozaki, A. Carbohydr Res 316, 179-183 (1999); Koizumi, S., Endo, T., Tabata, K. & Ozaki, A. Nat Biotechnol 16, 847-850 (1998)). However, prior to the invention described herein, there was a growing need to identify and characterize additional glycosyltransferases that are useful for the synthesis of HMOS in metabolically engineered bacterial hosts.

Human Milk Glycans

Human milk contains a diverse and abundant set of neutral and acidic oligosaccharides (Kunz, C., Rudloff, S., Baier, W., Klein, N., and Strobel, S. (2000). Annu Rev Nutr *20*, 699-722; Bode, L. (2006). J Nutr *136*, 2127-130). More than 130 different complex oligosaccharides have been identified in human milk, and their structural diversity and abundance is unique to humans. Although these molecules may not be utilized directly

by infants for nutrition, they nevertheless serve critical roles in the establishment of a healthy gut microbiome (Marcobal, A., Barboza, M., Froehlich, J. W., Block, D. E., et al. J Agric Food Chem 58, 5334-5340 (2010)), in the prevention of disease (Newburg, D. S., Ruiz-Palacios, G. M. & Morrow, A. L. Annu Rev Nutr 25, 37-58 (2005)), and in immune function (Newburg, D. S. & Walker, W. A. Pediatr Res 61, 2-8 (2007)). Despite millions of years of exposure to human milk oligosaccharides (HMOS), pathogens have yet to develop ways to circumvent the ability of HMOS to prevent adhesion to target cells and to inhibit infection. The ability to utilize HMOS as pathogen adherence inhibitors promises to address the current crisis of burgeoning antibiotic resistance. Human milk oligosaccharides produced by biosynthesis represent the lead compounds of a novel class of therapeutics against some of the most intractable scourges of society.

One alternative strategy for efficient, industrial-scale synthesis of HMOS is the metabolic engineering of bacteria. This approach involves the construction of microbial strains overexpressing heterologous glycosyltransferases, membrane transporters for the import of precursor sugars into the bacterial cytosol, and possessing enhanced pools of regenerating nucleotide sugars for use as biosynthetic precursors (Dumon, C., Samain, E., and Priem, B. (2004). Biotechnol Prog 20, 412-19; Ruffing, A., and Chen, R.R. (2006). Microb Cell Fact 5, 25). A key aspect of this approach is the heterologous glycosyltransferase selected for overexpression in the microbial host. The choice of glycosyltransferase can significantly affect the final yield of the desired synthesized oligosaccharide, given that enzymes can vary greatly in terms of kinetics, substrate specificity, affinity for donor and acceptor molecules, stability and solubility. A few glycosyltransferases derived from different bacterial species have been identified and characterized in terms of their ability to catalyze the biosynthesis of HMOS in *E. coli* host strains (Dumon, C., Bosso, C., Utille, J.P., Heyraud, A., and Samain, E. (2006). Chembiochem 7, 359-365; Dumon, C., Samain, E., and Priem, B. (2004). Biotechnol Prog 20, 412-19; Li, M., Liu, X.W., Shao, J., Shen, J., Jia, Q., Yi, W., Song, J.K., Woodward, R., Chow, C.S., and Wang, P.G. (2008). Biochemistry 47, 378-387). The identification of additional glycosyltransferases with faster kinetics, greater affinity for nucleotide sugar donors and/or acceptor molecules, or greater stability within the bacterial host significantly improves the yields of therapeutically useful HMOS. Prior to the invention described herein, chemical syntheses of HMOS were possible, but were limited by stereo-specificity issues, precursor availability, product impurities, and high overall cost (Flowers, H. M. Methods Enzymol 50, 93-121 (1978); Seeberger, P. H. Chem Commun (Camb) 1115-1121 (2003);

Koeller, K. M. & Wong, C. H. Chem Rev 100, 4465-4494 (2000)). The invention overcomes the shortcomings of these previous attempts by providing new strategies to inexpensively manufacture large quantities of human milk oligosaccharides (HMOS) for use as dietary supplements. Advantages include efficient expression of the enzyme, improved stability and/or solubility of the fucosylated oligosaccharide product (2'-FL, LDFT, LNF I, and LDFH I) and reduced toxicity to the host organism. The present invention features novel α(1,2) FTs suitable for expression in production strains for increased efficacy and yield of fucosylated HMOS compared to α(1,2) FTs currently utilized in the field.

As described in detail below, *E. coli* (or other bacteria) is engineered to produce selected fucosylated oligosaccharides (*i.e.*, 2'-FL, LDFT, LDHF I, or LNF I) in commercially viable levels. For example, yields are >5 grams/liter in a bacterial fermentation process. In other embodiments, the yields are greater than 10 grams/liter, greater than 15 grams/liter, greater than 20 grams/liter, greater than 25 grams/liter, greater than 30 grams/liter, greater than 35 grams/liter, greater than 40 grams/liter, greater than 45 grams/liter, greater than 50 grams/liter, greater than 55 grams/liter, greater than 60 grams/liter, greater than 65 grams/liter, greater than 70 grams/liter, or greater than 75 grams/liter of fucosylated oligosaccharide products, such as 2'-FL, LDFT, LDHF I, and LNF I.

Role of Human milk glycans in infectious disease

Human milk glycans, which comprise both unbound oligosaccharides and their glycoconjugates, play a significant role in the protection and development of the infant gastrointestinal (GI) tract. Neutral fucosylated oligosaccharides, including 2'-fucosyllactose (2'-FL), protect infants against several important pathogens. Milk oligosaccharides found in various mammals differ greatly, and the composition in humans is unique (Hamosh M., 2001 Pediatr Clin North Am, 48:69-86; Newburg D.S., 2001 Adv Exp Med Biol, 501:3-10). Moreover, glycan levels in human milk change throughout lactation and also vary widely among individuals (Morrow A.L. et al., 2004 J Pediatr, 145:297-303; Chaturvedi P et al., 2001 Glycobiology, 11:365-372). Approximately 200 distinct human milk oligosaccharides have been identified and combinations of simple epitopes are responsible for this diversity (Newburg D.S., 1999 Curr Med Chem, 6:117-127; Ninonuevo M. et al., 2006 J Agric Food Chem, 54:7471-7480l).

Human milk oligosaccharides are composed of 5 monosaccharides: D-glucose (Glc), D-galactose (Gal), N-acetylglucosamine (GlcNAc), L-fucose (Fuc), and sialic acid (N-acetyl neuraminic acid, Neu5Ac, NANA). Human milk oligosaccharides are usually divided into

two groups according to their chemical structures: neutral compounds containing Glc, Gal, GlcNAc, and Fuc, linked to a lactose (Galβ1-4Glc) core, and acidic compounds including the same sugars, and often the same core structures, plus NANA (Charlwood J. et al., 1999 Anal Biochem, 273:261-277; Martín-Sosa et al., 2003 J Dairy Sci, 86:52-59; Parkkinen J. and Finne J., 1987 Methods Enzymol, 138:289-300; Shen Z. et al., 2001 J Chromatogr A, 921:315-321).

Approximately 70-80% of oligosaccharides in human milk are fucosylated, and their synthetic pathways are believed to proceed as shown in FIG. 1. A smaller proportion of the oligosaccharides are sialylated or both fucosylated and sialylated, but their synthetic pathways are not fully defined. Understanding of the acidic (sialylated) oligosaccharides is limited in part by the ability to measure these compounds. Sensitive and reproducible methods for the analysis of both neutral and acidic oligosaccharides have been designed. Human milk oligosaccharides as a class survive transit through the intestine of infants very efficiently, being essentially indigestible (Chaturvedi, P., Warren, C. D., Buescher, C. R., Pickering, L. K. & Newburg, D. S. Adv Exp Med Biol 501, 315-323 (2001)).

Human milk glycans inhibit binding of enteropathogens to their receptors

Human milk glycans have structural homology to cell receptors for enteropathogens and function as receptor decoys. For example, pathogenic strains of *Campylobacter* bind specifically to glycans containing H-2, i.e., 2'-fucosyl-*N*-acetyllactosamine or 2'-fucosyllactose (2'FL); *Campylobacter* binding and infectivity are inhibited by 2'-FL and other glycans containing this H-2 epitope. Similarly, some diarrheagenic *E. coli* pathogens are strongly inhibited *in vivo* by human milk oligosaccharides containing 2-linked fucose moieties. Several major strains of human caliciviruses, especially the noroviruses, also bind to 2-linked fucosylated glycans, and this binding is inhibited by human milk 2-linked fucosylated glycans. Consumption of human milk that has high levels of these 2-linked fucosyloligosaccharides was associated with lower risk of norovirus, *Campylobacter*, ST of *E. coli*-associated diarrhea, and moderate-to-severe diarrhea of all causes in a Mexican cohort of breastfeeding children (Newburg D.S. et al., 2004 Glycobiology, 14:253-263; Newburg D.S. et al., 1998 Lancet, 351:1160-1164). Several pathogens utilize sialylated glycans as their host receptors, such as influenza (Couceiro, J. N., Paulson, J. C. & Baum, L. G. Virus Res 29, 155-165 (1993)), parainfluenza (Amonsen, M., Smith, D. F., Cummings, R. D. & Air, G. M. J Virol 81, 8341-8345 (2007), and rotoviruses (Kuhlenschmidt, T. B., Hanafin, W. P., Gelberg, H. B. & Kuhlenschmidt, M. S. Adv Exp Med Biol 473, 309-317 (1999)). The

sialyl-Lewis X epitope is used by *Helicobacter pylori* (Mahdavi, J., Sondén, B., Hurtig, M., Olfat, F. O., et al. Science 297, 573-578 (2002)), *Pseudomonas aeruginosa* (Scharfman, A., Delmotte, P., Beau, J., Lamblin, G., et al. Glycoconj J 17, 735-740 (2000)), and some strains of noroviruses (Rydell, G. E., Nilsson, J., Rodriguez-Diaz, J., Ruvoën-Clouet, N., et al. Glycobiology 19, 309-320 (2009)).

Identification of novel α(1,2) fucosyltransferases

The present invention provides novel α(1,2) fucosyltransferase enzymes. The present invention also provides nucleic acid constructs (*i.e.*, a plasmid or vector) carrying the nucleic acid sequence of a novel α(1,2) fucosyltransferases for the expression of the novel α(1,2) fucosyltransferases in host bacterium. The present invention also provides methods for producing fucosylated oligosaccharides by expressing the novel α(1,2) fucosyltransferases in suitable host production bacterium, as further described herein.

Not all α(1,2)fucosyltransferases can utilize lactose as an acceptor substrate. An acceptor substrate includes, for example, a carbohydrate, an oligosaccharide, a protein or glycoprotein, a lipid or glycolipid, *e.g.*, N-acetylglucosamine, N-acetyllactosamine, galactose, fucose, sialic acid, glucose, lactose, or any combination thereof. A preferred alpha (1,2) fucosyltransferase of the present invention utilizes GDP-fucose as a donor, and lactose is the acceptor for that donor.

A method of identifying novel α(1,2)fucosyltransferase enzymes capable of utilizing lactose as an acceptor was previously carried out (as described in PCT/US2013/051777, hereby incorporated by reference in its entirety) using the following steps: 1) performing a computational search of sequence databases to define a broad group of simple sequence homologs of any known, lactose-utilizing α(1,2)fucosyltransferase; 2) using the list of homologs from step 1 to derive a search profile containing common sequence and/or structural motifs shared by the members of the broad group, e.g. by using computer programs such as MEME (Multiple Em for Motif Elicitation available at http://meme.sdsc.edu/meme/cgi-bin/meme.cgi) or PSI-BLAST (Position-Specific Iterated BLAST available at ncbi.nlm.nih.gov/blast with additional information at cnx.org/content/m11040/latest/); 3) searching sequence databases (e.g., using computer programs such as PSI-BLAST, or MAST (Motif Alignment Search Tool available at http://meme.sdsc.edu/meme/cgi-bin/mast.cgi);, using this derived search profile as query, and identifying "candidate sequences" whose simple sequence homology to the original lactose-accepting α(1,2)fucosyltransferase is 40% or less; 4) scanning the scientific literature and

developing a list of "candidate organisms" known to express α(1,2)fucosyl-glycans; 5) selecting only those "candidate sequences" that are derived from "candidate organisms" to generate a list of "candidate lactose-utilizing enzymes"; and 6) expressing each "candidate lactose-utilizing enzyme" and testing for lactose-utilizing α(1,2)fucosyltransferase activity.

The MEME suite of sequence analysis tools (meme.sdsc.edu/meme/cgi-bin/meme.cgi) can also be used as an alternative to PSI-BLAST. Sequence motifs are discovered using the program "MEME". These motifs can then be used to search sequence databases using the program "MAST". The BLAST and PSI-BLAST search algorithms are other well known alternatives.

To identify additional novel α(1,2)fucosyltransferases, a multiple sequence alignment query was generated using four previously identified lactose-utilizing α(1,2)fucosyltransferase protein sequences: *H. pylori* futC (SEQ ID NO: 1), *H. mustelae* FutL (SEQ ID NO: 2), *Bacteroides vulgatus* futN (SEQ ID NO: 3), and *E. coli 0126* wbgL (SEQ ID NO: 4). These sequence alignment and percentage of sequence identity is shown in FIG. 3. An iterative PSI-BLAST was performed, using the FASTA-formatted multiple sequence alignment as the query, and the NCBI PSI-BLAST program run on a local copy of NCBI BLAST+ version 2.2.29. An initial position-specific scoring matrix file (.pssm) was generated by PSI-BLAST, which the program then used to adjust the score of iterative homology search runs. The process is iterated to generate an even larger group of candidates, and the results of each run were used to further refine the matrix.

This PSI-BLAST search resulted in an initial 2515 hits. There were 787 hits with greater than 22% sequence identity to FutC. 396 hits were of greater than 275 amino acids in length. Additional analysis of the hits was performed, including sorting by percentage identity to FutC, comparing the sequences by BLAST to existing α(1,2) fucosyltransferase inventory (of known α(1,2) fucosyltransferases), and manual annotation of hit sequences to identify those originating from bacteria that naturally exist in the gastrointestinal tract. An annotated list of the novel α(1,2) fucosyltransferases identified by this screen are listed in Table 1. Table 1 provides the bacterial species from which the candidate enzyme is found, the GenBank Accession Number, GI Identification Number, amino acid sequence, and % sequence identity to FutC.

Of the identified hits, 12 novel α(1,2) fucosyltransferases were further analyzed for their functional capacity: *Prevotella melaninogenica* FutO, *Clostridium bolteae* FutP, *Clostridium bolteae* +13 FutP, *Lachnospiraceae sp.* FutQ, *Methanosphaerula palustries*

FutR, *Tannerella sp.* FutS, *Bacteroides caccae* FutU, *Butyrivibrio* FutV, *Prevotellaa sp.* FutW, *Parabacteroides johnsonii* FutX, *Akkermansia muciniphilia* FutY, *Salmonella enterica* FutZ, and *Bacteroides sp.* FutZA. For *Clostridium bolteae* FutP, the annotation named the wrong initiation methionine codon. Thus, the present invention includes FutP with an additional 13 amino acids at the N-terminus of the annotated FutP (derived in-frame from the natural upstream DNA sequence), which is designated herein as *Clostridium bolteae* +13 FutP. The sequence identity between the 12 novel α(1,2) fucosyltransferases identified and the 4 previously identified α(1,2) fucosyltransferases is shown in Table 2 below.

Table 2. Sequence Identity

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
H.pylori futC 1		70.10	21.99	20.82	27.68	27.36	23.56	23.28	23.62	25.75	23.72	24.05	22.29	24.19	22.92	22.29
H.mustelae futL 2	70.10		23.87	19.88	26.38	28.21	24.30	23.38	24.67	25.31	25.31	24.47	23.56	25.15	23.55	23.26
Bacteroides vulgatus futN 3	21.99	23.87		25.16	32.05	28.71	28.04	25.79	37.46	32.27	26.11	61.27	71.63	27.67	25.15	84.75
E. coli O126 wbqL 4	20.82	19.88	25.16		24.25	22.73	22.32	26.04	25.45	24.77	21.49	23.29	26.71	24.83	21.45	25.16
Prevotella melaninogenica FutO WP_003814512.1 5	27.68	26.38	32.05	24.25		36.96	31.63	35.74	35.16	55.74	30.28	30.03	32.80	30.09	26.28	31.83
Clostridium bolteae+13 FutP WP_002570768.1 6	27.36	28.21	28.71	22.73	36.96		37.87	35.10	33.77	36.91	35.74	29.58	31.39	27.67	26.33	29.13
Lachnospiraceae sp. FutQ WP_009251343.1 7	23.56	24.30	28.04	22.32	31.63	37.87		29.87	29.17	32.90	51.02	28.53	30.00	27.69	24.00	27.74
Methanosphaerula palustris FutR YP_002467213.1 8	23.28	23.38	25.79	26.04	35.74	35.10	29.87		28.71	38.24	31.41	25.39	28.08	30.65	23.93	25.53
Tannerella sp. FutS WP_021929367.1 9	23.62	24.62	37.46	25.45	35.16	33.77	29.17	28.71		34.41	30.03	35.71	36.27	26.48	21.75	36.60
Bacteroides caccae FutU WP_005675707.1 10	25.75	25.31	32.27	24.77	55.74	36.91	32.90	38.24	34.41		31.21	29.94	33.33	29.28	24.46	33.01
Butyrivibrio FutV WP_022772718.1 11	23.72	25.31	26.11	21.49	30.28	35.74	51.02	31.41	30.03	31.21		27.62	26.20	26.46	22.15	26.52
Prevotella sp. FutW WP_022481266.1 12	24.05	24.47	61.27	23.29	30.03	29.58	28.53	25.39	35.71	29.94	27.62		57.60	25.79	22.15	59.01
Parabacteroides johnsonii FutX WP_008155883.1 13	22.29	23.56	71.63	26.71	32.80	31.39	30.00	26.08	36.27	33.33	26.20	57.60		28.71	24.00	74.02
Akkermansia muciniphilia FutY YP_001877555 14	24.19	25.15	27.67	24.83	30.09	27.67	27.69	30.65	26.48	29.28	26.46	25.79	28.71		21.45	28.08
Salmonella enterica FutZ WP_023214330 15	22.92	23.55	25.15	21.45	26.28	26.33	24.00	23.93	21.75	24.46	22.15	22.15	24.00	21.45		24.62
Bacteroides sp. FutZA WP_022161880.1 16	22.29	23.26	84.75	25.16	31.83	29.13	27.74	25.53	36.60	33.01	26.52	59.01	74.02	28.08	24.62	

Based on the amino acid sequences of the identified a(1,2) fucosyltransferases (*i.e.*, in Table 1), syngenes can be readily designed and constructed by the skilled artisan using standard methods known in the art. For example, the syngenes include a ribosomal binding site, are codon-optimized for expression in a host bacterial production strain (*i.e.*, *E. coli*), and have common 6-cutter restriction sites or sites recognized by endogenous restriction enzymes present in the host strain (*i.e.*, EcoK restriction sites) removed to ease cloning and expression in the *E. coli* host strain. In a preferred embodiment, the syngenes are constructed with the following configuration: EcoRI site – T7g10 RBS – α(1,2) FT syngene – XhoI site. The nucleic acid sequences of sample syngenes for the 12 identified α(1,2) fucosyltransferases are shown in Table 3. (the initiating methionine ATG codon is bolded)

Table 3. Nucleic acid sequences of 12 novel α(1,2) fucosyltransferase syngenes

Bacteria/ Gene name	Sequence	SEQ ID NO:
FutO	CAGTCAGTCAGAATTCAAGAAGGAGATATACAT**ATG**AAAATCGTCAAAATCCTGGGCGGT CTGGGCAATCAGATGTTCCAGTATGCTCTGTACCTGAGCCTGCAAGAAAGTTTTCCAAAA GAACGTGTGGCCCTGGACCTGTCCTCCTTCCACGGCTATCACCTGCATAATGGCTTTGAG CTGGAGAACATTTTCTCCGTTACCGCTCAGAAAGCATCCGCCGCAGATATCATGCGTATT GCTTATTACTACCCGAACTATCTGCTGTGGCGCATTGGCAAACGTTTTCTGCCGCGTCGT AAAGGTATGTGCCTGGAATCTAGCTCCCTGCGTTTCGATGAAAGCGTTCTGCGTCAGGAA GGTAACCGTTATTTTGACGGTTACTGGCAAGACGAACGCTACTTCGCAGCCTATCGTGAA	276

	AAAGTGCTGAAGGCTTTCACCTTTCCTGCATTCAAACGCGCAGAAAACCTGAGCCTGCTG GAAAAACTGGACGAAAACAGCATTGCTCTGCATGTTCGTCGCGGTGATTACGTAGGTAAT AACCTGTACCAAGGCATCTGTGACCTGGACTACTACCGTACCGCTATCGAGAAAATGTGT GCACACGTTACTCCGTCTCTGTTTTGTATCTTTTCCAACGACATCACGTGGTGCCAGCAG CACCTGCAACCGTACCTGAAGGCCCCTGTGGTGTACGTTACTTGGAACACCGGTGTTGAA TCCTACCGCGATATGCAGCTGATGTCCTGCTGCGCACATAACATCATCGCGAATAGCTCC TTCTCTTGGTGGGGTGCTTGGCTGAATCAGAACCGTGAAAAAGTTGTTATCGCCCCGAAA AAATGGCTGAACATGGAAGAATGTCACTTCACGCTGCCGGCAAGCTGGATCAAAATTTAG CTCGAGTGACTGACTG	
FutP	CAGTCAGTCAGAATTCAAGAAGGAGATATACAT**ATG**GTGATTATCAAAATGATGGGTGGT CTGGGCAACCAGATGTTCCAGTACGCACTGTACAAAGCATTCGAGCAGAAGCACATCGAT GTGTATGCAGACCTGGCATGGTACAAAAACAAATCCGTGAAATTTGAACTGTACAACTTC GGCATTAAAATCAACGTAGCATCCGAGAAAGACATCAACCGTCTGAGCGATTGCCAGGCG GACTTTGTTTCCCGCATCCGCCGTAAAATCTTTGGTAAAAAAAAGAGCTTCGTATCTGAA AAAAATGACTCCTGCTATGAAAACGACATCCTGCGTATGGACAACGTTTATCTGAGCGGT TATTGGCAGACCGAAAAATACTTCTCTAACACGCGTGAGAAGCTGCTGGAGGATTATTCC TTCGCTCTGGTAAACTCTCAGGTGTCCGAATGGGAAGACTCCATTCGCAACAAAAACAGC GTTAGCATCCATATCCGTCGTGGTGATTATCTACAGGGCGAACTGTATGGTGGTATTTGC ACCTCTCTGTACTACGCCGAAGCAATCGAGTACATTAAAATGCGTGTTCCGAACGCAAAA TTCTTCGTTTTCTCTGATGACGTTGAATGGGTTAAACAGCAAGAAGACTTCAAAGGCTTC GTAATCGTTGATCGCAACGAGTATTCTAGCGCTCTGTCCGATATGTACCTGATGTCCCTG TGCAAGCATAACATTATTGCTAACTCCTCTTTCAGCTGGTGGGCAGCTTGGCTGAACCGT AACGAAGAAAAAATTGTAATCGCGCCGCGCCGTTGGCTGAACGGCAAGTGCACCCCAGAT ATCTGGTGTAAAAAATGGATTCGTATCTAGCTCGAGTGACTGACTG	277
FutQ	CAGTCAGTCAGAATTCAAGAAGGAGATATACAT**ATG**GTGATCGTACAGCTGAGCGGCGGT CTGGGCAACCAGATGTTCGAATACGCGCTGTACCTGAGCCTGAAAGCAAAAGGCAAAGAA GTGAAAATTGACGATGTTACGTGTTACGAGGGCCCTGGCACCCGTCCGCGTCAACTGGAT GTTTTTGGTATCACGTACGATCGCGCGTCTCGTGAGGAGCTGACTGAGATGACGGACGCG AGCATGGATGCGCTGTCTCGTGTTCGTCGCAAACTGACCGGTCGCCGCACTAAAGCGTAC CGCGAACGCGACATCAACTTCGATCCACTGGTTATGGAAAAAGACCCGGCACTGCTGGAA GGCTGTTTCCAGTCTGACAAATACTTTCGTGATTGCGAAGGCCGCGTGCGCGAAGCGTAT CGTTTCCGCGGCATTGAATCCGGCGCGTTCCCGCTGCCGGAAGACTATCTGCGCCTGGAA AAGCAGATCGAAGATTGTCAGTCCGTATCCGTACACATCCGTCGTGGCGACTACCTGGAC GAATCTCATGGTGGTCTGTACACCGGCATTTGTACTGAGGCGTACTATAAAGAGGCTTTT GCTCGCATGGAACGTCTGGTTCCGGGCGCACGTTTCTTCCTGTTCTCTAACGATCCAGAA TGGACTCGTGAGCACTTTGAGAGCAAGAACTGCGTTCTGGTTGAAGGTAGCACCGAAGAC ACGGGTTACATGGACCTGTACCTGATGAGCCGCTGCCGCCACAATATTATTGCCAACTCT TCTTTCAGCTGGTGGGGCGCTTGGCTGAATGAGAACCCTGAGAAAAAAGTCATCGCACCG GCTAAATGGCTGAACGGTCGTGAGTGCCGTGATATCTATACCGAACGCATGATTCGTCTG TAGCTCGAGTGACTGACTG	278
FutR	CAGTCAGTCAGAATTCAAGAAGGAGATATACAT**ATG**ATCATTGTTCGTCTGAAAGGCGGT CTGGGCAACCAACTGTCTCAGTATGCACTGGGCCGTAAGATCGCGCATCTGCACAATACC GAACTGAAACTGGACACCACTTGGTTCACCACTATCTCCTCCGACACTCCACGTACCTAC CGTCTGAACAATTATAACATCATCGGCACTATTGCATCCGCAAAGGAAATCCAGCTGATC GAACGTGGTCGCGCGCAAGGCCGTGGCTACCTGCTGTCTAAAATTTCTGATCTGCTGACT CCGATGTACCGTCGTACCTACGTCCGTGAACGTATGCATACCTTCGATAAAGCTATCCTG ACCGTTCCGGACAACGTGTACCTGGATGGTTACTGGCAGACCGAAAAGTACTTCAAAGAC ATCGAAGAAATCCTGCGCCGTGAGGTTACGCTGAAAGATGAACCGGATAGCATCAACCTG GAAATGGCTGAACGTATTCAGGCTTGCCACAGCGTTTCCCTGCACGTGCGTCGTGGCGAC TACGTTTCCAACCCGACCACTCAACAATTCCACGGCTGTTGCTCCATTGACTACTACAAC CGCGCTATCTCTCTGATTGAAGAAAAAGTGGATGACCCGTCTTTCTTTATTTTTTCTGAC GATCTGCCGTGGGCTAAAGAAAACCTGGACATCCCTGGCGAAAAAACCTTCGTTGCGCAT AACGGCCCGGAAAAAGAGTATTGCGATCTGTGGCTGATGTCTCTGTGCCAGCACCATATC ATCGCAAACTCTTCTTTCAGCTGGTGGGGTGCCTGGCTGGGTCAAGACGCCGAAAAGATG GTGATCGCGCCGCGTCGCTGGGCCCTGTCCGAGAGCTTTGACACTTCTGACATCATTCCG GACTCTTGGATTACTATCTAGCTCGAGTGACTGACTG	279
FutS	CAGTCAGTCAGAATTCAAGAAGGAGATATACAT**ATG**GTACGCATTGTGGAAATCATCGGC GGTCTGGGTAACCAGATGTTCCAGTACGCATTCTCCCTGTACCTGAAAAACAAATCTCAC ATCTGGGACCGTCTGTATGTGGACATCGAGGCGATGAAAACCTACGATCGTCACTATGGT CTGGAACTGGAGAAAGTTTTCAATCTGAGCCTGTGTCCAATCTCTAACCGTCTGCACCGC AACCTGCAAAAACGCTCCTTCGCAAAACACTTTGTAAAGAGCCTGTACGAGCACTCTGAA	280

	TGCGAGTTCGACGAACCGGTGTACCGTGGCCTGCGTCCTTATCGCTATTATCGCGGCTAC TGGCAAAACGAAGGTTACTTCGTTGATATTGAACCGATGATCCGTGAGGCTTTTCAGTTC AACGTTAACATCCTGAGCAAAAAGACTAAAGCGATCGCATCCAAAATGCGCCGTGAACTG TCCGTATCTATCCATGTTCGCCGTGGTGATTACGAAAACCTGCCGGAAGCGAAAGCGATG CATGGCGGTATTTGTTCTCTGGACTATTACCACAAAGCGATCGACTTCATCCGCCAGCGT CTGGATAATAACATCTGTTTCTATCTGTTCTCCGACGATATCAATTGGGTAGAAGAAAAC CTGCAACTGGAAAACCGTTGCATCATCGACTGGAACCAGGGCGAAGATAGCTGGCAGGAC ATGTACCTGATGAGCTGCTGCCGCCACCACATTATCGCAAACAGCTCTTTCTCCTGGTGG GCGGCATGGCTGAATCCAAACAAGAACAAAATCGTACTGACCCCGAACAAATGGTTCAAC CATACTGACGCAGTGGGTATCGTCCCAAAGTCCTGGATTAAAATTCCTGTGTTTTAGCTC GAGTGACTGACTG	
FutU	CAGTCAGTCAGAATTCAAGAAGGAGATATACAT**ATG**AAAATCGTTAAAATCCTGGGCGGC CTGGGTAACCAGATGTTTCAGTACGCCCTGTTCCTGTCTCTGAAAGAACGCTTCCCGCAT GAACAGGTGATGATTGACACCAGCTGCTTCCGCAATTACCCACTGCACAACGGTTTCGAA GTGGATCGTATCTTCGCCCAGAAAGCACCGGTTGCCTCTTGGCGTAACATCCTGAAGGTT GCCTACCCGTACCCGAACTACCGTTTCTGGAAAATCGGTAAATACATCCTGCCTAAACGT AAAACCATGTGTGTAGAGCGTAAAAACTTCAGCTTTGACGCCGCAGTCCTGACCCGTAAA GGCGATTGCTACTATGATGGCTACTGGCAGCATGAGGAATATTTCTGTGATATGAAAGAA ACGATTTGGGAGGCTTTCTCCTTCCCTGAGCCGGTTGATGGTCGTAACAAGGAGATCGGT GCCCTGCTACAGGCATCTGATAGCGCTTCCCTGCACGTTCGTCGCGGTGACTACGTGAAC CACCCACTGTTTCGTGGTATTTGTGACCTGGACTATTATAAACGTGCCATCCACTACATG GAAGAACGCGTCAACCCACAGCTGTACTGCGTTTTCAGCAACGATATGGCCTGGTGCGAG TCCCACCTGCGTGCACTGCTGCCAGGCAAAGAAGTAGTTTATGTTGACTGGAACAAGGGT GCGGAATCTTACGTTGATATGCGTCTGATGAGCCTGTGCCGTCACAACATCATCGCTAAC TCTTCTTTCAGCTGGTGGGGCGCATGGCTGAACCGTAACCCGCAGAAAGTGGTGGTAGCG CCGGAACGTTGGATGAACAGCCCGATTGAAGACCCAGTGAGCGACAAATGGATTAAACTG TAGCTCGAGTGACTGACTG	281
FutV	CAGTCAGTCAGAATTCAAGAAGGAGATATACAT**ATG**ATCATCATCCAGCTGAAAGGTGGC CTGGGCAACCAAATGTTCCAGTACGCGCTGTACAAATCCCTGAAAAAACGTGGTAAAGAA GTTAAAATTGATGACAAAACTGGCTTCGTGAACGACAAACTGCGTATCCCGGTACTGTCC CGTTGGGGTGTTGAGTACGATCGTGCAACCGACGAAGAGATTATTAACCTGACCGACTCC AAAATGGACCTGTTCTCTCGCATCCGCCGTAAACTGACTGGCCGCAAAACGTTCCGTATC GACGAAGAATCCGGTAAATTCAACCCGGAAATCCTGGAAAAAGAGAACGCTTATCTGGTG GGTTATTGGCAGTGCGACAAGTACTTCGACGACAAAGATGTGGTTCGCGAAATTCGTGAA GCGTTCGAGAAAAAACCGCAGGAGCTGATGACCGACGCCAGCTCTTGGTCTACTCTACAG CAGATTGAATGCTGCGAGTCCGTATCCCTGCACGTACGTCGTACTGATTACGTGGACGAG GAACATATTCATATCCATAACATCTGTACGGAAAAATACTATAAAAACGCCATTGATCGT GTGCGTAAACAGTACCCGAGCGCAGTGTTCTTCATCTTCACCGATGATAAAGAATGGTGC CGCGACCACTTTAAAGGTCCGAACTTCATCGTAGTCGAACTGGAAGAAGGCGACGGTACC GACATCGCTGAAATGACTCTGATGTCCCGCTGTAAACATCACATCATCGCTAATTCTAGC TTTAGCTGGTGGGCGGCGTGGCTGAACGACTCCCCGGAAAAAATCGTGATCGCTCCTCAG AAATGGATTAACAACCGCGACATGGACGATATTTACACCGAGCGTATGACTAAAATCGCA CTGTAGCTCGAGTGACTGACTG	282
FutW	CAGTCAGTCAGAATTCAAGAAGGAGATATACAT**ATG**CGCCTGGTTAAAATGATCGGCGGT CTGGGTAATCAGATGTTCATCTACGCGTTTTACCTACAGATGCGTAAGCGTTTCTCCAAC GTTCGTATCGACCTGACCGATATGATGCACTACAACGTACACTATGGCTACGAACTGCAC AAAGTTTTCGGTCTGCCGCGCACCGAGTTCTGTATGAACCAGCCTCTGAAAAAGGTTCTG GAGTTCCTGTTCTTCCGTACCATTGTTGAACGTAAACAGCACGGTCGTATGGAGCCGTAT ACTTGCCAGTATGTTTGGCCGCTGGTTTACTTTAAGGGCTTCTATCAGTCCGAACGTTAC TTCTCCGAAGTTAAGGACGAAGTTCGTGAGTGTTTCACCTTCAATCCGGCACTGGCGAAT CGTTCTTCCCAACAGATGATGGAACAGATCCAGAATGATCCTCAGGCTGTCTCTATCCAC ATCCGTCGTGGCGACTATCTGAATCCGAAGCACTACGACACTATCGGTTGTATCTGTCAG CTGCCGTATTACAAGCACGCCGTTTCCGAAATTAAAAAGTACGTTTCTAACCCTCACTTT TACGTTTTCTCCGAAGACCTGGATTGGGTCAAAGCAAACCTGCCGCTGGAAAACGCACAG TACATCGATTGGAACAAAGGCGCAGATAGCTGGCAGGATATGATGCTGATGAGCTGTTGC AAACACCACATTATCTGTAACTCCACCTTTAGCTGGTGGGCGGCGTGGCTGAACCCATCT GTCGAAAAAACCGTGATCATGCCGGAACAGTGGACGTCTCGTCAAGATTCCGTGGACTTT GTGGCTAGCTGTGGCCGTTGGGTCCGTGTTAAAACGGAGTAGCTCGAGTGACTGACTG	283
FutX	CAGTCAGTCAGAATTCAAGAAGGAGATATACAT**ATG**CGTCTGATCAAGATGATCGGCGGC CTGGGTAACCAGATGTTTATCTACGCGTTCTACCTGAAAATGAAACACCATTACCCGGAT ACGAACATCGATCTGTCTGACATGGTTCATTATAAAGTTCACAACGGTTATGAGATGAAC	284

	CGTATCTTTGACCTGAGCCAGACTGAATTTTGCATCAACCGTACCCTGAAAAAAATCCTG GAGTTCCTGTTCTTCAAAAAAATCTACGAACGTCGCCAGGACCCGTCTACTCTGTATCCA TACGAAAAACGTTATTTTTGGCCGCTGCTGTACTTTAAAGGTTTCTACCAGTCTGAACGC TTCTTCTTCGATATCAAAGACGACGTTCGTAAAGCCTTCTCTTTTAACCTGAACATCGCT AACCCGGAAAGCCTGGAACTGCTGAAACAGATCGAAGTTGACGACCAAGCTGTTTCTATC CACATCCGCCGTGGTGACTACCTGCTGCCGCGTCACTGGGCAAACACGGGTTCCGTGTGC CAGCTGCCGTATTACAAGAACGCGATCGCGGAAATGGAGAACCGTATTACTGGCCCGAGC TACTACGTGTTCTCTGATGATATCTCTTGGGTTAAAGAAAACATCCCGCTGAAGAAAGCG GTCTACGTGACGTGGAACAAGGGCGAAGACAGCTGGCAGGATATGATGCTGATGAGCCAC TGTCGTCACCACATTATCTGTAATTCTACGTTCTCCTGGTGGGGTGCTTGGCTGAACCCA CGTAAAGAGAAAATCGTCATCGCGCCGTGTCGCTGGTTCCAGCATAAAGAAACCCCGGAC ATGTACCCGAAAGAATGGATCAAAGTACCGATTAACTAGCTCGAGTGACTGACTG	
FutZ	CAGTCAGTCAGAATTCAAGAAGGAGATATACAT**ATG**TATTCTTGCCTGTCTGGTGGCCTG GGTAACCAAATGTTTCAATACGCAGCAGCGTATATCCTGAAGCAGTATTTTCAGTCTACC ACTCTGGTCCTGGATGATAGCTATTACTATTCCCAGCCGAAACGTGATACCGTTCGTAGC CTGGAACTGAATCAGTTCAACATCTCTTATGATCGTTTTAGCTTCGCGGATGAAAAAGAG AAGATCAAACTGCTGCGCAAATTCAAACGTAACCCGTTCCCTAAACAGATTTCCGAGATC CTGTCTATTGCGCTGTTCGGCAAATACGCGCTGTCCGACCGTGCATTTTACACCTTCGAA ACTATCAAAAACATCGACAAAGCGTGCCTGTTCTCCTTTTACCAGGACGCCGATCTGCTG AATAAATATAAGCAGCTGATCCTGCCGCTGTTCGAACTGCGCGATGACCTGCTGGATATC TGCAAGAACCTGGAACTGTATTCCCTGATCCAACGCAGCAACAATACCACTGCACTGCAT ATCCGCCGTGGCGACTACGTGACCAACCAGCACGCCGCGAAATACCACGGCGTGCTGGAC ATCAGCTACTATAACCACGCAATGGAATACGTGGAACGTGAACGCGGCAAACAGAACTTC ATTATCTTCAGCGATGATGTACGTTGGGCACAGAAAGCGTTTCTGGAGAACGATAATTGC TACGTGATTAACAACTCCGACTACGATTTCTCTGCGATCGATATGTATCTGATGTCTCTG TGCAAAAACAACATCATCGCAAATTCCACCTACTCCTGGTGGGGTGCGTGGCTGAACAAA TACGAGGACAAACTGGTTATCTCTCCGAAACAATGGTTTCTGGGTAACAACGAAACCTCT CTGCGTAACGCGTCTTGGATCACCCTGTAGCTCGAGTGACTGACTG	285
FutZA	CAGTCAGTCAGAATTCAAGAAGGAGATATACAT**ATG**CGTCTGATCAAGATGACCGGTGGC CTGGGTAACCAGATGTTCATCTACGCGTTTTATCTGCGTATGAAAAAACGTTATCCGAAA GTTCGTATTGATCTGTCTGATATGGTTCATTATCACGTTCACCACGGCTATGAAATGCAC CGTGTTTTCAATCTGCCGCACACCGAATTTTGCATCAACCAGCCGCTGAAAAAAGTGATC GAGTTCCTGTTTTTCAAAAAGATTTACGAACGTAAACAGGACCCTAATTCTCTGCGTGCA TTCGAGAAGAAGTATCTGTGGCCGCTGCTGTACTTCAAAGGTTTCTATCAATCTGAGCGC TTCTTTGCTGACATCAAAGACGAGGTTCGTAAAGCATTCACCTTTGACTCTTCTAAAGTG AACGCTCGCTCTGCCGAACTGCTGCGTCGCCTGGATGCCGATGCTAACGCGGTTAGCCTG CACATTCGTCGCGGTGACTATCTACAGCCGCAGCATTGGGCTACCACTGGTTCTGTCTGC CAGCTGCCGTACTACCAGAACGCGATCGCTGAAATGAACCGTCGCGTTGCTGCCCCGAGC TACTACGTTTTCAGCGATGACATCGCGTGGGTGAAGGAAAACATCCCTCTACAGAACGCA GTGTACATCGACTGGAATAAAGGCGAAGAAAGCTGGCAGGATATGATGCTGATGAGCCAC TGCCGCCACCACATTATCTGTAACAGCACCTTCTCTTGGTGGGGCGCGTGGCTGGACCCG CACGAGGACAAAATTGTAATCGTTCCGAATCGTTGGTTCCAGCATTGCGAAACTCCTAAC ATCTATCCGGCAGGCTGGGTGAAAGTTGCGATTAATTAGCTCGAGTGACTGACTG	286

In any of the methods described herein, the α(1,2) fucosyltransferase genes or gene products may be variants or functional fragments thereof. A variant of any of genes or gene products disclosed herein may have 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to the nucleic acid or amino acid sequences described herein.

Variants as disclosed herein also include homolog, orthologs, or paralogs of the genes or gene products described herein that retain the same biological function as the genes or gene products specified herein. These variants can be used interchangeably with the genes recited in these methods. Such variants may demonstrate a percentage of homology or

identity, for example, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity conserved domains important for biological function, preferably in a functional domain, e.g. catalytic domain.

The term "% identity," in the context of two or more nucleic acid or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection. For example, % identity is relative to the entire length of the coding regions of the sequences being compared, or the length of a particular fragment or functional domain thereof.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

Percent identity is determined using search algorithms such as BLAST and PSI-BLAST (Altschul et al., 1990, J Mol Biol 215:3, 403-410; Altschul et al., 1997, Nucleic Acids Res 25:17, 3389-402). For the PSI-BLAST search, the following exemplary parameters are employed: (1) Expect threshold was 10; (2) Gap cost was Existence:11 and Extension:1; (3) The Matrix employed was BLOSUM62; (4) The filter for low complexity regions was "on".

Changes can be introduced by mutation into the nucleic acid sequence or amino acid sequence of any of the genes or gene products described herein, leading to changes in the amino acid sequence of the encoded protein or enzyme, without altering the functional ability of the protein or enzyme. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of any of sequences expressly disclosed herein. A "non-essential" amino acid residue is a residue at a position in the sequence that can be altered from the wild-type sequence of the polypeptide without altering the biological activity, whereas an "essential" amino acid residue is a residue at a position that is required for biological activity. For example, amino acid residues that are conserved among members of a family of proteins are not likely to be amenable to mutation. Other amino acid residues, however, (*e.g.*, those that are poorly conserved among members of the protein family) may not be as essential for activity and thus are more likely to be

amenable to alteration. Thus, another aspect of the invention pertains to nucleic acid molecules encoding the proteins or enzymes disclosed herein that contain changes in amino acid residues relative to the amino acid sequences disclosed herein that are not essential for activity (*i.e.*, fucosyltransferase activity).

An isolated nucleic acid molecule encoding a protein essentially retaining the functional capability compared to any of the genes described herein can be created by introducing one or more nucleotide substitutions, additions or deletions into the corresponding nucleotide sequence, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into a nucleic acid sequence by standard techniques such that the encoded amino acid sequence is altered, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. Certain amino acids have side chains with more than one classifiable characteristic. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, tryptophan, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tyrosine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a given polypeptide is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a given coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for given polypeptide biological activity to identify mutants that retain activity. Conversely, the invention also provides for variants with mutations that enhance or increase the endogenous biological activity. Following mutagenesis of the nucleic acid sequence, the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined. An increase, decrease, or elimination of a given biological activity of the variants disclosed herein can be readily measured by the ordinary person skilled in the art, *i.e.*, by measuring the capability for mediating oligosaccharide modification, synthesis, or degradation (via detection of the products).

The present invention also provides for functional fragments of the genes or gene products described herein. A fragment, in the case of these sequences and all others provided herein, is defined as a part of the whole that is less than the whole. Moreover, a fragment ranges in size from a single nucleotide or amino acid within a polynucleotide or polypeptide sequence to one fewer nucleotide or amino acid than the entire polynucleotide or polypeptide sequence. Finally, a fragment is defined as any portion of a complete polynucleotide or polypeptide sequence that is intermediate between the extremes defined above.

For example, fragments of any of the proteins or enzymes disclosed herein or encoded by any of the genes disclosed herein can be 10 to 20 amino acids, 10 to 30 amino acids, 10 to 40 amino acids, 10 to 50 amino acids, 10 to 60 amino acids, 10 to 70 amino acids, 10 to 80 amino acids, 10 to 90 amino acids, 10 to 100 amino acids, 50 to 100 amino acids, 75 to 125 amino acids, 100 to 150 amino acids, 150 to 200 amino acids, 200 to 250 amino acids, 250 to 300 amino acids, 300 to 350 amino acids, 350 to 400 amino acids, 400 to 450 amino acids, or 450 to 500 amino acids. The fragments encompassed in the present invention comprise fragments that retain functional fragments. As such, the fragments preferably retain the catalytic domains that are required or are important for functional activity. Fragments can be determined or generated by using the sequence information herein, and the fragments can be tested for functional activity using standard methods known in the art. For example, the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined. The biological function of said fragment can be measured by measuring ability to synthesize or modify a substrate oligosaccharide, or conversely, to catabolize an oligosaccharide substrate.

Within the context of the invention, "functionally equivalent", as used herein, refers to a gene or the resulting encoded protein variant or fragment thereof capable of exhibiting a substantially similar activity as the wild-type fucosyltransferase. Specifically, the fucosyltransferase activity refers to the ability to transfer a fucose sugar to an acceptor substrate via an alpha-(1,2)-linkage. As used herein, "substantially similar activity" refers to an activity level within 5%, 10%, 20%, 30%, 40%, or 50% of the wild-type fucosyltransferase.

To test for lactose-utilizing fucosylatransferase activity, the production of fucosylated oligossacharides (*i.e.*, 2'-FL) is evaluated in a host organism that expresses the candidate enzyme (or syngene) and which contains both cytoplasmic GDP-fucose and lactose pools. The production of fucosylated oligosaccharides indicates that the candidate enzyme-encoding sequence functions as a lactose-utilizing α(1,2)fucosyltransferase.

Engineering of *E. coli* to produce human milk oligosaccharide 2'-FL

Described herein is a gene screening approach, which was used to validate the novel α (1,2) fucosyltransferases (α (1,2) FTs) for the synthesis of fucosyl-linked oligosaccharides in metabolically engineered *E. coli*. Of particular interest are α (1,2) FTs that are capable of the synthesis of the HMOS 2'-fucosyllactose (2'-FL). 2'-FL is the most abundant fucosylated oligosaccharide present in human milk, and this oligosaccharide provides protection to newborn infants against infectious diarrhea caused by bacterial pathogens such as *Campylobacter jejuni* (Ruiz-Palacios, G.M., et al. (2003). J Biol Chem 278, 14112-120; Morrow, A.L. et al. (2004). J Pediatr 145, 297-303; Newburg, D.S. et al. (2004). Glycobiology 14, 253-263). Other α (1,2) FTs of interest are those capable of synthesis of HMOS lactodifucotetraose (LDFT), laco-N-fucopentaose I (LNFI), or lacto-N-difucohexaose I (LDFH I).

The synthetic pathway of fucosyl oligosaccharides of human milk is illustrated in FIG. 1. Structurally, 2'-FL consists of a fucose molecule α 1,2 linked to the galactose portion of lactose (Fucα1-2Galβ1-4Glc). An α (1,2) FT from *H. pylori* strain 26695 termed FutC has been utilized to catalyze the synthesis of 2'-FL in metabolically engineered *E. coli* (Drouillard, S. et al. (2006). Angew Chem Int Ed Engl 45, 1778-780).

Candidate α(1,2) FTs (*i.e.*, syngenes) were cloned by standard molecular biological techniques into an expression plasmid. This plasmid utilizes the strong leftwards promoter of bacteriophage λ (termed P_L) to direct expression of the candidate genes (Sanger, F. et al. (1982). J Mol Biol 162, 729-773). The promoter is controllable, e.g., a trp-cI construct is stably integrated the into the E.coli host's genome (at the ampC locus), and control is implemented by adding tryptophan to the growth media. Gradual induction of protein expression is accomplished using a temperature sensitive cI repressor. Another similar control strategy (temperature independent expression system) has been described (Mieschendahl et al., 1986, Bio/Technology 4:802-808). The plasmid also carries the *E. coli rcsA* gene to up-regulate GDP-fucose synthesis, a critical precursor for the synthesis of fucosyl-linked oligosaccharides. In addition, the plasmid carries a β-lactamase (*bla*) gene for maintaining the plasmid in host strains by ampicillin selection (for convenience in the laboratory) and a native *thyA* (thymidylate synthase) gene as an alternative means of selection in *thyA*$^-$ hosts. Alternative selectable markers include the proBA genes to complement proline auxotrophy (Stein et al., (1984), J Bacteriol 158:2, 696-700 (1984) or purA to complement adenine auxotrophy (S. A. Wolfe, J. M. Smith, *J Biol Chem* **263**, 19147-53

(1988)). To act as plasmid selectable markers each of these genes are first inactivated in the host cell chromosome, then wild type copies of the genes are provided on the plasmid. Alternatively a drug resistance gene may be used on the plasmid, e.g. beta-lactamase (this gene is already on the expression plasmid described above, thereby permitting selection with ampicillin). Ampicillin selection is well known in the art and described in standard manuals such as Maniatis et al., (1982) Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring, NY.

The nucleic acid sequence of such an expression plasmid, pEC2-(T7)FutX-rcsA-thyA (pG401) is provided below. The underlined sequence represents the FutX syngene, which can be readily replaced with any of the novel α(1,2) FTs described herein using standard recombinant DNA techniques.

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGT
CTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGG
CTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATATGCGGTGTGAAATACC
GCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCTCCTCAACCTGTATATTCGTAAACCACGCC
CAATGGGAGCTGTCTCAGGTTTGTTCCTGATTGGTTACGGCGCGTTTCGCATCATTGTTGAGTTTTTC
CGCCAGCCCGACGCGCAGTTTACCGGTGCCTGGGTGCAGTACATCAGCATGGGGCAAATTCTTTCCAT
CCCGATGATTGTCGCGGGTGTGATCATGATGGTCTGGGCATATCGTCGCAGCCCACAGCAACACGTTT
CCTGAGGAACCATGAAACAGTATTTAGAACTGATGCAAAAAGTGCTCGACGAAGGCACACAGAAAAAC
GACCGTACCGGAACCGGAACGCTTTCCATTTTTGGTCATCAGATGCGTTTTAACCTGCAAGATGGATT
CCCGCTGGTGACAACTAAACGTTGCCACCTGCGTTCCATCATCCATGAACTGCTGTGGTTTCTGCAGG
GCGACACTAACATTGCTTATCTACACGAAAACAATGTCACCATCTGGGACGAATGGGCCGATGAAAAC
GGCGACCTCGGGCCAGTGTATGGTAAACAGTGGCGCGCCTGGCCAACGCCAGATGGTCGTCATATTGA
CCAGATCACTACGGTACTGAACCAGCTGAAAAACGACCCGGATTCGCGCCGCATTATTGTTTCAGCGT
GGAACGTAGGCGAACTGGATAAAATGGCGCTGGCACCGTGCCATGCATTCTTCCAGTTCTATGTGGCA
GACGGCAAACTCTCTTGCCAGCTTTATCAGCGCTCCTGTGACGTCTTCCTCGGCCTGCCGTTCAACAT
TGCCAGCTACGCGTTATTGGTGCATATGATGGCGCAGCAGTGCGATCTGGAAGTGGGTGATTTTGTCT
GGACCGGTGGCGACACGCATCTGTACAGCAACCATATGGATCAAACTCATCTGCAATTAAGCCGCGAA
CCGCGTCCGCTGCCGAAGTTGATTATCAAACGTAAACCCGAATCCATCTTCGACTACCGTTTCGAAGA
CTTTGAGATTGAAGGCTACGATCCGCATCCGGGCATTAAAGCGCCGGTGGCTATCTAATTACGAAACA
TCCTGCCAGAGCCGACGCCAGTGTGCGTCGGTTTTTTTACCCTCCGTTAAATTCTTCGAGACGCCTTC
CCGAAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGC
TATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCC
CAGTCACGACGTTGTAAAACGACGGCCAGTGCCAAGCTTTCTTTAATGAAGCAGGGCATCAGGACGGT
ATCTTTGTGGAGAAAGCAGAGTAATCTTATTCAGCCTGACTGGTGGGAAACCACCAGTCAGAATGTGT
TAGCGCATGTTGACAAAAATACCATTAGTCACATTATCCGTCAGTCGGACGACATGGTAGATAACCTG
TTTATTATGCGTTTTGATCTTACGTTTAATATTACCTTTATGCGATGAAACGGTCTTGGCTTTGATAT
TCATTTGGTCAGAGATTTGAATGGTTCCCTGACCTGCCATCCACATTCGCAACATACTCGATTCGGTT
CGGCTCAATGATAACGTCGGCATATTTAAAAACGAGGTTATCGTTGTCTCTTTTTTCAGAATATCGCC
AAGGATATCGTCGAGAGATTCCGGTTTAATCGATTTAGAACTGATCAATAAATTTTTTCTGACCAATA
GATATTCATCAAAATGAACATTGGCAATTGCCATAAAAACGATAAATAACGTATTGGGATGTTGATTA
ATGATGAGCTTGATACGCTGACTGTTAGAAGCATCGTGGATGAAACAGTCCTCATTAATAAACACCAC
TGAAGGGCGCTGTGAATCACAAGCTATGGCAAGGTCATCAACGGTTTCAATGTCGTTGATTTCTCTTT
TTTTAACCCCTCTACTCAACAGATACCCGGTTAAACCTAGTCGGGTGTAACTACATAAATCCATAATA
ATCGTTGACATGGCATACCCTCACTCAATGCGTAACGATAATTCCCCTTACCTGAATATTTCATCATG
ACTAAACGGAACAACATGGGTCACCTAATGCGCCACTCTCGCGATTTTTCAGGCGGACTTACTATCCC
GTAAAGTGTTGTATAATTTGCCTGGAATTGTCTTAAAGTAAAGTAAATGTTGCGATATGTGAGTGAGC
TTAAAACAAATATTTCGCTGCAGGAGTATCCTGGAAGATGTTCGTAGAAGCTTACTGCTCACAAGAAA

AAAGGCACGTCATCTGACGTGCCTTTTTTATTTGTACTACCCTGTACGATTACTGCAGCTCGAGCTAG
TTAATCGGTACTTTGATCCATTCTTTCGGGTACATGTCCGGGGTTTCTTTATGCTGGAACCAGCGACA
CGGCGCGATGACGATTTTCTCTTTACGTGGGTTCAGCCAAGCACCCCACCAGGAGAACGTAGAATTAC
AGATAATGTGGTGACGACAGTGGCTCATCAGCATCATATCCTGCCAGCTGTCTTCGCCCTTGTTCCAC
GTCACGTAGACCGCTTTCTTCAGCGGGATGTTTTCTTTAACCCAAGAGATATCATCAGAGAACACGTA
GTAGCTCGGGCCAGTAATACGGTTCTCCATTTCCGCGATCGCGTTCTTGTAATACGGCAGCTGGCACA
CGGAACCCGTGTTTGCCCAGTGACGCGGCAGCAGGTAGTCACCACGGCGGATGTGGATAGAAACAGCT
TGGTCGTCAACTTCGATCTGTTTCAGCAGTTCCAGGCTTTCCGGGTTAGCGATGTTCAGGTTAAAAGA
GAAGGCTTTACGAACGTCGTCTTTGATATCGAAGAAGAAGCGTTCAGACTGGTAGAAACCTTTAAAGT
ACAGCAGCGGCCAAAAATAACGTTTTTCGTATGGATACAGAGTAGACGGGTCCTGGCGACGTTCGTAG
ATTTTTTTGAAGAACAGGAACTCCAGGATTTTTTTCAGGGTACGGTTGATGCAAAATTCAGTCTGGCT
CAGGTCAAAGATACGGTTCATCTCATAACCGTTGTGAACTTTATAATGAACCATGTCAGACAGATCGA
TGTTCGTATCCGGGTAATGGTGTTTCATTTTCAGGTAGAACGCGTAGATAAACATCTGGTTACCCAGG
CCGCCGATCATCTTGATCAGACGCATATGTATATCTCCTTCTTGAATTCTAAAAATTGATTGAATGTA
TGCAAATAAATGCATACACCATAGGTGTGGTTTAATTTGATGCCCTTTTTCAGGGCTGGAATGTGTAA
GAGCGGGGTTATTTATGCTGTTGTTTTTTTGTTACTCGGGAAGGGCTTTACCTCTTCCGCATAAACGC
TTCCATCAGCGTTTATAGTTAAAAAAATCTTTCGGAACTGGTTTTGCGCTTACCCCAACCAACAGGGG
ATTTGCTGCTTTCCATTGAGCCTGTTTCTCTGCGCGACGTTCGCGGCGGCGTGTTTGTGCATCCATCT
GGATTCTCCTGTCAGTTAGCTTTGGTGGTGTGTGGCAGTTGTAGTCCTGAACGAAAACCCCCCGCGAT
TGGCACATTGGCAGCTAATCCGGAATCGCACTTACGGCCAATGCTTCGTTTCGTATCACACACCCCAA
AGCCTTCTGCTTTGAATGCTGCCCTTCTTCAGGGCTTAATTTTTAAGAGCGTCACCTTCATGGTGGTC
AGTGCGTCCTGCTGATGTGCTCAGTATCACCGCCAGTGGTATTTATGTCAACACCGCCAGAGATAATT
TATCACCGCAGATGGTTATCTGTATGTTTTTTATATGAATTTATTTTTTGCAGGGGGGCATTGTTTGG
TAGGTGAGAGATCAATTCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGG
CGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGC
TCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAA
AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC
CCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA
CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACC
TGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCG
GTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT
ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTG
GTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTAC
GGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGT
TGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGA
TTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG
AACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTT
AAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT
GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCC
GTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA
CCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTG
GTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCG
CCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGG
TATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAA
AAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATG
GTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGA
GTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATAC
GGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATC
TTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAA
AGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT
TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGT
TCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCT
ATAAAAATAGGCGTATCACGAGGCCCTTTCGTC (SEQ ID NO: 287)

The expression constructs were transformed into a host strain useful for the production of 2'-FL. Biosynthesis of 2'-FL requires the generation of an enhanced cellular pool of both lactose and GDP-fucose (FIG. 2). The wild-type *Eschericia coli* K12 prototrophic strain W3110 was selected as the parent background to test the ability of the candidates to catalyze 2'-FL production (Bachmann, B.J. (1972). Bacteriol Rev 36, 525-557). The particular W3110 derivative employed was one that previously had been modified by the introduction (at the *ampC* locus) of a tryptophan-inducible P_{trpB} *cI+* repressor cassette, generating an *E.coli* strain known as GI724 (LaVallie, E.R. et al. (2000). Methods Enzymol 326, 322-340). Other features of GI724 include *lacIq* and *lacPL8* promoter mutations. *E.coli* strain GI724 affords economical production of recombinant proteins from the phage λ P_L promoter following induction with low levels of exogenous tryptophan (LaVallie, E.R. et al. (1993). Biotechnology (N Y) 11, 187-193; Mieschendahl, et al. (1986). Bio/Technology 4, 802-08). Additional genetic alterations were made to this strain to promote the biosynthesis of 2'-FL. This was achieved in strain GI724 through several manipulations of the chromosome using λ Red recombineering (Court, D.L. et al. (2002). Annu Rev Genet 36, 361-388) and generalized P1 phage transduction.

First, the ability of the *E. coli* host strain to accumulate intracellular lactose was engineered by simultaneous deletion of the endogenous β-galactosidase gene (*lacZ*) and the lactose operon repressor gene (*lacI*). During construction of this deletion, the *lacIq* promoter was placed immediately upstream of the lactose permease gene, *lacY*. The modified strain maintains its ability to transport lactose from the culture medium (via LacY), but is deleted for the wild-type copy of the *lacZ* (β-galactosidase) gene responsible for lactose catabolism. Therefore, an intracellular lactose pool is created when the modified strain is cultured in the presence of exogenous lactose. A schematic of the P_{lacIq} $lacY^+$ chromosomal construct is shown in FIG. 12.

Genomic DNA sequence of the P_{lacIq} $lacY^+$ chromosomal construct is set forth below (SEQ ID NO: 288):

CACCATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAAGTGTAGGCTGGAGC
TGCTTCGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCGGAATAGGAACTAAGGAGGAT
ATTCATATGTACTATTTAAAAAACACAAACTTTTGGATGTTCGGTTTATTCTTTTTCTTTTACTTTTTTATCATG
GGAGCCTACTTCCCGTTTTTCCCGATTTGGCTACATGACATCAACCATATCAGCAAAAGTGATACGGGTATTATT
TTTGCCGCTATTTCTCTGTTCTCGCTATTATTCCAACCGCTGTTTGGTCTGCTTTCTGACAAACTCGGGCTGCGC
AAATACCTGCTGTGGATTATTACCGGCATGTTAGTGATGTTTGCGCCGTTCTTTATTTTTATCTTCGGGCCACTG
TTACAATACAACATTTTAGTAGGATCGATTGTTGGTGGTATTTATCTAGGCTTTTGTTTTAACGCCGGTGCGCCA
GCAGTAGAGGCATTTATTGAGAAAGTCAGCCGTCGCAGTAATTTCGAATTTGGTCGCGCGCGGATGTTTGGCTGT
GTTGGCTGGGCGCTGTGTGCCTCGATTGTCGGCATCATGTTCACCATCAATAATCAGTTTGTTTTCTGGCTGGGC
TCTGGCTGTGCACTCATCCTCGCCGTTTTACTCTTTTTCGCCAAAACGGATGCGCCCTCTTCTGCCACGGTTGCC
AATGCGGTAGGTGCCAACCATTCGGCATTTAGCCTTAAGCTGGCACTGGAACTGTTCAGACAGCCAAAACTGTGG

TTTTTGTCACTGTATGTTATTGGCGTTTCCTGCACCTACGATGTTTTTGACCAACAGTTTGCTAATTTCTTTACT
TCGTTCTTTGCTACCGGTGAACAGGGTACGCGGGTATTTGGCTACGTAACGACAATGGGCGAATTACTTAACGCC
TCGATTATGTTCTTTGCGCCACTGATCATTAATCGCATCGGTGGGAAAAACGCCCTGCTGCTGGCTGGCACTATT
ATGTCTGTACGTATTATTGGCTCATCGTTCGCCACCTCAGCGCTGGAAGTGGTTATTCTGAAAACGCTGCATATG
TTTGAAGTACCGTTCCTGCTGGTGGGCTGCTTTAAATATATTACCAGCCAGTTTGAAGTGCGTTTTTCAGCGACG
ATTTATCTGGTCTGTTTCTGCTTCTTTAAGCAACTGGCGATGATTTTTATGTCTGTACTGGCGGGCAATATGTAT
GAAAGCATCGGTTTCCAGGGCGCTTATCTGGTGCTGGGTCTGGTGGCGCTGGGCTTCACCTTAATTTCCGTGTTC
ACGCTTAGCGGCCCCGGCCCGCTTTCCCTGCTGCGTCGTCAGGTGAATGAAGTCGCTTAAGCAATCAATGTCGGA
TGCGGCGCGAGCGCCTTATCCGACCAACATATCATAACGGAGTGATCGCATTGTAAATTATAAAAATTGCCTGAT
ACGCTGCGCTTATCAGGCCTACAAGTTCAGCGATCTACATTAGCCGCATCCGGCATGAACAAAGCGCAGGAACAA
GCGTCGCA

Second, the ability of the host *E. coli* strain to synthesize colanic acid, an extracellular capsular polysaccharide, was eliminated by the deletion of the *wcaJ* gene, encoding the UDP-glucose lipid carrier transferase (Stevenson, G. et al. (1996). J Bacteriol 178, 4885-893). In a *wcaJ* null background GDP-fucose accumulates in the *E. coli* cytoplasm (Dumon, C. et al. (2001). Glycoconj J 18, 465-474). A schematic of the chromosomal deletion of *wcaJ* is shown in FIG. 13.

The sequence of the chromosomal region of *E. coli* bearing the *ΔwcaJ*::FRT mutation is set forth below (SEQ ID NO: 289):

GTTCGGTTATATCAATGTCAAAAACCTCACGCCGCTCAAGCTGGTGATCAACTCCGGGAACGGCGCAGCGGGTCC
GGTGGTGGACGCCATTGAAGCCCGCTTTAAAGCCCTCGGCGCGCCCGTGGAATTAATCAAAGTGCACAACACGCC
GGACGGCAATTTCCCCAACGGTATTCCTAACCCACTACTGCCGGAATGCCGCGACGACACCCGCAATGCGGTCAT
CAAACACGGCGCGGATATGGGCATTGCTTTTGATGGCGATTTTGACCGCTGTTTCCTGTTTGACGAAAAAGGGCA
GTTTATTGAGGGCTACTACATTGTCGGCCTGTTGGCAGAAGCATTCCTCGAAAAAAATCCCGGCGCGAAGATCAT
CCACGATCCACGTCTCTCCTGGAACACCGTTGATGTGGTGACTGCCGCAGGTGGCACGCCGGTAATGTCGAAAAC
CGGACACGCCTTTATTAAAGAACGTATGCGCAAGGAAGACGCCATCTATGGTGGCGAAATGAGCGCCCACCATTA
CTTCCGTGATTTCGCTTACTGCGACAGCGGCATGATCCCGTGGCTGCTGGTCGCCGAACTGGTGTGCCTGAAAGA
TAAAACGCTGGGCGAACTGGTACGCGACCGGATGGCGGCGTTTCCGGCAAGCGGTGAGATCAACAGCAAACTGGC
GCAACCCGTTGAGGCGATTAACCGCGTGGAACAGCATTTTAGCCGTGAGGCGCTGGCGGTGGATCGCACCGATGG
CATCAGCATGACCTTTGCCGACTGGCGCTTTAACCTGCGCACCTCCAATACCGAACCGGTGGTGCGCCTGAATGT
GGAATCGCGCGGTGATGTGCCGCTGATGGAAGCGCGAACGCGAACTCTGCTGACGTTGCTGAACGAGTAATGTCG
GATCTTCCCTTACCCCACTGCGGGTAAGGGGCTAATAACAGGAACAACGATGATTCCGGGGATCCGTCGACCTGC
AGTTCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCGAAGCAGCTCCAGCCTACAGTTAACAAAGCGGCATA
TTGATATGAGCTTACGTGAAAAAACCATCAGCGGCGCGAAGTGGTCGGCGATTGCCACGGTGATCATCATCGGCC
TCGGGCTGGTGCAGATGACCGTGCTGGCGCGGATTATCGACAACCACCAGTTCGGCCTGCTTACCGTGTCGCTGG
TGATTATCGCGCTGGCAGATACGCTTTCTGACTTCGGTATCGCTAACTCGATTATTCAGCGAAAAGAAATCAGTC
ACCTTGAACTCACCACGTTGTACTGGCTGAACGTCGGGCTGGGGATCGTGGTGTGCGTGGCGGTGTTTTTGTTGA
GTGATCTCATCGGCGACGTGCTGAATAACCCGGACCTGGCACCGTTGATTAAAACATTATCGCTGGCGTTTGTGG
TAATCCCCCACGGGCAACAGTTCCGCGCGTTGATGCAAAAAGAGCTGGAGTTCAACAAAATCGGCATGATCGAAA
CCAGCGCGGTGCTGGCGGGCTTCACTTGTACGGTGGTTAGCGCCCATTTCTGGCCGCTGGCGATGACCGCGATCC
TCGGTTATCTGGTCAATAGTGCGGTGAGAACGCTGCTGTTTGGCTACTTTGGCCGCAAAATTTATCGCCCCGGTC
TGCATTTCTCGCTGGCGTCGGTGGCACCGAACTTACGCTTTGGTGCCTGGCTGACGGCGGACAGCATCATCAACT
ATCTCAATACCAACCTTTCAACGCTCGTGCTGGCGCGTATTCTCGGCGCGGGCGTGGCAGGGGGATACAACCTGG
CGTACAACGTGGCCGTTGTGCCACCGATGAAGCTGAACCCAATCATCACCCGCGTGTTGTTTCCGGCATTCGCCA
AAATTCAGGACGATACCGAAAAGCTGCGTGTTAACTTCTACAAGCTGCTGTCGGTAGTGGGGATTATCAACTTTC
CGGCGCTGCTCGGGCTAATGGTGGTGTCGAATAACTTTGTACCGCTGGTCTTTGGTGAGAAGTGGAACAGCATTA
TTCCGGTGCTGCAATTGCTGTGTGTGGTGGGTCTGCTGCGCTCCG

Third, the magnitude of the cytoplasmic GDP-fucose pool was enhanced by the introduction of a null mutation into the *lon* gene. Lon is an ATP-dependant intracellular

protease that is responsible for degrading RcsA, which is a positive transcriptional regulator of colanic acid biosynthesis in *E. coli* (Gottesman, S. & Stout, V. Mol Microbiol 5, 1599-1606 (1991)). In a *lon* null background, RcsA is stabilized, RcsA levels increase, the genes responsible for GDP-fucose synthesis in *E. coli* are up-regulated, and intracellular GDP-fucose concentrations are enhanced. The lon gene was almost entirely deleted and replaced by an inserted functional, wild-type, but promoter-less *E.coli lacZ*$^+$ gene (*Δlon::(kan, lacZ*$^+$). λ Red recombineering was used to perform the construction. A schematic of the *kan, lacZ*$^+$ insertion into the *lon* locus is shown in FIG. 14.

Genomic DNA sequence surrounding the lacZ+ insertion into the *lon* region in the *E.coli* strain is set forth below (SEQ ID NO: 290):

GTGGATGGAAGAGGTGGAAAAAGTGGTTATGGAGGAGTGGGTAATTGATGGTGAAAGGAAAGGGTTGGTGATTTA
TGGGAAGGGGGAAGGGGAAGAGGGATGTGGTGAATAATTAAGGATTGGGATAGAATTAGTTAAGGAAAAAGGGGG
GATTTTATGTGGGGTTTAATTTTTGGTGTATTGTGGGGGTTGAATGTGGGGGAAAGATGGGGATATAGTGAGGTA
GATGTTAATAGATGGGGTGAAGGAGAGTGGTGTGATGTGATTAGGTGGGGGAAATTAAAGTAAGAGAGAGGTGTA
TGATTGGGGGGATGGGTGGAGGTGGAGTTGGAAGTTGGTATTGTGTAGAAAGTATAGGAAGTTGAGAGGGGTTTT
GAAGGTGAGGGTGGGGGAAGGAGTGAGGGGGGAAGGGGTGGTAAAGGAAGGGGAAGAGGTAGAAAGGGAGTGGGG
AGAAAGGGTGGTGAGGGGGGATGAATGTGAGGTAGTGGGGTATGTGGAGAAGGGAAAAGGGAAGGGGAAAGAGAA
AGGAGGTAGGTTGGAGTGGGGTTAGATGGGGATAGGTAGAGTGGGGGGTTTTATGGAGAGGAAGGGAAGGGGAAT
TGGGAGGTGGGGGGGGGTGTGGTAAGGTTGGGAAGGGGTGGAAAGTAAAGTGGATGGGTTTGTTGGGGGGAAGGA
TGTGATGGGGGAGGGGATGAAGATGTGATGAAGAGAGAGGATGAGGATGGTTTGGGATGATTGAAGAAGATGGAT
TGGAGGGAGGTTGTGGGGGGGGGTTGGGTGGAGAGGGTATTGGGGTATGAGTGGGGAGAAGAGAGAATGGGGTGGT
GTGATGGGGGGGTGTTGGGGGTGTGAGGGGAGGGGGGGGGGGTTGTTTTTGTGAAGAGGGAGGTGTGGGGTGGGG
TGAATGAAGTGGAGGAGGAGGGAGGGGGGGTATGGTGGGTGGGGAGGAGGGGGGTTGGTTGGGGAGGTGTGGTGG
AGGTTGTGAGTGAAGGGGGAAGGGAGTGGGTGGTATTGGGGGAAGTGGGGGGGGGAGGATGTGGTGTGATGTGAGG
TTGGTGGTGGGGAGAAAGTATGGATGATGGGTGATGGAATGGGGGGGGTGGATAGGGTTGATGGGGGTAGGTGGG
GATTGGAGGAGGAAGGGAAAGATGGGATGGAGGGAGGAGGTAGTGGGATGGAAGGGGGTGTTGTGGATGAGGATG
ATGTGGAGGAAGAGGATGAGGGGGTGGGGGGAGGGGAAGTGTTGGGGAGGGTGAAGGGGGGATGGGGGAGGGGGA
GGATGTGGTGGTGAGGGATGGGGATGGGTGGTTGGGGAATATGATGGTGGAAAATGGGGGGTTTTGTGGATTGAT
GGAGTGTGGGGGGGTGGGTGTGGGGGAGGGGTATGAGGAGATAGGGTTGGGTAGGGGTGATATTGGTGAAGAGGT
TGGGGGGGAATGGGGTGAGGGGTTGGTGGTGGTTTAGGGTATGGGGGGTGGGGATTGGGAGGGGATGGGGTTGTA
TGGGGTTGTTGAGGAGTTGTTGTAATAAGGGGATGTTGAAGTTGGTATTGGGAAGTTGGTATTGTGTAGAAAGTA
TAGGAAGTTGGAAGGAGGTGGAGGGTAGATAAAGGGGGGGGTTATTTTTGAGAGGAGAGGAAGTGGTAATGGTAG
GGAGGGGGGGTGAGGTGGAATTGGGGGGATAGTGAGGGGGTGGAGGAGTGGTGGGGAGGAATGGGGATATGGAAA
GGGTGGATATTGAGGGATGTGGGTTGTTGGGGGTGGAGGAGATGGGGATGGGTGGTTTGGATGAGTTGGTGTTGA
GTGTAGGGGGTGATGTTGAAGTGGAAGTGGGGGGGGGAGTGGTGTGGGGGATAATTGAATTGGGGGGTGGGGGAG
GGGAGAGGGTTTTGGGTGGGGAAGAGGTAGGGGGTATAGATGTTGAGAATGGGAGATGGGAGGGGTGAAAAGAGG
GGGGAGTAAGGGGGTGGGGATAGTTTTGTTGGGGGGGTAATGGGAGGGAGTTTAGGGGGTGTGGTAGGTGGGGGA
GGTGGGAGTTGAGGGGAATGGGGGGGGGATGGGGTGTATGGGTGGGGAGTTGAAGATGAAGGGTAATGGGGATTT
GAGGAGTAGGATGAATGGGGTAGGTTTTGGGGGTGATAAATAAGGTTTTGGGGTGATGGTGGGAGGGGTGAGGGG
TGGTAATGAGGAGGGGATGAGGAAGTGTATGTGGGGTGGAGTGGAAGAAGGGTGGTTGGGGGTGGTAATGGGGGG
GGGGGTTGGAGGGTTGGAGGGAGGGGTTAGGGTGAATGGGGGTGGGTTGAGTTAGGGGAATGTGGTTATGGAGGG
GTGGAGGGGTGAAGTGATGGGGGAGGGGGGTGAGGAGTTGTTTTTTATGGGGAATGGAGATGTGTGAAAGAAAGG
GTGAGTGGGGGTTAAATTGGGAAGGGTTATTAGGGAGGTGGATGGAAAAATGGATTTGGGTGGTGGTGAGATGGG
GGATGGGGTGGGAGGGGGGGGGGAGGGTGAGAGTGAGGTTTTGGGGGAGAGGGGAGTGGTGGGAGGGGGTGATGT
GGGGGGGTTGTGAGGATGGGGTGGGGTTGGGTTGGAGTAGGGGTAGTGTGAGGGAGAGTTGGGGGGGGGTGTGGG
GGTGGGGTAGTTGAGGGAGTTGAATGAAGTGTTTAGGTTGTGGAGGGAGATGGAGAGGGAGTTGAGGGGTTGGGA
GGGGGTTAGGATGGAGGGGGAGGATGGAGTGGAGGAGGTGGTTATGGGTATGAGGGAAGAGGTATTGGGTGGTGA
GTTGGATGGTTTGGGGGGATAAAGGGAAGTGGAAAAAGTGGTGGTGGTGTTTTGGTTGGGTGAGGGGTGGATGGG
GGGTGGGGTGGGGAAAGAGGAGAGGGTTGATAGAGAAGTGGGGATGGTTGGGGGTATGGGGAAAATGAGGGGGGT
AAGGGGAGGAGGGGTTGGGGTTTTGATGATATTTAATGAGGGAGTGATGGAGGGAGTGGGAGAGGAAGGGGGGGT
GTAAAGGGGGATAGTGAGGAAAGGGGTGGGAGTATTTAGGGAAAGGGGGAAGAGTGTTAGGGATGGGGTGGGGGT
ATTGGGAAAGGATGAGGGGGGGGGTGTGTGGAGGTAGGGAAAGGGATTTTTTGATGGAGGATTTGGGGAGAGGGG
GGAAGGGGTGGTGTTGATGGAGGGGGGGGTAGATGGGGGAAATAATATGGGTGGGGGTGGTGTGGGGTGGGGGGG

GTTGATAGTGGAGGGGGGGGGAAGGATGGAGAGATTTGATGGAGGGATAGAGGGGGTGGTGATTAGGGGGGTGGG
GTGATTGATTGGGGAGGGAGGAGATGATGAGAGTGGGGTGATTAGGATGGGGGTGGAGGATTGGGGTTAGGGGTT
GGGTGATGGGGGGTAGGGAGGGGGGATGATGGGTGAGAGGATTGATTGGGAGGATGGGGTGGGTTTGAATATTGG
GTTGATGGAGGAGATAGAGGGGGTAGGGGTGGGAGAGGGTGTAGGAGAGGGGATGGTTGGGATAATGGGAAGAGG
GGAGGGGGTTAAAGTTGTTGTGGTTGATGAGGAGGATATGGTGGAGGATGGTGTGGTGATGGATGAGGTGAGGAT
GGAGAGGATGATGGTGGTGAGGGTTAAGGGGTGGAATGAGGAAGGGGTTGGGGTTGAGGAGGAGGAGAGGATTTT
GAATGGGGAGGTGGGGGAAAGGGAGATGGGAGGGTTGTGGTTGAATGAGGGTGGGGTGGGGGGTGTGGAGTTGAA
GGAGGGGAGGATAGAGATTGGGGATTTGGGGGGTGGAGAGTTTGGGGTTTTGGAGGTTGAGAGGTAGTGTGAGGG
GATGGGGATAAGGAGGAGGGTGATGGATAATTTGAGGGGGGAAAGGGGGGGTGGGGGTGGGGAGGTGGGTTTGAG
GGTGGGATAAAGAAAGTGTTAGGGGTAGGTAGTGAGGGAAGTGGGGGGAGATGTGAAGTTGAGGGTGGAGTAGAG
GGGGGGTGAAATGATGATTAAAGGGAGTGGGAAGATGGAAATGGGTGATTTGTGTAGTGGGTTTATGGAGGAAGG
AGAGGTGAGGGAAAATGGGGGTGATGGGGGAGATATGGTGATGTTGGAGATAAGTGGGGTGAGTGGAGGGGAGGA
GGATGAGGGGGAGGGGGTTTTGTGGGGGGGGTAAAAATGGGGTGAGGTGAAATTGAGAGGGGAAAGGAGTGTGGT
GGGGGTAAGGGAGGGAGGGGGGGTTGGAGGAGAGATGAAAGGGGGAGTTAAGGGGATGAAAAATAATTGGGGTGT
GGGGTTGGTGTAGGGAGGTTTGATGAAGATTAAATGTGAGGGAGTAAGAAGGGGTGGGATTGTGGGTGGGAAGAA
AGGGGGGATTGAGGGTAATGGGATAGGTGAGGTTGGTGTAGATGGGGGGATGGTAAGGGTGGATGTGGGAGTTTG
AGGGGAGGAGGAGAGTATGGGGGTGAGGAAGATGGGAGGGAGGGAGGTTTGGGGGAGGGGTTGTGGTGGGGGAAA
GGAGGGAAAGGGGGATTGGGGATTGAGGGTGGGGAAGTGTTGGGAAGGGGGATGGGTGGGGGGGTGTTGGGTATT
AGGGGAGGTGGGGAAAGGGGGATGTGGTGGAAGGGGATTAAGTTGGGTAAGGGGAGGGTTTTGGGAGTGAGGAGG
TTGTAAAAGGAGGGGGAGTGAATGGGTAATGATGGTGATAGTAGGTTTGGTGAGGTTGTGAGTGGAAAATAGTGA
GGTGGGGGAAAATGGAGTAATAAAAAGAGGGGTGGGAGGGTAATTGGGGGTTGGGAGGGTTTTTTTGTGTGGGTA
AGTTAGATGGGGGATGGGGGTTGGGGTTATTAAGGGGTGTTGTAAGGGGATGGGTGGGGTGATATAAGTGGTGGG
GGTTGGTAGGTTGAAGGATTGAAGTGGGATATAAATTATAAAGAGGAAGAGAAGAGTGAATAAATGTGAATTGAT
GGAGAAGATTGGTGGAGGGGGTGATATGTGTAAAGGTGGGGGTGGGGGTGGGTTAGATGGTATTATTGGTTGGGT
AAGTGAATGTGTGAAAGAAGG

Fourth, a *thyA* (thymidylate synthase) mutation was introduced into the strain by P1 transduction. In the absence of exogenous thymidine, *thyA* strains are unable to make DNA and die. The defect can be complemented in *trans* by supplying a wild-type *thyA* gene on a multicopy plasmid (Belfort, M., Maley, G.F., and Maley, F. (1983). Proc Natl Acad Sci U S A 80, 1858-861). This complementation was used here as a means of plasmid maintenance.

An additional modification that is useful for increasing the cytoplasmic pool of free lactose (and hence the final yield of 2'-FL) is the incorporation of a *lacA* mutation. LacA is a lactose acetyltransferase that is only active when high levels of lactose accumulate in the *E. coli* cytoplasm. High intracellular osmolarity (*e.g.*, caused by a high intracellular lactose pool) can inhibit bacterial growth, and *E. coli* has evolved a mechanism for protecting itself from high intra cellular osmlarity caused by lactose by "tagging" excess intracellular lactose with an acetyl group using LacA, and then actively expelling the acetyl-lactose from the cell (Danchin, A. Bioessays 31, 769-773 (2009)). Production of acetyl-lactose in *E. coli* engineered to produce 2'-FL or other human milk oligosaccharides is therefore undesirable: it reduces overall yield. Moreover, acetyl-lactose is a side product that complicates oligosaccharide purification schemes. The incorporation of a *lacA* mutation resolves these problems. Sub-optimal production of fucosylated oligosaccharides occurs in strains lacking either or both of the mutations in the colanic acid pathway and the lon protease. Diversion of lactose into a side product (acetyl-lactose) occurs in strains that do not contain the lacA

mutation. A schematic of the *lacA* deletion and corresponding genomic sequence is provided above (SEQ ID NO: 288).

The strain used to test the different α(1,2) FT candidates incorporates all the above genetic modifications and has the following genotype:

Δ*ampC*::P_{trp}^{B}*cI*, Δ(*lacI-lacZ*)::*FRT*, P_{lacIq}*lacY*$^{+}$, Δ*wcaJ*::FRT, *thyA*::Tn*10*, Δ*lon*:(*npt3*, *lacZ*$^{+}$), Δ*lacA*

The *E. coli* strains harboring the different α(1,2) FT candidate expression plasmids were analyzed. Strains were grown in selective media (lacking thymidine) to early exponential phase. Lactose was then added to a final concentration of 0.5%, and tryptophan (200 μM) was added to induce expression of each candidate α(1,2) FT from the P_L promoter. At the end of the induction period (~24 h) equivalent OD 600 units of each strain and the culture supernatant was harvested. Lysates were prepared and analyzed for the presence of 2'-FL by thin layer chromatography (TLC).

A map of plasmid pG217 is shown in FIG. 11, which carries the *B. vulgatus* FutN. The sequence of plasmid pG217 is set forth below (SEQ ID NO: 291):

TCTAGAATTCTAAAAATTGATTGAATGTATGCAAATAAATGCATACACCATAGGTGTGGTTTAATTTGATGCCCT
TTTTCAGGGCTGGAATGTGTAAGAGCGGGGTTATTTATGCTGTTGTTTTTTTGTTACTCGGGAAGGGCTTTACCT
CTTCCGCATAAACGCTTCCATCAGCGTTTATAGTTAAAAAAATCTTTCGGAACTGGTTTTGCGCTTACCCCAACC
AACAGGGGATTTGCTGCTTTCCATTGAGCCTGTTTCTCTGCGCGACGTTCGCGGCGGCGTGTTTGTGCATCCATC
TGGATTCTCCTGTCAGTTAGCTTTGGTGGTGTGTGGCAGTTGTAGTCCTGAACGAAAACCCCCCGCGATTGGCAC
ATTGGCAGCTAATCCGGAATCGCACTTACGGCCAATGCTTCGTTTCGTATCACACACCCCAAAGCCTTCTGCTTT
GAATGCTGCCCTTCTTCAGGGCTTAATTTTTAAGAGCGTCACCTTCATGGTGGTCAGTGCGTCCTGCTGATGTGC
TCAGTATCACCGCCAGTGGTATTTATGTCAACACCGCCAGAGATAATTTATCACCGCAGATGGTTATCTGTATGT
TTTTTATATGAATTTATTTTTTGCAGGGGGGCATTGTTTGGTAGGTGAGAGATCAATTCTGCATTAATGAATCGG
CCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGT
CGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC
AGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA
TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA
AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCT
GTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGT
CGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG
TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG
GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTAT
CTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGG
TAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT
TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGAT
CTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGA
CAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACT
CCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC
ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAAC
TTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCG
CAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTC
CCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT
TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCC
ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAG
TTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAA
ACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC
CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAA
AAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCA
GGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT

TCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCAC
GAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCAC
AGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGG
CTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATATGCGGTGTGAAATACCGCACAGA
TGCGTAAGGAGAAAATACCGCATCAGGCGCCTCCTCAACCTGTATATTCGTAAACCACGCCCAATGGGAGCTGTC
TCAGGTTTGTTCCTGATTGGTTACGGCGCGTTTCGCATCATTGTTGAGTTTTTCCGCCAGCCCGACGCGCAGTTT
ACCGGTGCCTGGGTGCAGTACATCAGCATGGGGCAAATTCTTTCCATCCCGATGATTGTCGCGGGTGTGATCATG
ATGGTCTGGGCATATCGTCGCAGCCCACAGCAACACGTTTCCTGAGGAACCATGAAACAGTATTTAGAACTGATG
CAAAAAGTGCTCGACGAAGGCACACAGAAAAACGACCGTACCGGAACCGGAACGCTTTCCATTTTTGGTCATCAG
ATGCGTTTTAACCTGCAAGATGGATTCCCGCTGGTGACAACTAAACGTTGCCACCTGCGTTCCATCATCCATGAA
CTGCTGTGGTTTCTGCAGGGCGACACTAACATTGCTTATCTACACGAAAACAATGTCACCATCTGGGACGAATGG
GCCGATGAAAACGGCGACCTCGGGCCAGTGTATGGTAAACAGTGGCGCGCCTGGCCAACGCCAGATGGTCGTCAT
ATTGACCAGATCACTACGGTACTGAACCAGCTGAAAAACGACCCGGATTCGCGCCGCATTATTGTTTCAGCGTGG
AACGTAGGCGAACTGGATAAAATGGCGCTGGCACCGTGCCATGCATTCTTCCAGTTCTATGTGGCAGACGGCAAA
CTCTCTTGCCAGCTTTATCAGCGCTCCTGTGACGTCTTCCTCGGCCTGCCGTTCAACATTGCCAGCTACGCGTTA
TTGGTGCATATGATGGCGCAGCAGTGCGATCTGGAAGTGGGTGATTTTGTCTGGACCGGTGGCGACACGCATCTG
TACAGCAACCATATGGATCAAACTCATCTGCAATTAAGCCGCGAACCGCGTCCGCTGCCGAAGTTGATTATCAAA
CGTAAACCCGAATCCATCTTCGACTACCGTTTCGAAGACTTTGAGATTGAAGGCTACGATCCGCATCCGGGCATT
AAAGCGCCGGTGGCTATCTAATTACGAAACATCCTGCCAGAGCCGACGCCAGTGTGCGTCGGTTTTTTTACCCTC
CGTTAAATTCTTCGAGACGCCTTCCCGAAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATC
GGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCC
AGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGCCAAGCTTTCTTTAATGAAGCAGGGCATCAGGAC
GGTATCTTTGTGGAGAAAGCAGAGTAATCTTATTCAGCCTGACTGGTGGGAAACCACCAGTCAGAATGTGTTAGC
GCATGTTGACAAAAATACCATTAGTCACATTATCCGTCAGTCGGACGACATGGTAGATAACCTGTTTATTATGCG
TTTTGATCTTACGTTTAATATTACCTTTATGCGATGAAACGGTCTTGGCTTTGATATTCATTTGGTCAGAGATTT
GAATGGTTCCCTGACCTGCCATCCACATTCGCAACATACTCGATTCGGTTCGGCTCAATGATAACGTCGGCATAT
TTAAAAACGAGGTTATCGTTGTCTCTTTTTTCAGAATATCGCCAAGGATATCGTCGAGAGATTCCGGTTTAATCG
ATTTAGAACTGATCAATAAATTTTTTCTGACCAATAGATATTCATCAAAATGAACATTGGCAATTGCCATAAAAA
CGATAAATAACGTATTGGGATGTTGATTAATGATGAGCTTGATACGCTGACTGTTAGAAGCATCGTGGATGAAAC
AGTCCTCATTAATAAACACCACTGAAGGGCGCTGTGAATCACAAGCTATGGCAAGGTCATCAACGGTTTCAATGT
CGTTGATTTCTCTTTTTTTAACCCCTCTACTCAACAGATACCCGGTTAAACCTAGTCGGGTGTAACTACATAAAT
CCATAATAATCGTTGACATGGCATACCCTCACTCAATGCGTAACGATAATTCCCCTTACCTGAATATTTCATCAT
GACTAAACGGAACAACATGGGTCACCTAATGCGCCACTCTCGCGATTTTTCAGGCGGACTTACTATCCCGTAAAG
TGTTGTATAATTTGCCTGGAATTGTCTTAAAGTAAAGTAAATGTTGCGATATGTGAGTGAGCTTAAAACAAATAT
TTCGCTGCAGGAGTATCCTGGAAGATGTTCGTAGAAGCTTACTGCTCACAAGAAAAAAGGCACGTCATCTGACGT
GCCTTTTTTATTTGTACTACCCTGTACGATTACTGCAGCTCGAGTTAGGATACCGGCACTTTGATCCAACCAGTC
GGGTAGATATCCGGTGCTTCGGAGTGCTGGAACCAACGGCTCGGCACAATAACAGTCTTATCCATATTAGGGTTC
AGCCAGGCACCCCACCAAGAAAACGTGCTGTTACAAATGATGTGATGTTTGCAATGAGACATCAGCATCATATCC
TGCCAGGAGTCTTCATCAGTGTTCCAGTCAATATAAACCGCATTCTGCAGTGGCAGATTTTCTTTAACCCACGCG
ATATCGTCGGAGAAGATATAGTAAGATGGGCTAGCAACACGACGGGACATTTCCGCGATAGCATTCTGGTAATAC
GGCAGCTGGCACACGGAACCGGTAGTAGCCCAGTGTTTCGGCTGCAGATAGTCACCACGACGAATGTGCAGGGAA
ACCGCGTTTTCATCTTTGTCCAGGATTTCCAGCATGTTCAGGCTGCGGGAATTTGCTTTGTTCTTATCAAAGGTG
AAGGATTCACGCACTTCGTCTTTGATATCAGCGAAGAAACGCTCGCTCTGATAGAAACCTTTAAAGTACAGCAGC
GGCCAGAAATACTTCTTCTCGAACGCACGCAGAGAGTTCGGCGCCTGCTTGCGTTCGTAGATTTTTTTAAAAAAC
AGGAATTCGATAACTTTTTTCAGCGGTTGGTTGATGCAGAATTCGGTGTGCGGCAGGTTGAACACGCGGTGCATT
TCGTAACCGTAATGGACTTTGTAATGCATCATGTCGCTCAGGTCGATACGGACCTTCGGGTAATACTTTTTCATA
CGCAGATAGAAAGCATAGATAAACATCTGGTTGCCCAGACCGCCAGTCACTTTGATCAGACGCATTATATCTCCT
TCTTG

Fucosylated oligosaccharides produced by metabolically engineered *E. coli* cells are purified from culture broth post-fermentation. An exemplary procedure comprises five steps. (1) Clarification: Fermentation broth is harvested and cells removed by sedimentation in a preparative centrifuge at 6000 x g for 30 min. Each bioreactor run yields about 5-7 L of partially clarified supernatant. (2) Product capture on coarse carbon : A column packed with coarse carbon (Calgon 12x40 TR) of ~1000 ml volume (dimension 5 cm diameter x 60 cm

length) is equilibrated with 1 column volume (CV) of water and loaded with clarified culture supernatant at a flow rate of 40 ml/min. This column has a total capacity of about 120 g of sugar. Following loading and sugar capture, the column is washed with 1.5 CV of water, then eluted with 2.5 CV of 50% ethanol or 25% isopropanol (lower concentrations of ethanol at this step (25-30%) may be sufficient for product elution.) This solvent elution step releases about 95% of the total bound sugars on the column and a small portion of the color bodies. In this first step capture of the maximal amount of sugar is the primary objective. Resolution of contaminants is not an objective. (3) Evaporation: A volume of 2.5 L of ethanol or isopropanol eluate from the capture column is rotary-evaporated at 56 C° and a sugar syrup in water is generated. Alternative methods that could be used for this step include lyophilization or spray-drying. (4) Flash chromatography on fine carbon and ion exchange media: A column (GE Healthcare HiScale50/40, 5x40cm, max pressure 20 bar) connected to a Biotage Isolera One FLASH Chromatography System is packed with 750 ml of a Darco Activated Carbon G60 (100-mesh): Celite 535 (coarse) 1:1 mixture (both column packings were obtained from Sigma). The column is equilibrated with 5 CV of water and loaded with sugar from step 3 (10-50 g, depending on the ratio of 2'-FL to contaminating lactose), using either a celite loading cartridge or direct injection. The column is connected to an evaporative light scattering (ELSD) detector to detect peaks of eluting sugars during the chromatography. A four-step gradient of isopropanol, ethanol or methanol is run in order to separate 2'-FL from monosaccharides (if present), lactose and color bodies. Fractions corresponding to sugar peaks are collected automatically in 120-ml bottles, pooled and directed to step 5. In certain purification runs from longer-than-normal fermentations, passage of the 2'-FL-containing fraction through anion-exchange and cation exchange columns can remove excess protein/DNA/caramel body contaminants. Resins tested successfully for this purpose are Dowex 22.

The gene screening approach described herein was successfully utilized to identify new α(1,2) FTs for the efficient biosynthesis of 2'-FL in metabolically engineered *E. coli* host strains. The results of the screen are summarized in Table 1.

Production Host Strains

E. coli K-12 is a well-studied bacterium which has been the subject of extensive research in microbial physiology and genetics and commercially exploited for a variety of industrial uses. The natural habitat of the parent species, *E. coli*, is the large bowel of mammals. *E. coli* K-12 has a history of safe use, and its derivatives are used in a large

number of industrial applications, including the production of chemicals and drugs for human administration and consumption. *E. coli* K-12 was originally isolated from a convalescent diphtheria patient in 1922. Because it lacks virulence characteristics, grows readily on common laboratory media, and has been used extensively for microbial physiology and genetics research, it has become the standard bacteriological strain used in microbiological research, teaching, and production of products for industry and medicine. *E. coli* K-12 is now considered an enfeebled organism as a result of being maintained in the laboratory environment for over 70 years. As a result, K-12 strains are unable to colonize the intestines of humans and other animals under normal conditions. Additional information on this well known strain is available at http://epa.gov/oppt/biotech/pubs/fra/fra004.htm. In addition to *E. coli* K12, other bacterial strains are used as production host strains, e.g., a variety of bacterial species may be used in the oligosaccharide biosynthesis methods, e.g., *Erwinia herbicola* (*Pantoea agglomerans*), *Citrobacter freundii*, *Pantoea citrea*, *Pectobacterium carotovorum*, *or Xanthomonas campestris.* Bacteria of the genus *Bacillus* may also be used, including *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus coagulans*, *Bacillus thermophilus*, *Bacillus laterosporus*, *Bacillus megaterium*, *Bacillus mycoides*, *Bacillus pumilus*, *Bacillus lentus*, *Bacillus cereus*, *and Bacillus circulans*. Similarly, bacteria of the genera *Lactobacillus* and *Lactococcus* may be modified using the methods of this invention, including but not limited to *Lactobacillus acidophilus, Lactobacillus salivarius*, *Lactobacillus plantarum*, *Lactobacillus helveticus*, *Lactobacillus delbrueckii*, *Lactobacillus rhamnosus*, *Lactobacillus bulgaricus*, *Lactobacillus crispatus*, *Lactobacillus gasseri*, *Lactobacillus casei*, *Lactobacillus reuteri*, *Lactobacillus jensenii*, *and Lactococcus lactis. Streptococcus thermophiles and Proprionibacterium freudenreichii* are also suitable bacterial species for the invention described herein. Also included as part of this invention are strains, modified as described here, from the genera *Enterococcus* (e.g., *Enterococcus faecium* and *Enterococcus thermophiles*), *Bifidobacterium* (e.g., *Bifidobacterium longum*, *Bifidobacterium infantis*, and *Bifidobacterium bifidum*), *Sporolactobacillus* spp., *Micromomospora* spp., *Micrococcus* spp., *Rhodococcus* spp., and *Pseudomonas* (e.g., *Pseudomonas fluorescens* and *Pseudomonas aeruginosa*).

Suitable host strains are amenable to genetic manipulation, e.g., they maintain expression constructs, accumulate precursors of the desired end product, e.g., they maintain pools of lactose and GDP-fucose, and accumulate endproduct, e.g., 2'-FL. Such strains grow well on defined minimal media that contains simple salts and generally a single carbon source. The strains engineered as described above to produce the desired fucosylated

oligosaccharide(s) are grown in a minimal media. An exemplary minimal medium used in a bioreactor, minimal "FERM" medium, is detailed below.

Ferm (10 liters): Minimal medium comprising:

40g $(NH_4)_2HPO_4$

100g KH_2PO_4

10g $MgSO_4.7H_2O$

40g NaOH

1X Trace elements:

1.3g NTA (nitrilotriacetic acid)

0. 5g $FeSO_4.7H_2O$

0.09g $MnCl_2.4H_2O$

0.09g $ZnSO_4.7H_2O$

0.01g $CoCl_2.6H_2O$

0.01g $CuCl_2.2H_2O$

0.02g H_3BO_3

0.01g $Na_2MoO_4.2H_2O$ (pH 6.8)

Water to 10 liters

DF204 antifoam (0.1ml/L)

150 g glycerol (initial batch growth), followed by fed batch mode with a 90% glycerol-1% $MgSO_4$-1X trace elements feed, at various rates for various times.

A suitable production host strain is one that is not the same bacterial strain as the source bacterial strain from which the fucosyltransferase-encoding nucleic acid sequence was identified.

Bacteria comprising the characteristics described herein are cultured in the presence of lactose, and a fucosylated oligosaccharide is retrieved, either from the bacterium itself or from a culture supernatant of the bacterium. The fucosylated oligosaccharide is purified for use in therapeutic or nutritional products, or the bacteria are used directly in such products.

EXAMPLES

Example 1: Identification of novel α(1,2) fucosyltransferases

To identify additional novel α(1,2)fucosyltransferases, a multiple sequence alignment query was generated using the alignment algorithm of the CLCbio Main Workbench package, version 6.9 (CLCbio, 10 Rogers Street #101, Cambridge, Massachusetts 02142, USA) using four previously identified lactose-utilizing α(1,2)fucosyltransferase protein sequences: *H. pylori* futC (SEQ ID NO: 1), *H. mustelae* FutL (SEQ ID NO: 2), *Bacteroides vulgatus* futN (SEQ ID NO: 3), and *E. coli 0126* wbgL (SEQ ID NO: 4). This sequence alignment and percentages of sequence identity between the four previously identified lactose-utilizing α(1,2)fucosyltransferase protein sequences is shown in FIG. 3. An iterative PSI-BLAST was performed, using the FASTA-formatted multiple sequence alignment as the query, and the NCBI PSI-BLAST program run on a local copy of NCBI BLAST+ version 2.2.29. An initial position-specific scoring matrix file (.pssm) was generated by PSI-BLAST, which was then used to adjust the score of iterative homology search runs. The process is iterated to generate an even larger group of candidates, and the results of each run were used to further refine the matrix.

A portion of the initial position-specific scoring matrix file used is shown below:

Last position-specific scoring matrix computed

		A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
1	M	-1	-1	-2	-3	-2	0	-2	-3	-2	1	2	-1	6	0	-3	-2	-1	-2	-1	1
2	A	2	-2	0	4	-2	-1	1	-1	-1	-2	-3	-1	-2	-3	-1	1	-1	-3	-3	-1
3	F	-2	-3	-3	-4	-3	-3	-3	-3	-1	0	0	-3	0	7	-4	-3	-2	1	3	-1
4	K	0	3	0	-1	-2	1	0	-1	-1	-3	-3	3	-2	-3	-1	2	0	-3	-2	-2
5	V	-1	-3	-3	-4	-1	-3	-3	-4	-3	4	2	-3	1	0	-3	-2	-1	-3	-1	3
6	V	-1	-3	-3	-3	-1	-3	-3	-4	-3	4	1	-3	1	-1	-3	-2	0	-3	-1	3
7	Q	-1	4	0	-1	-3	4	1	-2	0	-3	-2	3	-1	-3	-2	0	-1	-3	-2	-3
8	I	-1	-3	-3	-4	-1	-2	-3	-4	-3	3	2	-3	1	0	-3	-2	-1	-3	-1	3
9	C	-1	-1	0	-1	5	3	0	-2	4	-2	-2	0	-1	-2	-2	0	2	-2	-1	-1
10	G	0	-3	-1	-1	-3	-2	-2	6	-2	-4	-4	-2	-3	-3	-2	0	-2	-3	-3	-3
11	G	0	-3	-1	-1	-3	-2	-2	6	-2	-4	-4	-2	-3	-3	-2	0	-2	-3	-3	-3
12	L	-2	-2	-4	-4	-1	-2	-3	-4	-3	2	4	-3	2	0	-3	-3	-1	-2	-1	1
13	G	0	-3	-1	-1	-3	-2	-2	6	-2	-4	-4	-2	-3	-3	-2	0	-2	-3	-3	-3
14	N	-2	-1	6	1	-3	0	0	-1	1	-4	-4	0	-2	-3	-2	1	0	-4	-2	-3
15	Q	-1	1	0	0	-3	6	2	-2	0	-3	-2	1	-1	-3	-1	0	-1	-2	-2	-2
16	M	-1	-2	-3	-4	-2	-1	-2	-3	-2	1	3	-2	5	0	-3	-2	-1	-2	-1	1
17	F	-2	-3	-3	-4	-3	-3	-4	-3	-1	0	0	-3	0	7	-4	-3	-2	1	3	-1
18	Q	-1	0	-1	-1	-3	5	1	-2	0	1	-1	1	0	-2	-2	-1	-1	-2	-2	0
19	Y	-2	-2	-3	-3	-3	-2	-3	-3	1	-1	-1	-2	-1	5	-3	-2	-2	2	6	-1
20	A	4	-1	-1	-1	-1	-1	-1	0	-2	-2	-2	-1	-1	-2	-1	2	0	-3	-2	-1
21	F	-2	-3	-3	-4	-3	-3	-4	-3	-1	0	0	-3	0	7	-4	-3	-2	1	3	-1
22	A	3	-2	-1	-2	-1	-1	-1	4	-2	-2	-2	-1	-2	-3	-1	1	-1	-3	-2	-1
23	K	-1	0	-1	-2	-3	0	-1	-3	1	-2	-2	3	-1	2	-2	-1	-1	1	6	-2
24	S	2	-1	-1	-2	-1	-1	-1	-1	-2	-1	1	-1	0	-1	-1	3	0	-3	-2	0
25	L	-2	3	-2	-3	-2	-1	-2	-3	-2	1	3	0	1	0	-3	-2	-1	-2	-1	0
26	Q	0	0	0	-1	-2	4	1	-2	-1	-1	0	0	3	-2	-2	2	0	-2	-2	-1
27	K	-1	2	0	-1	-3	1	0	-2	-1	-2	-2	4	-1	-3	-1	0	2	-3	-2	-2
28	H	-1	0	0	-2	-3	0	0	-2	6	1	-1	2	-1	-1	-2	-1	-1	-3	0	0
29	S	-1	-1	3	-1	-2	-1	-1	-2	0	-1	1	-1	0	1	-2	1	0	0	4	-1
30	N	-1	-1	4	0	-3	-1	-1	3	0	-3	-3	-1	-2	0	-3	0	-1	-1	4	-3
31	T	-1	-2	-1	-2	-2	-1	-2	-2	-2	1	-1	-1	-1	-2	5	0	3	-3	-2	0
32	P	-1	0	-2	-1	-3	0	-1	-2	-2	-3	-3	2	-2	-4	7	-1	-1	-4	-3	-3
33	V	-1	-3	-3	-4	-1	-2	-3	-4	-3	2	2	-3	1	-1	-3	-2	0	-3	-1	4
34	L	-2	3	-2	-3	-2	-1	-2	-3	0	0	2	0	1	1	-3	-2	-1	0	4	-1
35	L	-2	-3	-4	-4	-2	-3	-3	-4	-3	3	3	-3	1	3	-3	-3	-1	-1	1	1
36	D	-2	-2	1	6	-4	0	1	-2	-1	-4	-4	-1	-3	-4	-2	0	-1	-5	-3	-4

The command line of PSI-BLAST that was used is as follows: psiblast-db<LOCAL NR database name> -max_target_seqs 2500-in_msa<MSA file in FAST format> -out <results output file> -outfmt "7sskingdoms sscinames scomnames sseqid stitle evalue length pident" –out_pssm<PSSM file output> -out_ascii_pssm<PSSM (ascii) output> -num_iterations 6 –num_threads 8

This PSI-BLAST search resulted in an initial 2515 hits. There were 787 hits with greater than 22% sequence identity to FutC. 396 hits were of greater than 275 amino acids in length. Additional analysis of the hits was performed, including sorting by percentage identity to FutC, comparing the sequences by BLAST to an existing α(1,2) fucosyltransferase inventory (of known α(1,2) fucosyltransferases, to eliminate known lactose-utilizing enzymes and duplicate hits), and manual annotation of hits to identify those originating from bacteria that naturally exist in the gastrointestinal tract. An annotated list of the novel α(1,2) fucosyltransferases identified by this screen are listed in Table 1. Table 1 provides the bacterial species from which the enzyme is found, the GenBank Accession Number, GI Identification Number, amino acid sequence, and % sequence identity to FutC.

Multiple sequence alignment of the 4 known α(1,2) FTs used for the PSI-BLAST query and 12 newly identified α(1,2) FTs is shown in FIG. 4.

Example 2: Validation of novel α(1,2) FTs

To test for lactose-utilizing fucosylatransferase activity, the production of fucosylated oligossacharides (*i.e.*, 2'-FL) is evaluated in a host organism that expresses the candidate enzyme (*i.e.*, syngene) and which contains both cytoplasmic GDP-fucose and lactose pools. The production of fucosylated oligosaccharides indicates that the candidate enzyme-encoding sequence functions as a lactose-utilizing α(1,2)fucosyltransferase. Of the identified hits, 12 novel α(1,2) fucosyltransferases were further analyzed for their functional capacity to produce 2'-fucosyllactose: *Prevotella melaninogenica* FutO, *Clostridium bolteae* FutP, *Clostridium bolteae* +13 FutP, *Lachnospiraceae sp.* FutQ, *Methanosphaerula palustries* FutR, *Tannerella sp.* FutS, *Bacteroides caccae* FutU, *Butyrivibrio* FutV, *Prevotellaa sp.* FutW, *Parabacteroides johnsonii* FutX, *Akkermansia muciniphilia* FutY, *Salmonella enterica* FutZ, and *Bacteroides sp.* FutZA.

Syngenes were constructed comprising the 12 novel α(1,2) FTs in the configuration as follows: EcoRI – T7g10 RBS – syngene – XhoI. FIG. 5A and FIG. 5B show the syngene fragments after PCR assembly and gel-purification.

The candidate α(1,2) FTs (*i.e.*, syngenes) were cloned by standard molecular biological techniques into an exemplary expression plasmid pEC2-(T7)-Fut syngene-rcsA-thyA. This plasmid utilizes the strong leftwards promoter of bacteriophage λ (termed P_L) to direct expression of the candidate genes (Sanger, F. et al. (1982). J Mol Biol 162, 729-773). The promoter is controllable, e.g., a trp-cI construct is stably integrated the into the E.coli host's genome (at the ampC locus), and control is implemented by adding tryptophan to the growth media. Gradual induction of protein expression is accomplished using a temperature sensitive cI repressor. Another similar control strategy (temperature independent expression system) has been described (Mieschendahl et al., 1986, Bio/Technology 4:802-808). The plasmid also carries the *E. coli rcsA* gene to up-regulate GDP-fucose synthesis, a critical precursor for the synthesis of fucosyl-linked oligosaccharides. In addition, the plasmid carries a β-lactamase (*bla*) gene for maintaining the plasmid in host strains by ampicillin selection (for convenience in the laboratory) and a native *thyA* (thymidylate synthase) gene as an alternative means of selection in *thyA*$^-$ hosts.

The expression constructs were transformed into a host strain useful for the production of 2'-FL. The host strain used to test the different α(1,2) FT candidates incorporates all the above genetic modifications described above and has the following genotype:

Δ*ampC*::$P_{trp}{}^{B}$*cI*, Δ(*lacI-lacZ*)::*FRT*, P_{lacIq}*lacY*$^+$, Δ*wcaJ*::FRT, *thyA*::Tn*10*, Δ*lon*:(*npt3*, *lacZ*$^+$), Δ*lacA*

The *E. coli* strains harboring the different α(1,2) FT candidate expression plasmids were analyzed. Strains were grown in selective media (lacking thymidine) to early exponential phase. Lactose was then added to a final concentration of 0.5%, and tryptophan (200 μM) was added to induce expression of each candidate α(1,2) FT from the P_L promoter. At the end of the induction period (~24 h) the culture supernatants and cells were harvested. Heat extracts were prepared from whole cells and the equivalent of $0.2OD_{600}$ units of each strain analyzed for the presence of 2'-FL by thin layer chromatography (TLC), along with 2μl of the corresponding clarified culture supernatant for each strain.

FIG. 6 shows the oligosaccharides produced by the α(1,2) FT-expressing bacteria, as determined by TLC analysis of the culture supernatant and extracts from the bacterial cells. 2'FL was produced by exogenous expression of WbgL (used as control), FutO, FutP, FutQ, FutR, FutS, FutU, FutW, FutX, FutZ, and FutZA.

Table 4 summarizes the fucosyltransferase activity for each candidate syngene as determined by the 2'FL synthesis screen described above. 11 of the 12 candidate α(1,2) FTs were found to have lactose-utilizing fucosyltransferase activity.

Table 4. 2'FL synthesis screen results

	syngene	24h OD (induced)	2'-FL culture medium	2'-FL cell extract			
Escherichia coli	WbgL	9.58	5	5	**pG204**	pEC2-WbgL-rcsA-thyA	E640
Prevotella melaninogenica	FutO	12.2	3	2	**pG393**	pEC2-(T7)FutO-rcsA-thyA	E985
Clostridium bolteae	FutP	10.4	1	2	**pG394**	pEC2-(T7)FutP-rcsA-thyA	E986
Lachnospiraceae sp.	FutQ	10.6	3	4	**pG395**	pEC2-(T7)FutQ-rcsA-thyA	E987
Methanosphaerula palustris	FutR	11.9	0	1	**pG396**	pEC2-(T7)FutR-rcsA-thyA	E988
Tannerella sp.	FutS	11.3	2	3	**pG397**	pEC2-(T7)FutS-rcsA-thyA	E989
Bacteroides caccae	FutU	12.1	0	2	**pG398**	pEC2-(T7)FutU-rcsA-thyA	E990
Butyrivibrio	FutV	11.3	0	1	**pG399**	pEC2-(T7)FutV-rcsA-thyA	E991
Prevotella sp.	FutW	10.5	3	3	**pG400**	pEC2-(T7)FutW-rcsA-thyA	E992
Parabacteroides johnsonii	FutX	10.7	3	5	**pG401**	pEC2-(T7)FutX-rcsA-thyA	E993
Akkermansia muciniphilia	FutY	9.1	0	0	**pG402**	pEC2-(T7)FutY-rcsA-thyA	E994
Salmonella enterica	FutZ	11.0	0	3	**pG403**	pEC2-(T7)FutZ-rcsA-thyA	E995
Bacteroides sp.	FutZA	9.9	3	3	**pG404**	pEC2-(T7)FutZA-rcsA-thyA	E996

Example 3: Characterization of cultures expressing novel α(1,2) FTs

Further characterization of the bacterium expressing novel α(1,2) FTs FutO, FutQ, and FutX was performed. Specifically, proliferation rate and exogenous α(1,2) FT expression was examined.

Expression plasmids containing fucosyltransferases WbgL (plasmid pG204), FutN (plasmid pG217), and novel α(1,2) FTs FutO (plasmid pG393), FutQ (plasmid pG395), and FutX (pG401) were introduced into host bacterial strains. For example, the host strains utilized has the following genotype: Δ*ampC*::P_{trp}^{B}*cI*, Δ(*lacI-lacZ*)::*FRT*, P_{laclq}*lacY*$^{+}$, Δ*wcaJ*::FRT, *thyA*::Tn*10*, Δ*lon*:(*npt3*, *lacZ*$^{+}$), Δ*lacA*

Bacterial cultures expressing each exogenous fucosyltransferase were induced by addition of tryptophan (to induce expression of the exogenous fucosyltransferases) in the presence of lactose. Growth of the cultures was monitored by spectrophotometric readings at A600 at the following timepoints: 4 hours and 1 hour before induction, at the time of induction (time 0), and 3 hours, 7 hours, and 24 hours after induction. The results are shown in FIG. 7, and indicate that expression of the exogenous fucosyltransferase did not prevent cell proliferation. Furthermore, the growth curve for the bacterial cultures expressing the novel α(1,2) fucosyltransferases FutO, FutQ, and FutX is similar to those expressing the known α(1,2)FT enzymes WbgL and FutN.

Protein expression was also assessed for the bacterial cultures expressing each fucosyltransferase after induction. Cultures were induced as described previously, and

protein lysates were prepared from the bacterial cultures at the time of induction (0 hours), 3 hours, 7 hours, and 24 hours after induction. The protein lysates were run on an SDS-PAGE gel and stained to examine the distribution of proteins at each time point. As shown in FIG. 8, induction at 7hours and 24 hours showed increases in a protein band at around 20-28 kDa for bacterial cultures expressing exogenous FutN, FutO, and FutX. These results indicate that induction results in significant expression of the exogenous fucosyltransferases.

Finally, additional TLC analysis to assess the efficiency and yield of 2'FL production in bacterial cultures expressing novel α(1,2) FTs FutO, FutQ, and FutX compared to known fucosyltransferases WbgL and FutN. Cultures were induced at 7 hours and 24 hours, and run out on TLC. FIG. 9A shows the level of 2'FL in the cell supernatant. The level of 2'FL found in the bacterial cells were also examined. As shown in FIG. 9B, 2'FL was produced in cell lysates from bacteria expressing the novel α(1,2) FTs FutO, FutQ, and FutX at 7 hours and 24 hours after induction.

Example 4: FutN exhibits increased efficiency for production of 2'FL

Fucosylated oligosaccharides produced by metabolically engineered *E. coli* cells to express *B. vulgatus* FutN was purified from culture broth post-fermentation.

Fermentation broth was harvested and cells were removed by sedimentation in a preparative centrifuge at 6000 x g for 30 min. Each bioreactor run yields about 5-7 L of partially clarified supernatant. A column packed with coarse carbon (Calgon 12x40 TR) of ~1000 ml volume (dimension 5 cm diameter x 60 cm length) was equilibrated with 1 column volume (CV) of water and loaded with clarified culture supernatant at a flow rate of 40 ml/min. This column had a total capacity of about 120 g of sugar. Following loading and sugar capture, the column is washed with 1.5 CV of water, then was eluted with 2.5 CV of 50% ethanol or 25% isopropanol (lower concentrations of ethanol at this step (25-30%) may be sufficient for product elution.) This solvent elution step released about 95% of the total bound sugars on the column and a small portion of color bodies (caramelized sugars). A volume of 2.5 L of ethanol or isopropanol eluate from the capture column was rotary-evaporated at 56 C° and a sugar syrup in water was generated. A column (GE Healthcare HiScale50/40, 5x40cm, max pressure 20 bar) connected to a Biotage Isolera One FLASH Chromatography System was packed with 750 ml of a Darco Activated Carbon G60 (100-mesh): Celite 535 (coarse) 1:1 mixture (both column packings were obtained from Sigma). The column was equilibrated with 5 CV of water and loaded with sugar from step 3 (10-50 g, depending on the ratio of 2'-FL to contaminating lactose), using either a celite loading

cartridge or direct injection. The column was connected to an evaporative light scattering (ELSD) detector to detect peaks of eluting sugars during the chromatography. A four-step gradient of isopropanol, ethanol or methanol was run in order to separate 2'-FL from monosaccharides (if present), lactose and color bodies. Fractions corresponding to sugar peaks were collected automatically in 120-ml bottles, pooled.

The results from two fermentation runs are shown in FIG. 10A and FIG. 10B. The cultures were grown for 136 (run 36B) or 112 hours (run 37A), and the levels of 2'-FL produced was analyzed by TLC analysis. As shown in both FIG. 10A and FIG. 10B, the 2'-fucosyllactose was produced at 40 hours of culture, and production continued to increase until the end point of the fermentation process. The yield of 2'-FL produced from run 36B was 33 grams per liter. The yield of 2'-FL produced from run 37A was 36.3 grams per liter. These results indicate that expression of exogenous FutN is suitable for high yield of 2'-fucosyllactose product.

Table 1. Hits from PSI-BLAST multiple sequence alignment query for novel α(1,2) fucosyltransferases

Bacterium names	Accession No.	GI No.	Gene name [bacterium]	% identity FutC	Alias	SEQUENCE	SEQ ID NO
Helicobacter pylori	AAD29869.1	4808599	alpha-1,2-fucosyltransferase [Helicobacter pylori]	98	FutC	MAFKVVQICGGLGNQMFQYAFAKSLQKHSNTPVLLDITSFDWSDRKMQLELFPINLPYASAKEIAIAKMQH LPKLVRDALKCMGFDRVSQEIVFEYEPELLKPSRLTYFYGYFQDPRYFDAISPLIKQTFTLPPPPENNKNNKKE EEYHRKLSLILAAKNSVFVHIRRGDYVGIGCQLGIDYQKKALEYMAKRVPNMELFVFCEDLEFTQNLDLGYPF MDMTTRNKEEEAYWDMLLMQSCQHGIIANSTYSWWAAYLIENPEKIIIGPKHWLFGHENILCKEWVKIESH FEVKSQKYNA	1
Helicobacter mustelae;Helicobacter mustelae 12198	YP_003517185.1	291277413	alpha-1,2-fucosyltransferase [Helicobacter mustelae 12198]	70.85	FutL	MDFKIVQVHGGLGNQMFQYAFAKSLQTHLNIPVLLDTTWFDYGNRELGLHLFPIDLQCASAQQIAAAHMQ NLPRLVRGALRRMGLGRVSKEIVFEYMPELFEPSRIAYFHGYFQDPRYFEDISPLIKQTFTLPHPTEHAEQYSR KLSQILAAKNSVFVHIRRGDYMRLGWQLDISYQLRAIAYMAKRVQNLELFLFCEDLEFVQNLDLGYPFVDMT TRDGAAHWDMMLMQSCKHGIITNSTYSWWAAYLIKNPEKIIIGPSHWIYGNENILCKDWVKIESQFETKS	2
Bacteroides;Bacteroides vulgatus ATCC 8482;Bacteroides sp. 4_3_47FAA;Bacteroides sp. 3_1_40A;Bacteroides vulgatus PC510;Bacteroides vulgatus CL09T03C04;Bacteroides vulgatus dnLKV7;Bacteroides vulgatus CAG:6	YP_001300461.1	150005717	glycosyl transferase family protein [Bacteroides vulgatus ATCC 8482]	24.83	FutN	MRLIKVTGGLGNQMFIYAFYLRMKKYYPKVRIDLSDMMHYKVHYGYEMHRVFNLPHTEFCINQPLKKVIEFL FFKKIYERKQAPNSLRAFEKKYFWPLLYFKGFYQSERFFADIKDEVRESFTFDKNKANSRSLNMLEILDKDENA VSIHIRRGDYLQPKHWATTGSVCQLPYYQNAIAEMSRRVASPSYYIFSDDIAWVKENLPLQNAVYIDWNTD EDSWQDMMLMSHCKHHIICNSTFSWWGAWINPNMDKTVIVPSRWFQHSEAPDIYPTGWIKVPVS	3
Escherichia coli;Escherichia coli UMEA 3065-1	WP_021554465.1	545259828	protein [Escherichia coli]	23.13	Wbgl	MSIIRLQGGLGNQLFQFSFGYALSKINGTPLYFDISHYAENDDHGGYRLNNLQIPEEYLQYYTPKINNIYKLLVR GSRLYPDIFLFLGFCNEFHAYGYDFEYIAQKWKSKKYIGYWQSEHFFHKHILDLKEFFIPKNVSEQANLLAAKIL ESQSSLSIHIRRGDYIKNKTATLTHGVCSLEYYKKALNKIRDLAMIRDVFIFSDDIFWCKENIETLLSKKYNIYYSE DLSQEEDLWLMSLANHHIIANSSFSWWGAYLGSSASQIVYPTPWYDITPKNTYIPIVNHWINVDKHSSC	4

Helicobacter bilis;Helicobacter bilis ATCC 43879	WP_00521 9731.1	491361813	predicted protein [Helicobacter bilis]	36.79	FutD		MGDYKIVELTCGLGNQMFQYAFAKALQKHLQVPVLLDKTWYDTQDNSTQFSLDIFNVDLEYATNTQIEKAK ARVSKLPGLLRKMFGLKKHNIAYSQSFDFHDEYLLPNDFTYFSGFFQNAKYLKGLEQELKSIFYYDSNNFSNFG KQRLELILQAKNSIFIHIRRGDYCKIGWELGMDYYKRAIQYIMDRVEEPKFFIFGATDMSFTEQFQKNLGLNE NNSANLSEKTITQDNQHEDMFLMCYCKHAILANSSYSFWSAYLNNDANNIVIAPTPWLLDNDNIICDDWIKI	5
Escherichia coli	AAO37698.1	37788088	fucosyltransferase [Escherichia coli]	25.94	WbsJ		MEVKIIGGLGNQMFQYATAFAIAKRTHQNLTVDISDAVKYKTHPLRLVELSCSSEFVKKAWPFEKYLFSEKIPH FMKKGMFRKHYVEKSLEYDPDIDTKSINKKIVGYFQTEKYFKEFRHELIKEFQPKTKFNSYQNELLNLIKENDTC SLHIRRGDYVSSKIANETHGTCSEKYFERAIDYLMNKGVINKKTLLFIFSDDIKWCRENIFFNNQICFVQGDAY HVELDMLLMSKCKNNIISNSSFSWWAAWLNENKNKTVIAPSKWFKKDIKHDIIPESWVKL	6
Vibrio cholerae	BAA33632.1	3721682	probable beta-D-galactoside 2-alpha-L-fucosyl transferase [Vibrio cholerae]	25.94	WbIA		MIVMKISGGLGNQLFQYAVGRAIAIQYGVPLKLDVSAYKNYKLHNGYRLDQFNINADIANEDEIFHLKGSSN RLSRIILRRLGWLKKNTYYAEKQRTIYDVSVFMQAPRYLDGYWQNEQYFSQIRAVLLQELWPNQPLSINAQA HQIKIQQTHAVSIHVRRGDYLNHPEIGVLDIDYYKRAVDYIKEKIEAPVFFVFSNDVAWCKDNFNFIDSPVFIE DTQTEIDDLMLMCQCQHNIVANSSFSWWAAWLNSNVDKIVIAPKTWMAENPKGYKWVPDSWREI	7
Bacteroides fragilis;Bacteroides fragilis NCTC 9343;Bacteroides fragilis YCH46;Bacteroides fragilis HMW 615	YP_099118.1	53713126	alpha-1,2-fucosyltransferase [Bacteroides fragilis YCH46]	24.58	Bft2		MIVSSLRGGLGNQMFIYAMVKAMALRNNVPFAFNLTTDFANDEVYKRKLLSYFALDLPENKKLTFDFSYGN YYRRLSRNLGCHILHPSYRYICEERPPHFESRLISSKITNAFLEGYWQSEKYFLDYKQEIKEDFVIQKKLEYTSYLE LEEIKLLDKNAIMIGVRRYCESDVAPGGVLEDDYYKCAMDIMASKVTSPVFFCFSQDLEWVEKHLAGKYPVR LISKKEDDSGTIDDMFLMMHFRNYIISNSSFYWWGAWLSKYDDKLVIAPGNFINKDSVPESWFKLNVR	8
Escherichia coli;Escherichia coli KTE84	WP_00159 2236.1	486356116	protein [Escherichia coli]	24.25	Wbg N		MSIVVARLAGGLGNQMFQYAKGYAESVERNSYLKLDLRGYKNYTLHGGFRLDKLNIDNTFVMSKKEMCIFP NFIVRAINKFPKLSLCSKRFESEQYSKKINGSMKGSVEHGFWQNERYFLEHKEKLREITPINLNLDAKELSDVI RCTNSVSVHIRRGDYVSNVEALKIHGLCTERYYIDSIRYLKERFNNLVFFVFSDDIEWCKKYKNEIFSRSDDVKFI EGNTQEVDMWLMSNAKYHIIANSSFSWWGAWLKNYDLGITIAPTPWFEREELNSFDPCPEKWVRIEK	9
Prevotella melaninogenica;Prevotella melaninogenica ATCC 25845	YP_003814512.1	302346214	glycosyltransferase family 11 [Prevotella melaninogenica ATCC 25845]	31.1	FutO		MKIVKILGGLGNQMFQYALYLSLQESFPKERVALDLSSFHGYHLHNGFELENIFSVTAQKASAADIMRIAYYY PNYLLWRIGKRFLPRRKGMCLESSSLREDESVLRQEGNRYFDGYWQDERYFAAYREKVLKAFTFPAFKRAEN LSLLEKLDENSIALHVRRGDYVGNNLYQGICLDYYYRTAIEKMCAHVTPSLFCFSNDITWCQQHLQPYLKAP VVYVTWNTGVESVRDMQLMSCCAHNIIANSSFSWWGAWLNQNREKVVIAPKKWLNMEECHFTLPASWI KI	10

Clostridium bolteae;Clostridium bolteae 90A9;Clostridium bolteae 90B3;Clostridium bolteae 9088	WP_00257 0768.1	488634090	protein [Clostridium bolteae]	29.86	FutP		MFQYALYKAFEQKHIDVYADLAWYKNKSVKFELYNFGIKINVASEKDINRLSDCQADFVSRIRRKIFGKKKSFV SEKNDSCYENDILRMDNVYLSGYWQTEKYFSNTREKLLEDYSFALVNSQVSEWEDSIRNKNSVSIHIRRGDYL QGELYGGICTSLYYAEAIEYIKMRVPNAKFFVFSDDVEWVKQQEDFKGFVIVDRNEYSSALSDMYLMSLCKH NIIANSSFSWWAAWLNRNEEKIVIAPRRWLNGKCTPDIWCKKWIRI	11
Lachnospiraceae bacterium 3_1_57FAA_CT1	WP_00925 1343.1	496545268	protein [Lachnospiracea e bacterium 3_1_57FAA_CT1]	29.25	FutQ		MVIVQLSGGLGNQMFEYAYLSLKAKGKEVKIDDVTCYEGPGTRPRQLDVFGITYDRASREELTEMTDASM DALSRVRRKLTGRRTKAYRERDINFDPLVMEKDPALLEGCFQSDKYFRDCEGRVREAYRFRGIESGAFPLPED YLRLEKQIEDCQSVSVHIRRGDYLDESHGGLYTGICTEAYYKEAFARMERLVPGARFFLFSNDPEWTREHFES KNCVLVEGSTEDTGYMDLYLMSRCRHNIIANSSFSWWGAWLNENPEKKVIAPAKWLNGRECRDIYTERMI RL	12
Methanosphaerula palustris;Methanosphaerula palustris E1-9c	YP_002467 213.1	219852781	glycosyl transferase family protein [Methanosphae rula palustris E1-9c]	28.52	FutR		MIIVRLKGGLGNQLSQYALGRKIAHLHNTELKLDTTWFTTISSDTPRTYRLNNYNIIGTIASAKEIQLIERGRAQ GRGYLLSKISDLLTPMYRRTIYVRERMHTFDKAILTVPDNVYLDGYWQTEKYFKDIEEILRREVTLKDEPDSINLE MAERIQACHSVSLHVRRGDYVSNPTTQQFHGCCSIDYYNRAISLIEEKVDDPSFFIFSDDLPWAKENLDIPGE KTFVAHNGPEKEYCDLWLMSLCQHHIIANSSFSWWGAWLGQDAEKMVIAPRRWALSESFDTSDIIPDSWI TI	13
Tannerella sp. CAG:118	WP_02192 9367.1	547187521	glycosyl transferase family 11 [Tannerella sp. CAG:118]	28.38	FutS		MVRIVEIIGGLGNQMFQYAFSLYLKNKSHIWDRLYVDIEAMKTYDRHYGLELEKVFNLSLCPISNRLHRNLQK RSFAKHFVKSLYEHSECEFDEPVYRGLRPYRYYRGYWQNEGYFVDIEPMIREAFQFNVNILSKKTKAIASKMR RELSVSIHVRRGDYENLPEAKAMHGGICSLDYYHKAIDFIRQRLDNNICFYLFSDDINWVEENLQLENRCIID WNQGEDSWQDMYLMSCCRHHIIANSSFSWWAAWLNPNKNKIVLTPNKWFNHTDAVGIVPKSWIKIPVF	14
Bacteroides caccae;Bacteroides caccae ATCC 43185	WP_00567 5707.1	491925845	protein [Bacteroides caccae]	28.09	FutU		MKIVKILGGLGNQMFQYALFLSLKERFPHEQVMIDTSCFRNYPLHNGFEVDRIFAQKAPVASWRNILKVAYP YPNYRFWKIGKYILPKRKTMCVERKNFSFDAAVLTRKGDCYYDGYWCQHEEYFCDMKETIWEAFSFPEPVDG RNKEIGALLQASDSASLHVRRGDYVVNHPLFRGICDLDYYKRAIHYMEERVNPQLYCVFSNDMAWCESHLRA LLPGKEVVYVDWNKGAESYVDMRLMSLCRHNIIANSSFSWWGAWLNRNPQKVVVAPERWMNSPIEDPV SDKWIKL	15
Butyrivibrio sp. AE2015	WP_02277 2718.1	551028636	protein [Butyrivibrio sp. AE2015]	27.8	FutV		MIIIQLKGGLGNQMFQYALYKSLKKRGKEVKIDDKTGFVNDKLRIPVLSRWGVEYDRATDEEIINLTDSKMDL FSRIRRKLTGRKTFRIDEESGKFNPEILEKENAYLVGYWQCDKYFDDKDVVREIREAFEKKPQELMTDASSWS TLQQIECCESVSLHVRRTDYVDEEHIHHNICTEKYYKNAIDRVRKQYPSAVFFIFTDDKEWCRDHFKGPNFIV VELEEGDGTDIAEMTLMSRCKHHIIANSSFSWWAAWLNDSPEKIVIAPQKWINNRDMDDIYTERMTKIAL	16

Prevotella sp. CAG:891	WP_02248 1266.1	548264264	uncharacterized protein [Prevotella sp. CAG:891]	27.4	FutW	MRLVKMIGGLGNQMFIYAFYLQMRKRFSNVRIDLTDMMHYNVHYGYELHKVFGLPRTEFCMNQPLKKVL EFLFRTIVERKQHGRMEPYTCQYVWPLVYFKGFYQSERYFSEVKDEVRECFTFNPALANRSSQQMMEQIQ NDPQAVSIHIRRGDYLNPKHYDTIGCICQLPYYKHAVSEIKKYVSNPHFYVFSEDLDWVKANLPLENAQYIDW NKGADSWQDMMLMSCCKHHIICNSTFSWWAAWLNPSVEKTVIMPEQWTSRQDSWDFVASCGRWVRV KTE	17
Parabacteroides johnsonii;Parabacteroides CL02T12C29	WP_00815 5883.1	495431188	glycosyl transferase [Parabacteroides johnsonii]	26.69	FutX	MRLIKMIGGLGNQMFIYAFYLKMKHHYPDTNIDLSDMVHYKVHNGYEMNRIFDLSQTEFCINRTLKKILEFL FFKKIYERRQDPSTLYPYEKRYFWPLLYFKGFYQSERFFFDIKDDVRKAFSFNLNIANPESLELLKQIEVDDQAV SIHIRRGDYLLPRHWANTGSVCQLPYYKNAIAEMENRITGPSYYVFSDDISWVKENIPLKKAVYVTWNKGED SWQDMMLMSHCRHHIICNSTFSWWGAWLNPRKEKIVIAPCRWFQHKETPDMYPKEWIKVPIN	18
Akkermansia muciniphila;Akkermansia muciniphila ATCC BAA-835	YP_001877 555.1	187735443	glycosyl transferase family protein [Akkermansia muciniphila ATCC BAA-835]	25.67	FutY	MRLFGGLGNQLFQYAFLFALSRQGGKARLETSSYEHDDKRVCELHHFRVSLPIEGGPPPWAFRKSRIPACLRS LFAAPKYPHFREEKRHGFDPGLAAPPRRHTYFKGYFQTEQYFLHCREQLCREFRLKTPLTPENARILEDIRSCCS ISLHIRRTDYLSNPYLSPPPLEYYLRSMAEMEGRLRAAGAPQESLRYFIFSDDIEWARQNLRPALPHVHVDIND GGTGYFDLELMRNCRHHIIANSTFSWWAAWLNEHAEKIVIAPRIWFNREEGDRYHTDDALIPGSWLRI	19
Salmonella enterica;Salmonella enterica subsp. enterica serovar Poona str. ATCC BAA-1673	WP_02321 4330.1	555221695	fucosyltransferase [Salmonella enterica]	25.99	FutZ	MYSCLSGGLGNQMFQYAAAYILKQYFQSTTLVLDDSYYYSQPKRDTVRSLELNQFNISYDRFSFADEKEKIKLL RKFKRNPFPKQISEILSIALFGKYALSDRAFYTFETIKNIDKACLFSFYQDADLLNKYKQLILPLFELRDDLLDICKN LELYSLIQRSNNTTALHIRRGDYVTNQHAAKYHGVLDISYYNHAMEYVERERGKQNFIIFSDDVRWAQKAFL ENDNCYVINNSDYDFSAIDMYLMSLCKNNIIANSTYSWWGAWLNKYEDKLVISPKQWFLGNNETSLRNAS WITL	20
Bacteroides sp. CAG:633	WP_02216 1880.1	547748823	glycosyltransferase family 11 [Bacteroides sp. CAG:633]	26.01	FutZA	MRLIKMIGGLGNQMFIYAFYLRMKKRYPKVRIDLSDMVHYHVHHGYEMHRVFNLPHTEFCINQPLKKVIEF LFFKKIYERKQDPNSLRAFEKKYLWPLLYFKGFYQSERFFADIKDEVRKAFTFDSSKVNARSAELLRRLDADAN AVSLHIRRGDYLQPQHWATTGSVCQLPYYQNAIAEMNRRVAAPSYYVFSDDIAWVKENIPLQNAVYIDWN KGEESWQDMMLMSHCRHHIICNSTFSWWGAWLDPHEDKIVVPNRWFQHCETPNIYPAGWVKVAIN	21
Clostridium sp. CAG:306	WP_02224 7142.1	547839506	alpha-1 2-fucosyltransferase [Clostridium sp. CAG:306]	34.28		MEKIKIVKLQGGMGNQMFQYAFGKGLESKFGCKVLFDKINYDELQKTIINNTGKNAEGICVRKYELGIFNLNI DFATAEQIQECIGEKLNKACYLPGFIRKIFNLSKNKTVSNRIFEKKYGEYDEEILKDYSLAYYDGYFQNPKYFEDI SDKIKKEFTLPEIKNHDIYNKKLLEKITQFENSVFIHVRRDDYLNINCEIDLDYYQKAVKYILKHIENPKFFVFCAE DPDYIKNHFDIGYDFELVGENNKTQDTYYENMRLMMACKHAIIANSSYSWWAAWLSDYDNKIVIAPTPWL PGISNEIICKNWIQIKRGISNE	22

Prevotella sp. oral taxon 306;Prevotella sp. oral taxon 306 str. F0472	WP_00943 4595.1	497004957	protein [Prevotella sp. oral taxon 306]	32.11		MKIVKILGGLGNQMFQYALYLSLQESFPKERVALDLSCFNGYHLHNGFELERIFSLTAQKASAATIMRIAYYYP NYLLWRIGKRLLPRRKTMCLESSTFRYDESVLTREGNRYFDGYWQDERYFVACREKVLKAFTFPAFKRTENLS LLRKLDKNSVAIHVRRGDYIGNQLYQGICDLDYYRAAIDKISTYVTPSVFCIFSNDIAWCQTHLQPYLKAPVVY VTWNTGTESYRDMQLMSCCAHNIIANSSFSWWGAWLNQNNEKVVIAPKRWLNNMDDCQFPLPASWVKI	23
Brachyspira sp. CAG:484	WP_02191 7109.1	547139308	glycosyl transferase family 11 [Brachyspira sp. CAG:484]	30.14		MQLVKLMGGLGNQMFQYAFAKALGDKNILFYGDYKKHSLRKVELNRFKCKAVYIPRELFKYLKFVFTKFDKIE YMRSGIYVPEYLNRDGNHIYIGFWQTEKYFKQIRPRLLKDFTPRKKLDRENAGIISKMQQINSVSVHIRRTDYV DESHIYGDTNLDYYKRAIEYISSKIENPEFFFFSDDMAYVKEKFAGLKFPHSFIDINSGNNSYKDLILMKNCKHN IIANSTFSWWGAWLNENEEKIVIAPAKWFVTGENDKDIVPDEWIKL	24
Thalassospira profundimaris;Thalassospira profundimaris WP0211	WP_00888 9330.1	496164823	glycosyl transferase family 11 [Thalassospira profundimaris]	30		MVIVKLLGGLGNQMFQYATGRAVASRLDVELLLDVSAFAHYDLRRYELDDWNITARLATSEELARSGVTAA PPSFFDRIARFLRIDLPVNCFREASFTYDPRILEVSSPVYLDGYWQSERYFLDIEKKLRQEFQLKASIDANNHSFK KKIDGLGKQAVSLHVRRGDYVTNPQTASYHGVCSLDYYRAAVDYIAEHVSDPCFFVFSDDLEWVQTNLNIK QPIVLVDANGPDNGAADMALMMACRHHIIANSSFSWWGSWLNPLNDKIVAPKKWFGRANHDTTDLVP DSWVRL	25
Acetobacter sp. CAG:267	WP_02207 8656.1	547459369	alpha-1 2-fucosyltransfera se [Acetobacter sp. CAG:267]	29.9		MAVSPQESKYSAHVSPDKPLRIVRLGGGLGNQMFQYAFGLAAGDVLWDNTSFLTNHYRSFDLGLYNISGDF ASNEQIKKCKNEIRFKNILPRSIRKKFNLGKFIYLKTNRVCERQINRYEPELLSKDGDVYYDGVFQTEKYFKPLRE RLLHDFTLTKPLDAANLDMLAKIRAADAVAVHIRRGDYLNPRSPFTYLDKDYFLNAMDYIGKRVDKPHFFIFS SDTDWVRTNIQTAYPQTIVEINDEKHGYFDLELMRNCRHNIIANSTFSWWGAWLNTNPDKIVVAPKQWFR PDAAEYSGDIVPNDWIKL	26
Dysgonomonas mossii;Dysgonomonas mossii DSM 22836	WP_00684 2165.1	493896281	protein [Dysgonomonas mossii]	29.9		MVTVLLSGGLGNQMFQYAAAKSLAIRLNTALSVDLYTFSKKTQATVRPYELGIFNIEDVVETSSLKAKAVIKAR PHIQRHRSFFQRFGVFTDTYAILYQPTFEALTGGVIMSGYFQNESYFKNISELLRKDFSFKYPLIGENKDVAGQI SENQSVAVHIRRGDYLNKNSQSNFAILEKDYYEKAINYISAHVKNPEFYVFSEDFDWIKDNLNFKEPVTFID WNKGKDSYIDMQLMSLCKHNIIANSSFSWWSAWLNNSEERKIVAPERWFVDEQKNELLPCFYPQGWIKI	27
Clostridium sp. KLE 1755	WP_02163 6924.1	545396671	glycosyltransfer ase, family 11 [Clostridium sp. KLE 1755]	29.83		>gi|545396671|ref|WP_021636924.1| glycosyltransferase, family 11 [Clostridium sp. KLE 1755]MIIIEISGGLGNQMFQYALGQKFISMGKEVKYDLSFYNDRVQTLRQFELDIFDLDCPVASNSELSRFGK GNSLKSRLKQKLGWDKEKIYEENLDLGYQPRIFELDDIYLSGYWQSELYFKDIREQILRLYTFPIQLDYMNGVFL RKIENSNSVSIHIRRGDYLNENNLKIYGNICTLNYYNKAIQIAKKITNPIIFVFTNDIEWVRKELEIPNMVIVDC NSGKLSYWDMYLMSKCKANIVANSSFSWWGAWLNKNENRIIISPKRWLNNHEQTSTLCDNWIRCGDD	28

Gillisia limnaea;Gillisia limnaea DSM 15749	WP_006988068.1	494045950	alpha-1,2-fucosyltransferase [Gillisia limnaea]	29.28		MFISKNTVIIKLVGGLGNQMFQFAIAKIIAEKEKSEVLVDITFYTELTENTKKFPRHFSLGIFNSSFAIASKKEIDYF TKLSNFNKFKKKLGLNYPTIFHESSFNFKAQVLELKAPIYLNGYFQSFRYFLGKEYVIRKIFKFPDEALDKDNDNI KRKIIGKTSVSLHIRRGDYVNNKKTQQFHGNCTIDYYQSAIAYLSSKLTDFNLIFFSDDIHWVRQQFKNISNQKI YVSGNLNHNSWKDMYLMSLCDHNIIANSSFSWWGAWLNKNPEKIIAPKRWFADTEQDKNSIDLIPSEWY RI	29
Methylotenera mobilis;Methylotenera mobilis JLW8	YP_003048467.1	253996403	glycosyl transferase family protein [Methylotenera mobilis JLW8]	29.19		MLVSRIIGGLGNQMFEYAAARAASLRISVQLKLDLSGFETYDLHAYGLNNFNIVEDVAKKDDYFIGAPESLLKK IKKYLRGLIQLESFRESDLSFDSKVLELNDNTYLDGYWQCERYFIDFDKQIRQDFSFKFAPDALNQRYLELIDSV NAVSVHIRRGDYVSNSTTNEIHGVCDLDYYQRAAEFMRARIGPENLHFFVFSDDTDWVKENISFGSDTTFIS HNDAAKNYEDMRLMSACKHHIIANSSFSWWAAWLNPSKQKVVIAPRQWFKSTLLNSDDIVPASWVRL	30
Runella slithyformis;Runella slithyformis DSM 19594	YP_004658567.1	338214504	glycosyl transferase family protein [Runella slithyformis DSM 19594]	29.14		MIIVKLSGGLGNQLFQYAFGRHLATVNQKELKLDTSALTKTSDWTNRSYALDAFNIRAQEATPEEIKALAGKP NRLLQRVGRKVGITPIQYFQEPHFHFYSSALSIKSSHYLEGYWQSEKYFEAITPILREEFAFTISPSTHAQTIKEKI SNGTSVSIHLRRGDYVKTSKANRYLRPLTMDYYQKAIDYINQRVKNPNFFLFSDDIKWAKSQVTFPPTTHFST GTSAHEDLWLMTHCRHHIIANSTFSWWGAWLNCQPDKVIAPQKWFSTERFDTKDLLPEPWIQL	31
Pseudoalteromonas haloplanktis;Pseudoalteromonas haloplanktis ANT/505	WP_008295845 4.1	489048235	alpha-1,2-fucosyltransferase [Pseudoalteromonas haloplanktis]	29.1		MIKVKAIGGLGNQLFQYATARAIAEKRGDGVVVDMSDFSSYKTHPFCLNKFRCKATYESKPKLINKLLSNEKIR NLLQKLGFIKKYYFETQLPFNEDVLLNNSINYLTGYFQSEKYFLSIRECLLDELTLIEDLNIAETAVSKAIKNAKNSI SIHIRRGDYVSNEGANKTHGVCDSDYFKKALNYFSERKLLDEHTELFISDDIEWCRNNLSFDYKMINFVDGSS ERPEVDMVLMSQCKHQVISNSTFSWWGAWLNKNDEKVVVAPKEWFKSTDLDSTDIVPNQWIKL	32
uncultured bacterium	EKE06679.1	406985989	glycosyl transferase family 11 [uncultured bacterium]	28.67		MLTLKLKGGLGNQMFQYAASHNLAKNKKTKINFDLSFFSDIEVRDIKRDYLLDKFNISADISFDQKNSISGFRK FLVKVISKFFGEVFYYRLKFLSSKYLDGYFQSEKYFKNVEEDIRKDFTLKDEMGVEAKKIEQQIVNSKNSVSLHIR RGDYVDDLKTNIYHGVCNLDYYKRSIKYLKENFGEINIFVFSDDIAWVKENLAFENLQFVSRPDIKDYEELML MSKCEHNIIANSSFSWWGAWLNENKNKIIIAPKEWFQKFNINEKHIVPKSWIRL	33
Clostridium sp. KLE 1755	WP_021636949.1	545396696	glycosyltransferase, family 11 [Clostridium sp. KLE 1755]	28.57		MVIVQLSGGLGNQMFEYALYLSLKAKGKVVKIDDITCYEGPGTRPKQLDVFGVSYERATKQELTEMTDSSLD PYSRIRRKLTGRKTKAYREKDINFDPQVMERDPALLEGCFQSEKYFQDCREQVREAYRFRGIESGAYPLPEAY RRLEKEIADCKSVSVHIRRGDYLEESHGGLYTGICTEQYYQEAFARMEKEVPGAKFFLFSNDPDWTREHFKGE NRILVEGSTEDTGYLDLYLMSKCKHNIIANSSFSWWGAWLNDNPEKKVTAPARWLNGRECRDIYTERMIRI	34
Francisella philomiragia;Francisella philomiragia subsp. philomiragia ATCC 25015	WP_004287502.1	490414974	alpha-1,2-fucosyltransferase [Francisella philomiragia]	28.57		MKIIKIQGGLGNQMFQYAFYKSLKNNCIDCYVDIKNYDTYKLHYGFELNRIFKNIDLSFARKYHKKEVLGKLFSI IPSKFIVKFNKNYILQKNFAFDKAYFEIDNCYLDGYWQSEKYFKKITKDIYDAFTFEPLDSINFEFLKNIQDYNLV SIHVRRGDYVNHPLHGGICDLEYYNKAISFIRSKVANVHFLVFSNDILWCKDNLKLDRVTYIDHNRWMDSYK DMHLMSLCKHNIIANSSFSWWGAWLNQNDDKIVIAPSKWFNDDKINQKDICPNSWVRI	35

Pseudomonas fluorescens;Pseudomonas fluorescens NCIMB 11764	WP_01733 7316.1	515906733	protein [Pseudomonas fluorescens]	28.52		MVIAHLIGGLGNQMFQYAAARALSSAKKEPLLLDTSSFESYTLHQGFELSKLFAGEMCIARDKDINHVLSWQ AFPRIRNFLHRPKLAFLRKASLIIEPSFHYWNGIQKAPADCYLMGYWQSERYFQDAAEEIRKDFTFKLNMSPQ NIATADQILNTNAISLHVRRGDYVNNSVYAACTVEYYQAAIQLLSKRVDAPTFFVFSDDIDWVKNNLNIGFPH CYVNHNKGSESYNDMRLMSMCQHNIIANSSFSWWGAWLNSNADKIVVAPKQWFINNTNVNDLFPPAW VTL
Herbaspirillum sp. YR522	WP_00811 7381.1	495392680	glycosyl transferase family 11 [Herbaspirillum sp. YR522]	28.48		MIATRLIGGLGNQMFQYAAGRALALRVGSPLLLDVSGFANYELRRYELDGFRIDATAASAQQLARLGVNATP GTSLLARVLRKVWPQPADRILREASFTYDARIEQASAPVYLDGYWQSERYFARIRQHLLDEFTLKGDWGSDN AAMAAQIATAGAGAVSLHVRRGDYVSNAHTAQYHGVCSLDYYRDAVAHIGGRVEAPHFFVFSDDHEWVR ENLQIGHPATFVQINSADHGIYDMMLMKSCRHHIIANSSFSWWGAWLNPAEDKIVVAPQRWFKDATNDT RDLIPAAWVRL
Prevotella histicola;Prevotella histicola F0411	WP_00882 2166.1	496097659	rotein [Prevotella histicola]	28.43		MKIVKILGGLGNQMFQYAIYLSLKETFPQENVTVDLSCFHGYHLHNGFEIARIFSLHPDKATVMEILRIAYYYP NYFFWQIGKRVLPQRKTMCTESTKLLFDKSVLQREGDRYFDGYWQDERYFHDCRRTILNTFKFPPFTDDNNL ALLKKMDTNSVSIHVRRGDYVGNKLYQGICDLNYYREAIMKISSYISPSMFCVFSNDIEWCRDNLESFIKAPIY YVDWNSGTESYRDMQLMSCCGHNIIANSSFSWWGAWLNQNSSKIVIAPKRWINLKNCGFMLPSRWVKI
Flavobacterium sp. WG21	WP_01749 4954.1	516064371	protein [Flavobacterium sp. WG21]	28.42		MIVVQLIGGLGNQLFQYAAAKALALQTKQKFSLDVSQFESYKLHNYALNHFNVISKNYKKPNRYLRKIKSFYQ KNVFYKEVDFGYNPDLIHLKGGIIFLEGYFQSEKYFIKYEKEIREDFELRTPLKKETKAAIAKIESVNSVSIHIRRGD YINNPLHNTSKEEYYNKAIEIVENKINNPVYFVFSDDMEWVKANFSTKQFTIFIDFNDASTNFEDLKLMTSCK HNIIANSSFSWWGGWLNKNPDKIVIAPKRWFNDSSINTNDIIPTNWVKI
Polaribacter franzmannii	WP_01894 4517.1	517774309	protein [Polaribacter franzmannii]	28.42		MIIVRIVGGLGNQMFQYAYAKALQQKGYQVKIDITKFKKYNLHGGYQLDQFKIDLETSSPIANVLCRIGLRRS VKEKSLIFDEKFLEIPQREYIKGYFQTEKYFSSITPILRKQFIVQKELCNTTLRYLKEITQKNACSLHIRRGDYISDE KANSVHGTCDLPYYKKSIKRIQDEYKDAHFFIFSDDISWAKKNLTNKNTTFHEHVMPHEDMHLMSLCKHNIT ANSSFSWWGAWLNQHENKTVIAPKNWFVNRENEVACANWIQL
Polaribacter sp. MED152	YP_007670 847.1	472321325	glycosyl transferase family 11 [Polaribacter sp. MED152]	28.42		MVVVRILGGLGNQMFQYAYAKSLAEKGYEVQIDISKFKSYKLHGGYHLDKFRIDLETANSSSAFLSKIGLKKTIK EPNLLFHKDLLKVNNNAFIKGYFQAEQYFSDIREILINQFKIKKELAKSTLAIKNQIELLKTTCSLHVRRGDYISDK KANKVHGTCDLDYYSSAIEHISKQNSNVHFFVFSDDIAWVKDNLNITNATYIDHNVIPHEDMYLMTLCNHNI TANSSFSWWGAWLNQNPDKIVIAPKNWFVDKENEVACKSWITL

Methanococcus maripaludis;Methanococcus maripaludis C7	YP_001329558.1	150402264	glycosyl transferase family protein [Methanococcus maripaludis C7]	28.19		MKIIQLKGGLGNQMFQYALYKSLKKRGQEVLLDISWYLKNNAHNGYELEWVFGLSPEYASIRQCFKLGDIPI NLIYNVKRKVFPKKTHFFEKSNFNYDNNVFEVTNGYFEGYWQNENYFKNFRSEILNDFSFKNIDKRNAEFSEY LKSINSVSVHVRRGDYVTNQKALNVHGNICNLEYYNKAINLANNNLKNPKFVIFSDDITWCKSNLGIDDPVYV DWNTGPYSYQDMYLMSNCKNNIIANSSFSWWGAWLNQNTEKKVFSPKKWVNDRNNVNIVPNGWIKIK	42
Gallionella sp. SCGC AAA018-N21	WP_018293379.1	517104561	protein [Gallionella sp. SCGC AAA018-N21]	28.15		MIIAHIIGGLGNQMFQYAAGRALSLARGVPFKLDISGFEGYDLHQGFELQRVFNCAAGIASEAEVRDSLGW QFSSPIRRIVARPSLAVLRRSTFVVEPHFHYWAGIKQVPDNCYLAGYWQSEQYFQSHAAVIRTDFAFKPPLSG QNSKLAMQIAQGNAVSLHIRRGDYANNPKTTATHGLCSLDYYRAAIQHIAERVQSPHFFIFSDDIAWVKSNL AINFPHQYVDHNQGTESYNDMRLMSLCQHNIIANSSFSWWGAWLNTNAHKIVIAPKQWFANTTIHVADLI PSSWERL	43
Azospira oryzae;Dechlorosoma suillum PS	YP_005026324.1	372486759	Glycosyl transferase family 11 [Dechlorosoma suillum PS]	28.04		MQSPACIAGARAWWVGYGMAEAMQPVVVGLSGGLGNQMFQYAAGRALAHRLGHPLSLDLSWFQGRG DRHFALAPFHIAASLERAWPRLPPAMQAQLSRLSRRWAPRIMGAPVFREPHFHYVPAFAALAAPVFLEGY WQSERYFRELREPLLQDFSLRQPLPASCQPILAAIGNSDAICVHVRRGDYLSNPVAAKVHGVCPVDYYQQAV AELSASLARPHCFVFSDDPEWVRGSLAFPCPMTVVDVNGPAEAHFDLALMAACQHFVIANSSLSWWGAW LGQAAGKRVIAPSRWFLTSDKDARDLLPPSWERR	44
Prevotella paludivivens	WP_018463017.1	517274199	protein [Prevotella paludivivens]	28		MKIVKIIGGLGNQMFQYALAMALNKNFTDEEVKLDIHCFNGYTKHQGFEIDRVFGNEFFLASYRDVAKVAY PYFNFQLWRIGSRIFPDRRHMISEDTSFKIMPEVITSHNYKYYDGYWQHEEYFKNIHDEILDAFKFPKFQDER NKALAERLSDSNSISIHIRRGDYLNDELFKGTCGIEYYKKAIEEINERTVPTLFCVFSNDIHWCKENIEPLLNGKE TIYVDWNTGSDNYRDMQLMTKCKHNIIANSSFSWWGAWLNNTKDKIVIAPRIWYNTKEKVSPVANSWIKL	45
Gramella forsetii;Gramella forsetii KT0803	YP_860609.1	120434923	alpha-1,2-fucosyltransferase [Gramella forsetii KT0803]	27.96		MSNKNPVIVEIMGGLGNQMFQFAVAKLLAEKNSSVLLVDTNFYKEISQNLKDFPRYFSLGIFDISYKMGTEN GMVNFKNLSFKNRVSRKLGLNYPKIFKEKSYRFDADLFNKKTPIYLKGYFQSYKYFIGVESKIRQWFEFPYENL GVGNEEIKSKILEKTSVSVHIRRGDYVENKKTKEFHGNCSLEYYKNAITYFLDIVKEFNIVFFSDDISWVRDEFK DLPNEKVFVTGNLHENSWKDMYLMSLCDHNIIANSSFSWWAAWLNNNSEKNVIAPKKWFADIDQEQKSL DLLPPSWIRM	46
Mariprofundus ferrooxydans;Mariprofundus ferrooxydans PV-1	WP_009849029.1	497534831	alpha-1,2-fucosyltransferase [Mariprofundus ferrooxydans]	27.92		MIIVQFTGGLGNQMFQYALGRRLSLLHDVELKFDLSFYQHDILRDFMLDRFQVNGQVATEKEIEAYTNTPIF ALDRPLLDRLVRWGLYRGIVSVSDEPPGKQALMVYNSRVLQAPRNTYVQGYWQSEKYFMPIRQKLLDDFSL VDKADQANGAMLEKIRQCHSVSLHVRRGDYVSNPLTNHSHGTCGLEYYEKAIALIGSKVDDPHFFVFSDDPE WTRDHILKCRFPMTYVTCNSADSCEWDMELMRHCRDHIIANSSFSWWGAWLNMNPDKVVVAPAAWFN NFSADTSDLIPDSWVRI	47

Bacillus cereus;Bacillus cereus VD107	WP_00217 4293.1	488102896	protein [Bacillus cereus]	27.91		MKIIQVSSGLGNQMFQYALYKKISLNDNDVFLDSSTSYMMYKNQHNGYELERIFHIKPRHAGKEIIDNLSDLD SELISRIRRKLFGAKKSMYVELKEFEYDPIIFEKKETYFKGYWQNYNYFKDIEQELRKDFVFTEKLDKRNEKLANE IRNKNSVSIHIRRGDYYLNKVYEEKFGNIANLEYYLKAINLVKKKIEDPKFYIFSDDIDWAQKNINLTNDVVYISH NCQGNESYKDMQLMSLCKHNIIANSTFSWWGAFLNNNDDKIVVAPKKWINIKGLEKVELFPENWITY	48
Firmicutes bacterium CAG:534	WP_02235 2106.1	547951299	protein [Firmicutes bacterium CAG:534]	27.81		MIIIRMTGGLGNQMFQYALYLKLRAMGKEVKMDDFTEYEGREARPLSLWAFGIEYDRASREELCRMTDGFL DPVSRIRRKLFGRKSLEYMEKDCNFDPEILNRDPAYLTGYFQSEKYFADIEEEVRQAFRFSERIWEGIPSQLLER IRSYEQQIKTTMAVSVHIRRGDYLQNEEAYGGICTERYYKTAIEYVKKRQQDASFFVFTNDPDYAGEWILKNF GQEKERFVLIEGTQEENGYLDLYLMSLCRHHILANSSFSWWGAYLNPSREKMVIVPHKWFGNQECRDIYME NMIRIAKEQS	49
Sideroxydans lithotrophicus;Sidero xydans lithotrophicus ES-1	YP_003525 501.1	291615344	glycosyl transferase family 11 [Sideroxydans lithotrophicus ES-1]	27.81		MVISNIIGGLGNQMFQYAAARALSLKLEVPLKLDISGFTNYALHQGFELDRIFGCKIEIASEADVHEILGWCQSA SGIRRVVSRPGMSIFRRKGFVVEPHFSYWNGIRKITGDCYLAGYWQSEKYFLDAAVEIRKDFSFKLPLDSHNA ELAEKIDQENAVSLHIRRGDYANNPLTAATHGLCSLDYYFKSIKHIAGQVRNPYFFVFSDDIAWVKDNLEIEFP SQYVDYNHGSMSFNDMRLMSLCKHHIIANSSFSWWGAWLNPNPEKVVIAPERWFANRTDVQDLLPPGW VKL	50
zeta proteobacterium SCGC AB-137-C09	WP_01828 1578.1	517092760	protein [zeta proteobacterium SCGC AB-137-C09]	27.81		MIVSQIIGGLGNQMFQYATGRALSHRLHDTFFLDLDGFSGYQLHQGFELSNVFQCEVNVATRSQMQALLG WRSFSSVRRLLMKRSLKWARGHRVMIEPHFHYWSRFAEINEGCYLSGYWQSERYFKPIENIIRQDFKFNHLL KGVNLDLAQQMTEVNSVSLHVRRGDYASDANTNHTHGLCPLDYYRDAILYIAQNTVAPSFFIFSDDIEWCRE HLKLSFPATYIDHNKGSNSYCDMQLMSLCHHHIIANSSFSWWGAWLNTRLDKIVIAPKQWFANGNRTDDLI PAEWLVM	51
Pedobacter heparinus;Pedobacte r heparinus DSM 2366	YP_003090 434.1	255530062	glycosyl transferase family protein [Pedobacter heparinus DSM 2366]	27.8		MKIIRFLGGLGNQMFQYAFYKSLQHRFPHVKADLQGYQEYTLHNGFELEHIFNIKVNSVSSFTSDLFYNKKW LYRKLRRILNLRNTYIEEKKLFSFDPSLLNNPKSAYYWGYWQNFQYFEHIADDLRKDFQFRAPLSAQNQEVLD QTKLSNSISLHIRRGDYIKDPLLGGLCGPEYYQTAINYITSKVNAARFFIFSDDIDWCIANLKLQDCSFISWNKG TSSYIDMQLMSSCKHHIVANSSFSWWAAWLNPNPDKIVIAPEKWTNDKDINVRMSFPQGWISL	52
Methylophilus methylotrophus	WP_01898 5060.1	517814852	protein [Methylophilus methylotrophus]	27.78		MFQYAMGLSLAENNQTPLKLDLSQFTDYKLHNGFELSKVFNCSAETASVTQIETLLGICKYSFIRRILKNTYLKN LRPAQYVVEPFFGYWDGVNFLGDNVYLEGYWQSQKYFIDYESTIRTHFTFKNILSGENLKLSDRIKGSNSVSL HIRRGDYVTNKNNAFIGTCSLIYYQNAIEYFSTKIADPIFFIFSDDITWAKSNLRLANEHYFVGHNQGEDSHFD MQLMSLCKHHIIANSSFSWWGAWLNPSKDKIIIAPKKWFASGLNDQDLVPKDWLRI	53

Rhodobacterales bacterium HTCC2255	WP_00803 3953.1	495309205	alpha-1,2-fucosyltransferase [Rhodobacterales bacterium HTCC2255]	27.7		MIYTRIRGGLGNQLFQYSAARSLADYLNVSLGLDTREFDENSPYKMSLNHFNIRADLNPPDLIKHKKDGKIAYI IDHIKGNQKKVYKEPFLSFDKNLFSNVDGTYLKGYWQSEKYFLRNRKNILSDINLIKKTDKFNTINLKEIKKSTSI SLHIRRGDYLSNESYNETHGICSLSYYTDAVEYIKNRLGENIKVFAFSDDPDWVLENLKLSVDIKIINNNTSANS FEDLRLMLNCDHNIIANSSFSWWGAWLNQNPEKIVISPKKWYNKKQLQNADIVPSSWLKY	54
Spirulina subsalsa	WP_01730 2658.1	515872075	protein [Spirulina subsalsa]	27.69		MAKIIARIRGGIGNQIFIYAAARRLELINNAELVLDSVSGFVHDLQYRQHYQLDHFHIPCRKATPAERFEPFSR VRRYLKRQLNQRLPFEQRRYVIQESIDFDPRLIEFKPRGTVHLEGYWQSEDYFKDIEATIRQDLQIQPPTDPTN LAIVQHIHQHTSVAVHIRFFDQPNADTMNNAPSDYYHRAVEAMETFVPGAHYYLFSDQPEAAKSRIPLPDE RVTLVNHNRGNKLAYADLWLMTQCQHFIIANSTFSWWGAWLAENQKKQVIAPGFEKRREGVSWWGFKGL LPKQWIKL	55
Vibrio cyclitrophicus	WP_01043 3911.1	498119755	glycosyl transferase family 11 [Vibrio cyclitrophicus]	27.67		MVIVKITGGLGNQLFQYATGSALANKLSCELVLDLSFYPTQTLRKYELAKFNINARVATDREIFLAGGGNDFFS KAIKKLGLTSIIFPEYIKEQESIKYVGKIDLCKSGAYLDGYWQNPLYFSQNKIELTREFLPRAQLSPSALAWKDHI SQASNSVSLHVRRGDYVENAHTNNIHGTCSLEYYQHAIEKIRSEVHNPVFFVFSDDIEWCKLNLSSLAEVEFV DNTTSAIDDLMLMRQCKHSIIANSTFSWWGAWLKLDGLVIAPRNWFSSASRNLKGIYPKEWHIL	56
Lachnospiraceae bacterium NK4A179	WP_02278 3177.1	551039510	protein [Lachnospiraceae bacterium NK4A179]	27.65		MRSVVDIKGGYGNQLFCYSFGYAVSKETGSELIIDTSMLDMNNVKDRNYQLGVLGITYDSHISYKYGKDFLSR KTGLNRLRKKSAIGFGTVVFKEKEQYVYDPSVFEIKRDTYFDGFWQSSRYFEKYSDDLRKMLKPKKISNAAEKL AEDARDCLSVSVHIRRGDYVSLGWTLKDDYYIKALDIIKERYGSEPVFFVFSDNKKYADDFFSAAGLKYRLMD YETDDAVRDDMFLMSRCSHNIMANSSYSWWGAFLNDNKDKTVICPETGVWGGDFYPEGWMKVTASSG K	57
uncultured bacterium	EKE02186.1	406980610	glycosyl transferase family protein [uncultured bacterium]	27.57		MIIVNLYGGLGNQMFQYALGRHLAEKNNTEIKLDISAFESYKLRKYELGNLNIIEKFALPEEISRLSTLPTGKIER FIRKTLRKPVKKPESYIKENITGGFNPKILDLQNNIYLEGYWQSEKYFIEIEDIIRKEFSFKFPATGKNKEILENILNI NSVSLHIRRGDYVTNPEVNQVHGVCSLDYYKSCVDFIEKKLESPYFYIFSDDIEWVKNNLQIQSQVYYVDHNT VDNAIEDMRLMFSCKHNILANSSFSWWGAWLNSNPDKMVITPRKWFNTTYDSNDLIPERWIKL	58

						MKIGIIYIVTGPYIKFWNEFYSSSQLYFCVEAEKNYEVFTDSSELASQRLPNVHMHLIEDKGWIVNWSSKSKFIC EIRNQLTSYDYIFYLNGNFKFISPIYCDEILPQAEHNYLTALSFSHYLTIHPDHYPYDRNKNCNAFIPYGQGKYYF QGGFYGGRTQEVLSLSEWCRDAIEADFNKKVIARFHDESYINRYLLTQHPKVLNDKYAFQDIWPYEGEYKAIV LNKEEVPEDNNLQEMKQNYIDPSLSFLLNDELKFIPISIVCLYGGLGNQMFGYAFYLYIRHISTQERKLLIDPAP CKRYGNHNGYELPSIFSKICQDIHISDETKNNIRKLRKGTSLSIEEVRASMPQSFKEKKQPIIFYSGCWQCVTYV ETVKDEIKKDFIFDESKLNEPSAQSMLRIIRRSNSVSVHIRRNDYLIGNNEFLYGGICTKSYYEKAISQMYTLLKDE PIFIYFTDDPEWVRSNFALDKSYLVDWNKNKDNWQDMYLMSACRHHIIANSSFSWWAAWLGGFPEKKVI APSTWLNGNMQTPDILPTEWIKIPITPDKKILDRICNHLILHSSYMKQLGLNSGKMGVVIFFHHYARYTQNPLYE NYAGDLFDELYEEIHKGISFSFLDGLCGIAWAVEYLVHEQFIEGNTDDSLAEIDFKVMQIDPRRFTDYSFETGL EGIACYVLSRLLSPRVCSSSLTLDSVYLKDLTEACRKVPVDKANYTRLFLNYIESKEVGYSFKDVLMQVLNHSEK AFGSDGLTWQTGLTMIMR	59
Butyrivibrio sp. AE3009	WP_02277 8576.1	*551034739*	*protein [Butyrivibrio sp. AE3009]*	27.46		MIIQLKGGLGNQMFQYALYKELRSRGKEVKIDDVTGFVDDELRTPVLQRFGIEYDRATREEVVKLTDSKMDI FSRIRRKLTGRKTCRIDEESGTFNPDILELDEAYLVGYWQSDKYFRNEDVIAQLRQEFQKRPQEIMTDSASWA TLQQIECCQSVSLHIRRTDYIDEEHNHIHNLCTEKYYKGAIDRIRSQYPSAVFFITDDKEWCRNHFRGPNFFV VELAEKENTDIAEMLLMSSCKHHICANSSFSWWSAWLNDSPEKMVIVPNKWINNRDMDDIYTDRMTKM AI	60
Bacteroides ovatus;Bacteroides ovatus CL02T12C04	WP_00431 7929.1	490447027	protein [Bacteroides ovatus]	27.43		MKQTIILSGGLGNQMFQYAFFLSMKAKGKSCSLDTTLFQTNKMHNGFELKSVFDIPDSPNQASALHSLLIKM LRRYKPKSILTIDEPYTFCPDALESKKSFLMGDWLSPKYFESIKDVVVNAYRFHNIGNKNVDTANEMHGNNS VSIHIRRGDYLKLPYYCVCNENYYRQAIEQIKDRVDNPIFVVFSNEPSWCDSFMKEFRVNFKIVNWNQGKDS YQDMYLMTQCKHNIIANSTFSWWGAWLNNNTDKIIVAPSKWFKNSEHNIICKEWLLIDTSK	61
Desulfospira joergensenii	WP_02266 4368.1	550911345	protein [Desulfospira joergensenii]	27.42		MGKKYVETVVNGGLGNQIFQFSAGFALSKRLNLDLVLNISTFDSCQKRNFELYTFPKIKNSFACIKDDDPGVFS RLRIPFLNFKEKIKQFHESHFFFDPAFFDIREPVRIEGYFQSYKYFEKYSDQLKDILLDIPLTSRLKTVLKVISSKKES VSVHIRRGDYISDQGINEVHGTLNEAYYLNSIKLMEKMFPESFFFLFTDDPHYVEENFKFLEDTSCIISDNDCLP YEDMYLMANCHHNIIANSSFSWWGAWLNQNPEKIVIAPRKWFSRKILMEKPVMDLLPDDWILL	62
Lachnospiraceae bacterium 10-1	EOS74299.1	507817890	protein C819_03052 [Lachnospiracea e bacterium 10-1]	27.39		MNIIRMTGGLGNQMFQYALFLRLKAQGKEVKFDDRTEYKGEEARPILLWAFGIDYPAAGEEEVNELTDGV MKFSHRLRRKLFGRKSKEYREKSCNFDQQILEKEPAYFTGYFQSERYFEEVKEQVRKAFQFSGKIWGSVSKEL EERIREYQTKIENKSQMPVSVHIRRGDYLENDEAYGGICTDAYYRKAIEMMEEKFPNTVFYIFSNDTGWAKQ WIDHFYKEKSRFIVIEGTTEDTGYLDLFLMSKCRAHIIANSSFSWWGAWLDPDQEKIVIAPSKWVNNQDMK DIYTREMIKISPKGEVR	63

Bacteroides dorei;Bacteroides dorei DSM 17855	WP_00783 2461.1	495107639	protein [Bacteroides dorei]	27.33		MVVVVGAGLANRMFQYAFALSLREKGLDVFIDEDSFIPRFIDFERTKLDSVFVNVNIQRCDKNSFPLVLRED RFYKLLKRISEYMSDNRYIERWNLDYLPYIHKKASTNCIFIGFWISYKYFQSSEDAVRKAFTFKPLDSIRNVELAT KLVTENSVAVHFRKNIDYLKNLPNTCPPSYYYEAINYIKKYVPNPKFYFFSDNWDWVRENIRGVEFTAVDWN PSSGIHSHCDMQLMSLCKHNIIANSTYSWWSAYLNENNNKIVVCPKDWYGGMVKKLDTIIPESWIIING	64
Firmicutes bacterium CAG:24	WP_02191 6201.1	547127421	protein [Firmicutes bacterium CAG:24]	27.33		MVIVKMSGGLGNQMFQYALYRKIQQTGKDVKLDLFSFQDKNAFRRFSLDIFPIEYQTANLEECRKLGECSYR PVDKIRRKMFGLKESYYQEDLDKGYQPEILEMNPVYLDGYWQCERYFQDIREKILEDYTFPKKISIESSRLQERI KNTESVSIHIRRGDYLDAANYKIYGNICTIEYYQSAISRMRKLCEKPNFYLFSNDPEWAKEIFGDTEDITIVEEDK ERPDYEDMFLMSRCKHNIIANSSFSWWAAWLNQNENKRVIAPVKWFNNHSVTDVICDDWIRIDGDHKGA	65
Clostridium hathewayi CAG:224	WP_02203 1822.1	547299420	epsH [Clostridium hathewayi CAG:224]	27.3		MIYVNIRGRLGNQLFIYAFARALQKSTNQQITLNYTSFRKHYNNTAMDLEQFNIPEDIMFENSKELPWFANT DGKVIRILRHYFPKLIRSILQKMNVLMWLGDEYVEVKVNKRRDIYIDGFWQSSRYFKSVYKELKNELIPKMEM SKEIKTMGDLINQKESVCVSVRRGDYVTVKKNRDVYYICDEKYLNTSIMRMVELVPNVTWFIFSDDADWVK DNIVFPGEVFYQPPRVTPLETLYLMKACKHFIISNSSFSWWGQYLSNNDNKIVIGPAKWYVDGRKTDIIEEE WIKIEV	66
Syntrophus aciditrophicus SB;Syntrophus aciditrophicus	YP_462663.1	85860461	alpha-1,2-fucosyltransferase [Syntrophus aciditrophicus SB]	27.3		MVIVRLTGGIGNQMFQYAAARRVSLVNNAPLFIDLGWFQETGSWTPRKYELDAFRIAGESASVGDIKDFKS RRQNAFFRRLPLFLKKRIFHTRQTHIIEKSYNFDPEILNLQGNVVLDGYWQSEKYFSDVDSEIRREFSFQTDPAE RNRKILERIASCESVSIHIRRGDYVTLPDANAFHGLCTPAYYRLAVEQISRKVVEPVFFVFSDDIAWARGNLKL GFETCFMDQNGPDRGDEDLRLMIACRHHIIANSSFSWWGAWLCSNPEKIVYAPRKWFNNGLDTPDNIPAS WIRI	67
Bacteroides caccae;Bacteroides caccae ATCC 43185	WP_00567 8148.1	491931393	protein [Bacteroides caccae]	27.27		MKIVKIIGGLGNQMFQYALYLSLKKKYPKEKIKIDISMFETYGLHNGFELKRIFDIDAEYASREEIRELSFYIKIYKL QRIFRKIFPVRKTECVEKYDFKFMSEVWSNCDRYYEGYWQNWEYFIEAQTEVRSTFTFKKELVGRNAKVIREI QYAKMPVSLHIRRGDYLHHKLFGGLCDLNYYKKAIDYVLNNNYDTPQFYLFSNDIEWCKTYILPLVQGYPFILVD WNSGVESYIDMQLMSCCRINIIANSSFSWWAAWLNDSSEKIVIAPKLWAHSPYGKEIQLKSWLLF	68
Butyrivibrio fibrisolvens	WP_02275 6304.1	551011888	protein [Butyrivibrio fibrisolvens]	27.24		MIIIEMSGGLGNQMFQYALYKSMLHKGLDVTIDKSIYRDVDHKEQVDLDRFPNVSYIEADRKLSSTLRGYGY NDSIIDKIRNKLNKSKRNLYHEDLDKGYQPEIFEFDNVYLNGYWQCERYFKDIKNEIKKDFIFPCTQSGDDKIK ALTIEMESCNSVSLHVRRGDYLKPGLIEIYGNICTEEYYKKSIEYIKERVDNPVFYIFSNDMAWVRDNFKSDDFR YYVNEDGAFDGMTDMYLMTRCRHNIVANSSFSWWGAWLNKHDDNIVICPNRWVNTHTVTDIICEDWIRI DV	69

Parabacteroides distasonis;Parabacteroides distasonis CL03T12C09	WP_005857874.1	492476819	protein [Parabacteroides distasonis]	27.24		MIVGGNDYCKVKVNIIGGLGNQMFQYAFALSLKEHFPKEEIRIDISHFNYLFVNKVGAANLHNGYELDKIFF NIELKKANAWQLMKLTWFIPNYLISRIARKILPVRNSEYIQNSSDCFFYDPMVYNKQGSCYYEGYWQAIGYYE SMRDKLCKIFQHPSPEGKNKQYIENMESSNSVGIHIRRGDYLLSDNFRGICEVDYYKRAIDKILQDGEKHVFYL FSNDQKWCEEYILPLLGNYEIIFVTGNIGRDSCWDMFLMTHCKDLIIANSSFSWWGAFLNKRGGRVVTPKR WMINRNIRYDLWMPEWIRI	70
Geobacter uraniireducens;Geobacter uraniireducens Rf4	YP_001230447.1	148263741	glycosyl transferase family protein [Geobacter uraniireducens Rf4]	27.21		MIIARLQGGLGNQMFQYAVGLHLALTHNVELKIDITMFSDYKWHTYSLRPFNIRESIATEEEIKALTDVKMDR PYKKIDNFLCRLLRKSQKISATHVKEKHFHYDPDILKLPDNVYLDGYWQSEKYFKEIENIIRQTFIIKNPQLGRDK ELACKILSTESVCLHIRRGNYVTDKTTNSVLGPCDLSYYSNCIKSLAGNNKDPHFFVFSNDHEWVSKNLKLDYP TIYVDHNNEDKDYEDLRLMSQCKHHIIANSTFSWWSAWLCSNPDKVIYAPQKWFRVDEYNTKDLLPSNWLI L	71
Lachnospiraceae bacterium A4	WP_016280341.1	511026085	protein [Lachnospiraceae bacterium A4]	27.21		MIIVKIYEGLGNQLFQYAFARSIQVNGKKVFLDTSGYTDQLFPLCRTSTRRYQLNCFNIRIKEVEKKNIEKYSFL IQEDMFGKLISKLAKLHLWMYKVTIQQNAQEYKESYLNTRGNVYYKGWFQNPKYFSSIRRLLLKEITPKYKIRI PAELRELLQEDNIVAVHCRRGDYQYIRNCLPVNYYKKAMAYMEKKLGVPRYLFFSDDLSWVKRQFGNKDN NYYIEDYGKFEDYQELMIMSRCRNFIIANSTFSWWAAWLCSYENKVVIMPRVWTYVGGQGVEMSDFPAD WIRI	72
Colwellia psychrerythraea;Colwellia psychrerythraea 34H	YP_270849.1	71282201	alpha-1,2-fucosyltransferase [Colwellia psychrerythraea 34H]	27.15		MKVVRVCGGFGNQLFQYAFYLAVKHKFNETTKLDIHDMASYELHNGYELERIFNLNENYCSAEEKLAVQSTK NIFTKLLKEIKKYTPFIPRTYIKEKKHLHFSYQEVDLGTKDTSIYYRGSWQNPQYFNSIASEIREKLTFPEFTEPKSL ALHQEISEHETVAVHIRRGDYLKHKALGGICDLPYYQNAIKEIEGLVEKPLFVIFSDDITWCRANINVEKVRFVD WNSGEQSFQDMHLMSLCTHNIIANSSFSWWGAWLNANPNKIVISPNKWIHYTDSMGIVPSEWIKVETSI	73
Roseobacter sp. MED193	WP_009810150.1	497495952	alpha-1,2-fucosyltransferase [Roseobacter sp. MED193]	26.96		MITSRLHGRLGNQMFQYAAARALAHRLGCGVALDGRGAELRGEGVLTRVFDLPLSAAPKLPPLKQHAPLRY GLWRGLGLAPRFRRERGLGYNTAFETWEDGCYLHGYWQSERYFEEISDLIRADFTFPDFSNRQNAEMAARI MEDNAISLHVRRGDYVALSAHVLCDQAYYEAALTRLLEGLSQDAPTVYVFSDDPDWAKANLPLPCKKVVVD FNGPETDFEDMRLMSLCKHNIIGNSSFSWWAAWLNANPQKRVAGPANWFGDPKLSNPDILPSQWLKVA P	74
Cesiribacter andamanensis;Cesiribacter andamanensis AMV16	WP_009197396.1	496488826	Glycosyl transferase family 11 [Cesiribacter andamanensis]	26.89		MMIVRLCGGLGNQLFQYAVGKQLSVKNNIPLKIDDSWLRLPDARKYRLQFFQIEEPLASPQEVERFVGPYES QSLYARLYRKVQNMILPRHRRRYFQESGFWAYEPELMRIRSQVFLEGFWQHHAYFTRLHPQVLEALQLREEY RQEPYAVLDQIREDAASVSLHIRRGDYVSDPYNLQFFGVMPLSYYQQAVAYMQEQLHAPTFYIFSDDLDWA RAHLKLQAPMVFVDIEGGRKEYLELEAAMRLCRHNILANSSFSWWGAYLNTNPHKRVIAPRQWVADPELKD KVQIQMPDWILL	75

Rhodopirellula sallentina;Rhodopirellula sallentina SM41	WP_008679055.1	495954476	glycosyl transferase family 11 [Rhodopirellula sallentina]	26.89		MIATRLIGGLGNQMFQYAYGFSLARRRSERLVLDVSAFESYDLHALAIDQFDISAARMTQAEFARIPGRYRG KSRWAERVANFAGGLQSCDKRPLRLRREKPFGFAEKYLAEGSDLYLDGYWQSERYFPGLQAELKKEFQLKRG LSDESSRVLDEIQSSMSVAMHVRRGDYVTNAETLRIYRRLDAEYYRKCLNDIRQRFSNLNVFVFSNDIQWCQ DHLDVGLKQRPVTHNDATTAIEDMFLMSQCDHSIIANSSFSWWAAFLGRSDAQRRVVYPDPWFNPGTLN GDSLGCANWVSESSISVSRPSRAA	76
Butyrivibrio sp. AD3002	WP_022762282.1	551018054	protein [Butyrivibrio sp. AD3002]	26.85		MIIRMMGGLGNQMFQYALYLQLKALGKEVKIDDVYGFRDDPQRDPVLEKMYGITYTKASDAEVVDITDSH LDIFSRIRRKLFGRKSHEYIEETGLFDPKVFEFETAYLNGYFQSDKYFPDKEVLAQLRREFVIKPDDVFTSADSW ELYRQIRETESVSIHVRRGDYLLPGTVETFGGICNDYYKRAIDRMVSEHPDAIFFVFTSDKEWCEQNVSGKK FRIVDTKEENDDAADLLLMSLCKHHILANSSYSWWSAWMNDSPEKTVIVPSKWLNTKPMDDIYTSRMTKI	77
Segetibacter koreensis	WP_018611017.1	517440157	protein [Segetibacter koreensis]	26.78		MVVVKLIGGMGNQMFQYAIGRHLAIKNKCPLYFDHIELENKNTANTPRNYELDIFNVQYQKNPFLQSNRFV AKVYHKLFSVQRIKEPDFTFHPHILNVQGNIHLNGYWQNENYFKEIEEIRQDFTFKTPANEKIESILQQIAATN SVSLHVRRGDYITLTEANQFHGVCSDTYYQKAIAKIKEAIPAPHLFVFSDDIHWVKQNMPFTEEHTFVDGNT GKNSFEDLRLMAACRHNILANSSFSWWAGWLNKNPEKMVIAPEKWFRAVHTDIVPPSWIKM	78
Amphritea japonica	WP_019621022.1	518450815	protein [Amphritea japonica]	26.76		MVIVRLIGGLGNQLFQYAYALSLLEQGYDVKLDASAFESYTLHGGFGLGEYAERLEVATTEEVDMVSRVGRIS TLLRKLQGKKSRRVIKESNFSYDEKMLTPEDSHYLVGYFQSELYFNKIRGELLSALDLKHKLSPYTEASVLAIADA SVSVSMHIRRGDYVSDKAAHNTHGVCSLDYYYAAVTFFEERYPDVDFYIFSDDIEWVKENLNVQRAHYISSEE KRFAGEDIYLMSQCDHNIVANSSFSWWGAWLNANEDKIVVAPRQWYADSNMQRLSKTLVPDTWIRL	79
Desulfovibrio vulgaris;Desulfovibrio vulgaris str. 'Miyazaki F'	YP_002437106.1	218887785	glycosyl transferase family protein [Desulfovibrio vulgaris str. 'Miyazaki F']	26.76		MRPVVVDIFGGLGNQMFQYAAAKSLAERLGVRLELDVSMFSGDPLRAFSLGEFAITDHVRGKSRSSLLVRFA RSLGFGSSSKCVEPFHYWEGINEIEAPVHMHGYWQSEKYFKAYEDLIRRTFSFSACEGVASSGKYAGVSSP MSVSVHLRRGDYKEQKNVVVHGILGREYYDAAYSIIKQGCPSACFFVFTDAINEAVDFFSHWNDVLFVDGN NQYQDMYLMSQCRHHIIANSSYSWWGAWLGAFSDGMTVAPKMWFAYDVLKEKSIKDLFPEDWIVL	80
Spirosoma spitsbergense	WP_020606886.1	522095677	protein [Spirosoma spitsbergense]	26.76		MISRITSGLGNQLFQYAVARHLSLKNKTSLYVDLSYYLYQYHDDTSRNFKLGNFSVPYHTLQQSPVEYVSKAT KLLPNRSLRPFFLFQKERQFHFDEQILQSRAGCVILEGFWQSEAYFRDNADTIRRDLQLSGTPSPEFNQYRELI RETPMSVSIHVRRSDYVNHPEFSQTFGFVGIDYYKRAIELARKELANPRFFVFSDDKEWSKTNLPLGEDSVFV QNTGLNGDVADLVLMSHCQHHIIANSSFSWWGAWLNPNAGKLVITPKNWYKNKPAWNTKDLLPPTWLS I	81

Lachnospiraceae bacterium 28-4	WP_01629 2012.1	511037988	protein [Lachnospiracea e bacterium 28-4]	26.73		MNIIRMSGGIGNQMFQYALYLKLVSLGKEVKFDDVTFYELDNARPIMLSVFGIDYPKASREELVELTDASMD FLSRVRKIFGRKSGEYHEASADYDETVLEKEHAYLCGCFQSERYFKDIEYEVREAYRFRNVVVPEEIRGGIETY ERQIGESLSVSIHIRRGDYLDAADVYGGICTDAYYNQAIRYMIKKYENPSFFVFTNDTFWAEKWCEVRERETG KRFTVIKGTDEETGYIDLMLMSRCKAHIIANSSFSWWGAWLDASPDKCVVAPVKWINTRECRDIYTEDMVR IGSNGKISFSNCSSL	82
Lachnospiraceae bacterium COE1	WP_01630 2211.1	511048325	protein [Lachnospiracea e bacterium COE1]	26.71		MVVVRIWEGLGNQLFQYAYARALSLRTKDRVYLDISEYEMSPKPVRKYELCHFKIKQPVINCGRIFPFVNKDS FYTKNNQYLRYFPAGLIKEEDCYFKRDFCELKGLLYLKGWFQSEKYFKEFESHIREEIYPRNKIKITRGLRKILNSD NTVSVHIRRGDFGKDHNILPIEYYENSKRVILERVDNPYFIIFSDDILWVKENMNFGLNCFYMDKEYSYKDYEE LMIMSRCKHNIIANSTFSWWGAWLNPSKDKIVIAPKKWFLYNPKKDFDIVPNDWIRV	83
Parabacteroides;Parabacteroides sp. 20_3;Parabacteroides distasonis CL09T03C24	WP_00586 7692.1	492502331	alpha-1,2-fucosyltransferase [Parabacteroides]	26.69	FutZ B	MKIVNIIGGLGNQMFQYAFAVALKAKYPNEEVFIDTQHYKNAFIKVYHGNNFYHNGYEIDKVFPNATLEPAR PKDLMKVSFYIPNQVLARAVRRIFPKRKTEFVTDQQPYVFIPEALSVIDDCYFDGYWMTPLYFDKYRDRILKEF TFRPFDTKENLELEPLLKQDNSVTVHIRRGDYVGSSSFGGICTLDYYRNAIREAYNLITSPEFFIFSNDQKWCM ENMRNEFGDAKVHFIAHNRGADSYRDMQLLSIARCNILANSSFSWWGAYLNQRKNCFIICPHKWHNTLEY SDLYLPTWIKI	84
Bacteroides sp. HPS0048	WP_00256 1428.1	488624717	protein [Bacteroides sp. HPS0048]	26.62		MFVIRLIGGVGNQLFQYTFGQFLRHKFGVEVCYDIVAFDTVDKGRNLELQLLDESLPLFETSNFFFSKYKSWK KRLFLYGFLLKKNNKYYTKYAPEEISLFTEKGLSYFDGWWQYPALLRDTINNMEDFFIPKQPIPVQIQKYYNEIL LNNFAVALHVRRGDYFTSKYAKTYAVCNVEYYTSAVNLMCEKLRSCKFYVFSDDLDWVKSNLILPSNTVYVK NYDINSYWYIYLMSLCRHIIISNSSFSWWGATLNRNPHKIVIAPKYWSTKKNNTLCDNSWIKI	85
Bacteroides thetaiotaomicron;Bacteroides thetaiotaomicron dnLKV9	WP_01626 7863.1	511013468	protein [Bacteroides thetaiotaomicron]	26.58		MKIINILGGLGNQMFEYAMYLALKNAHSEEEILCSTRSFCGYGLHNGYELGRIFGIQVKEASLLQLTKLAYPFF NYKSWQVMRHWLPVRKTMTRGAINIPFDYSQVMREDSVYYDGYWQNEKNFLHIREEILTAYTFPKFDDEK NQELADIIVKSNAVSCHIRRGDYLKEINMCVCTSSYYAHAISYMNEEINPNLYCVFSDDIEWCRNNICELMGE DKKIIFIDWNKGEKSFRDMQLMSLCKHNIIANSSFSWWGAWLNRNDKKIVVAPTRWIASEVKNDPLCDSW KRIE	86
Desulfovibrio alaskensis;Desulfovibrio alaskensis G20	YP_389367.1	78357918	glycosyl transferase [Desulfovibrio alaskensis G20]	26.56		MKFVGVWILGGLGNQMFQFAAAYALAKRMGGELRLDLSGFKKYPLRSYSLDLFTVDTPLWHGLPMSQRRF RIPMDAWTRGSRLPLVPSPPFVMAKEKNFAFSPIVYELQQSCYLYGYWQSYRYFQDVEDDIRTLFSLSRFATL ELAPVVAQLNEVESVAVHLRRGDYITDAASNAVHGVCGIDYYQRSMSLVRRSTTKPIFYIFSDEPEVAKKLFAT EDDVVVMPSRRQEEDLLLMSRCKHHIIANSSFSWWAAWLGKRASGLCIAPRYWFARPKLESTYLFDLIPDE WLLL	87

Prevotella oralis CC98A	ETD21592.1	564721540	protein HMPREF1199_00667 [Prevotella oralis CC98A]	26.56		MDIVVIFNGLGNQMSQYAFYLAKRKSGSRCHCIFHNVSTGFHNGSELDKVFGIKYEKGIFSKLLSKIYDIFDGIP KLRKKLNSLGIHIIREPRNYDYTASLLPRVSRWGLNYFVGGWHSEKYYTEILQEIKNTFSFKIDDEIKDIDFYEFYS LHHNDINSVSLHIRRGDYVGANEYSYFQFGGVATLEYYHKAIDEIYQRIENPTFYVFSDDIGWCKTTFLKNNFIF VDCNCGEKSWRDMFLISQCKHHIIANSTFSWWGAWLSIFHNSITICPKEFIKGVVIRDVYPDTWIKLSS	88
Comamonadaceae bacterium CR	YP_008680725.1	550990115	glycosyltransferase [Comamonadaceae bacterium CR]	26.54		MASKSKIIPRIFGGLGNQLFIYAAARRLALVNGAELALDDVSGFVRDHEYNRHYQLDHFNIPCRKATAAERLE PFARVRRYILKRKWNQRLPFEQRKYLVQESVDFDERLLTFKPRGTVYLEGYWQSEDYFKDIEPQIRADLRIHPP TDTVNQQMAERIRATNAVAVHVRFFDAPAQSALGVGGNNAPGDYYQRAIKVMQEQAPDAQYYIFSDQP QAARARIPLRDDHVTLVNHNQCDAVAYADLWLISQCQHFIIANSTFSWWGAWLGKTPESIVIAPGFEKREG AMFWGFRGLLPDRWVKL	89
Vibrio nigripulchritudo;Vibrio nigripulchritudo AM115;Vibrio nigripulchritudo FTn2;Vibrio nigripulchritudo Pon4;Vibrio nigripulchritudo SO65	WP_022596860.1	550250577	WblA protein [Vibrio nigripulchritudo]	26.51		MKDSRIVKLNGGLGNQMFQFALAFALKKKLNVAVKFDTELLDTNRTEFKLSLERFGLIVDKLTITEKFKYKGLE SCKYRKICNWISNFTTINIHKGYYKEKERGVYDRGIFDSNVKYIDGYWQNQEYFNDFRSELLNKFNLNGKVSN HAIQYLKEITSVQNSVSIHVRRGDYLLLDVYRNLTLDYYSEAIKLVRITNPDSKFFIFSNDINWCKSNFKSVDNAI FVDSTVDEFDDMFLMSKCKTNIIANSTFSWWAAWLNNNSGKIVYCPKKWRNDTTEVHKGLPEGWNIIDK	90
Sulfurospirillum deleyianum;Sulfurospirillum deleyianum DSM 6946	YP_003304837.1	268680406	glycosyl transferase family protein [Sulfurospirillum deleyianum DSM 6946]	26.48		MIIIKIMGGLTSQMHKYALGRVLSLKYNVPLKLDLTWFDNPKSDTPWEYQLDYFNINATIATVSEIKKLKGNN LFNRIARKIEKFFSIRIYKKSYINKSFISISDFHKLKSDIYLDGEWNGFKYFEDYQDTIKNELTLKRGSSINIQNTIKE LKSSDNSVFLHIRRGDYLSNKNAAAFHAKCSLDYYYKAIQIVKEKIDNPIFYIFSDDILWVKKNFVINESCRFME KNQNFEDLLLMSYCKHGITANSGFSLMAGWLNCNKDKMIIVPQTWVNDDRININILNSLEQDNFTIIR	91

Escherichia coli;Escherichia coli Jurua						MTFIVRLTGGLGNQMFQYALARSLAKKYNARLKLDISYYHNQPHKDTPRTFELNQLCIVDNILNSSSFSEKFLY	
18/11;Escherichia coli 180600;Escherichia coli P0304777.1;Escheric hia coli P0304777.2;Escheric hia coli P0304777.3;Escheric hia coli P0304777.4;Escheric hia coli P0304777.7;Escheric hia coli P0304777.9;Escheric hia coli P0304777.10;Escheri chia coli P0304777.11;Escheri chia coli P0304777.12;Escheri chia coli P0304777.13;Escheri chia coli P0304777.14;Escheri chia coli P0304777.15	WP_00158 1194.1	486318742	glycosyl transferase 11 family protein [Escherichia coli]	26.47		IYDKLRVKLSKKISLPYFRNIVTPVNFNCIDFAEDKDYYFLGHFQELSNIYSIDESLRSEFKPNQEIMNLAHQSKIY ELIKQSRGSVALHIRRGDYVTNKNAAEHHGVIGLSYYVNALSYLENVSEFFDVFVFSDDPEWARKNIKNSRNL FFCDEGNCRYSKKYSTIDMYLMSQCDHFIIANSTYSWWAAWLGNYPSKHVVAPARWNANNSPYPILQNW KAIHE	92
Firmicutes bacterium CAG:24	WP_02191 4998.1	547109632	protein [Firmicutes bacterium CAG:24]	26.44		MVGVQLSGGLGNQMFEYALYLKLKSMGKDVRIDDVTCYGAQEKQRVNQLSVFGVSYEHMTKQEYEQITD SSMSPLHRARRLLCGRKDLSYREASCNYDPEILRREPALLLGYFQTERYFADIKDQVREAFTFRNLTLTKESAA MEQQMKECESVSVHIRRGDYLTPANQALFGGICDLDYYHRAVAEIRKRKPDVKFFLFSNDMEWTKEHFCGS EFVPVEGNSEQAGEQDLYLMSCCKNHILANSSFSWWGAWLDNGKDKLVIAPEKWMNGRGCCDIYTDEMI RV	93
Amphritea japonica	WP_01962 2926.1	518452719	protein [Amphritea japonica]	26.42		MVKIKIIGGLGNQMFQYAAAKSLAVLNNTRVSANVSVFSNYKTHPLRLNKLNCDCEFDFTRDFRLVLSGFPLL GSAFSKKSMLLNHYVEKDLLFDSSFFDLDDNVLLSGYFQSEKYFSNIRELLIQEFSLDDRLTEAELAINNKIESCN SIAIHIRRGDYITDLSANNIHGICSEEYFEKALNYLDSINVLSDPTTTLFIFSDDILWCKDNLAFKYRTVFVEGSVD RPEVDIHLMSKCKHQVISNSTFSWWGAWLNTNLDKCVIAPLKWFNSLHDSTDIVPKQWMRL	94

Bacteroides salyersiae;Bacteroides salyersiae WAL 10018 = DSM 18765 = JCM 12988	WP_00592 3045.1	492689153	protein [Bacteroides salyersiae]	26.41		MKQTIIMSGGLGNQMFQYALYCSMREKGIRVKIDISLYEFNRMHNGYMLDYAFGLNISHNKINKYSVLWTR LIRSNRAPFLLFREDESRFCDDVFTTYKPYIDGCWIDERYFFNIKKKIISQFSFHNIDQKNLMVANMMKVCNS VSLHIRRGDYLSQSMYNICNESYYKSAIEYIISRVEDSKFFIFSDDPEWCKYFMEKFNVDYEIIQHNFGKDSYKD MYLMTQCKHNIIANSTFSWWGAWLNNNAGKNVVCPSVWINGRDFNPCLEEWYHI	95
Bacteroides fragilis;Bacteroides fragilis CL03T00C08;Bacteroides fragilis CL03T12C07	WP_00578 6334.1	492241663	protein [Bacteroides fragilis]	26.38		MDIILLHNGLGNQMSQYAFYLSKKNGIHTSYICLSNDHNGIELDKVFGVECQMGCKKIFLLFILRLLMSNRT GFLIRKVNLLFSKIKIKLITENLDYSFHPSFLSASPYCLAFWVGGWHHPQYYSEISSQIKEAFTFKRSLLDERNICI EKRMREPNSVCLHIRRGDYLTGINYELFGKVCNEQYYQKAIDYIEGKLSDICYYVFSNDMEWAKKILGKNAV FVDWNRGEESWKDMYLMSKCSNLIIPNSTFSWWAAWLCEHPVNIVCPKLFVYGDEQSDIYLDNWHKIE	96
Bacteroides nordii;Bacteroides nordii CL02T12C05	WP_00748 6843.1	494751435	protein [Bacteroides nordii]	26.37		MMGIEKTNMVIVRLWGGIGNQLFQYSFGEFLREKYQVDVIYDIASFGKSDKLRKLELSVVVPGIPVTTDISFSK YVGTKNRLLRFIYGLKNSFIEEKYFSDEQLFKYLSKRGDVYLQGYWQKTIYAETLRRKGSFFLSQEEPIVLHTIKA KIQEAEGAIALHVRRGDYFSSKHINTFGVCDAHYYEKAVDIMRGRVSNAMIFVFSDDLDWVRRYVNLPTNVI YVPNNYDIPQYWYIYLMSLCRHNIISNSSFSWWGAFLNMNTNKIVVSPSKWTLNSDKTIALDEWFKI	97
Butyrivibrio proteoclasticus;Butyrivibrio proteoclasticus B316	YP_003829 743.1	302669783	glycosyl transferase 11 [Butyrivibrio proteoclasticus B316]	26.37		MECSMIIIKFCGALGNQLFQYALYEKMRILGKDVKADISAFGDGNEKRFFYLDELGIEFNIASADEIAEYLNRKT IRFVPGFLQHRHYYFEKKPYVYNKKILSYDDCYLEGYWQNYRYFDDIKDELLKHMKFPCLPLEQKKLAEKMEN ENSVAVHVRMGDYLNLQDLYGGICDADYYDRAFSYIEGNISNPVYYGFSDDVDKASALLAKHKINWIDYNSE KGAIYDLILMSKCKNNIIANSSFSWWGAYLEYNNGKVVVSPNRWMNCFENSNIAYWGWISL	98
Prevotella ruminicola;Prevotella ruminicola 23	YP_003574 648.1	294674032	family 11 glycosyl transferase [Prevotella ruminicola 23]	26.33		MRIVKVLGGLGNQMFQFALYKALQKQYPEERVLLDLHCFNGYHKHRGFEIDSVFGVTYEKATLKEVASLAYP YPNYQCWRIGSRILPVRKTMLKEEPNFTLEPSALSLPDSTYYDGYWQHEEYFMHIREEILSTYAFPAFDDERN KTTAQLAASTNSCSIHIRRGDYLTDPLRKGTTNGNYYIAAIKEMQQEVKPEKWLVFSDDIAWCQQHLASTLD ATNTIYIDWNTGANSIHDMHLMALCRHHIIANSSFSWWGAWLSQQDGITIAPSNWMNLKDVCSPVPDN WIKI	99
Prevotella salivae;Prevotella salivae DSM 15606	WP_00713 5533.1	494223898	protein [Prevotella salivae]	26.33		MKIIKIIGGLGNQMFQYALAIALQQQYKDEEIRLDLNCFRGYNKHQGYLLDEIFGRRFRAASLQEVARLAWPY PHYQLWRVGSRVLPRRQTMVCEPADGSFSPDVLTLEGNRYYDGYWQDERYFKAYRKEIIEAFKFSPFVGDG NRHVENNMLRNERFASLHVRRGDYLNDALYQNTCGIDYYQRAISQMNAMANPSCYFIFSDDIAWCKTHIEPL CEGHRPYYIDWNKGKEAYRDMQLMALCKYHIIANSSFSWWGAWLNDAEDGITIAPQQWYSHGNKPSPAS ESWIKV	100

SUBSTITUTE SHEET (RULE 26)

Lachnospiraceae bacterium COE1	WP_01629 9568.1	511045640	protein [Lachnospiracea e bacterium COE1]	26.3		MNIVRISDGLGNQMFQYAYARKISILSRQRTYLDIRFINNEDLVKKGNHVQFRKKLGHRKYGLSHFNVSLQIA DLKMLSHWEYLIQSNCMQQLIYSLSMQDKWIWRYRHEEVNYDGMLSKVELLFPTYYQGYFFALKYYDDIKH ILQHDFSLKDKMKLLPELRDALYNRNTISLHVRRGDFLEINRDISGSEYYEKAVQMIGSKVESPIFLIFSDDIEW VKEHIRIPNDKIYVSGIGYEDYEELTIMKHCKHNIIANSTFSYWAAYLNSNKDKIVICPKHWREIIPKDWICI	101
Bacteroides dorei;Bacteroides dorei 5_1_36/D4	WP_00784 2931.1	495118115	alpha-1,2-fucosyltransfera se [Bacteroides dorei]	26.28		MIVVNVNAGLANQMFHYAFGRGLEAKGWNIYFDQTNFKPRKEWSFENVQLQDAFPNLGLKMMPEGKFK WICVNNTNKLSKGLHLAMINLHNIIGDEKYIFETTYGYDPDIEKEITKNCILKGFWQSEKYFAHCKDDIRKQFS FLPFDEEKNIVIMNKMVKENSVAIHLRKGADYLKSELMGKGLCGVEYYIKAIEYIKKNIDNPVFVFTDNPVW VKNNLPKFDYILVDWNEVAGKKNFRDMQLMSCAKHNIIANSTYSWWGAWLNPNPNKIVIGPAKFFNPIN NFFSSDIMCEDWVKI	102
Roseobacter sp. SK209-2-6	WP_00821 0047.1	495485361	alpha-1,2-fucosyltransfera se [Roseobacter sp. SK209-2-6]	26.28		MLSKDPGMITTRLHGRLGNQMFQYAAGRALAARLGVPLALDSRGAKLRGEGVLTRVFDLPLAQPLSLPPLK QDAPLRYAAWRLTGRTPRFRREQGLGYNPAFETWGDDSYLHGYWQSEAYFDSIADQIRQDFTFPEFSNSQ NREMAQRIAGSTAISLHVRRGDYVALAAHVLCDQAYYEAALTRILEGVEGSPTVYVFSDDPNWAKENLPLPC EKVVVDFNGPDTDFEDMRLMSLCQHNIIGNSSFSWWAAWLNTHNEKRVAGPAHWFGNPKLQNPDILPE SWLKISV	103
alpha proteobacterium SCGC AAA076-C03	WP_02005 6701.1	518900826	protein [alpha proteobacteriu m SCGC AAA076-C03]	26.26		MIYSRIRGGLGNQLFQYCVARSLADNLGTSLGLDVRDFNENSPYLMGLKHFNIRADFNPPGMIEHKKNGYF RYLIDVVNGKQKFVYKEPHLNFDKNIFSLPNSSYLKGYWQTEKYFIKNKVNILNDLKISHQSDKNKTISSKIAN NTSVSLHIRRGDYISNSAYNSTHGTCSLAYYTNAVNFLVNKIGGNFKVFAFSDDPEWVSSNLKLPVDICFVKN NSSEYNYEDLRLMSECNHNIIANSSFSWWGAWLNTNHNKTVITPCKWYADNSTKNADITPSNWIKI	104
Helicobacter bilis;Helicobacter bilis WiWa	WP_00408 7499.1	490188900	protein [Helicobacter bilis]	26.26		MGGGGQDLRLFELMLYNISLPLCFDYKTLVKYFYSNDKSLKYNFPLQYIRYATRSKYHKLYWLALKHYKYFYDE DPQGDNIVKMYLNNSLEKHAYPFGYFQNLIYFDEIDSIIREEFCLKIPLKPHNQALKEKIEKTENSVFLHVRLGD YLKMEATDGGYVRLGKTYYQSALEILKTRLGQPHIFIFSNDIEWCEKNLCNLLDFTGCHIEFVKANGEGNAAE EMELMRACKHAVIANSTFSWWASYLIDNPDKQIIMPTQVFNDTRRIPKSNMLAKKGYILIDPFWGMHSIV	105
Ralstonia sp. GA3-3	WP_01081 3809.1	498513378	glycosyl transferase family protein [Ralstonia sp. GA3-3]	26.26		MIVTRVIGGLGNQMFQYAAGRALARRLGVPLKIDSSGFADYPLHNYGLHHFALKAVQAGDREIPSGRAENR WAKALRRFGLGTELRVFRERGFAVDPEVMKLPDGTYLDGYWQSEYFAEMTQELRRDFQIATPPTSENAE WLARIGGDEGAVSIHVRRGDYVTNASANAVHGICSLDYYMRAARYVAENIGVKPTFYVFSDDPDWVAGNL HLGHETRYVRHNDSARNYEDLRLMSACRHHIIANSTFSWWGAWLNASEKKVVIAPAQWFRDEKYDTRDLL PPTWTKL	106

Bacteroides ovatus;Bacteroides ovatus 3_8_47FAA	WP_00430 3999.1	490431888	protein [Bacteroides ovatus]	26.25		MVVVYIAAGLANKMFQYAFSRGLMSHGLDVFLDQTSFQPEWSFEDIALEEVFPNIEIKNAPNNMFSLAYKK DLLSRIYRRMSAFFPNNRYLMERPFIYDELIYKKATNNCIFCGLWQTELYFNFCERDVRRNFVFTPFQDDQNI KLAEKMKNENSVAIHIRKGADYLKRNIWDGTCSVEYYNQAINYLKEHVSNPVFYLFTDNPEWVEENLKNIDY KLVDWNPVSGKQSYRDMQLMSCAKHNIIANSTYSWWGAWLNNNPQKIVVAPKIWFNPKIEKAPYIIPDR WIRL	107
Loktanella vestfoldensis	WP_01995 5906.1	518799952	protein [Loktanella vestfoldensis]	26.23		MIITKLIGGLGNQMFQYAAGRSLAMRHGVPLLLDITELRSYPKHQGYQFEDVFAGRFEIAGLIPLIRVLGRKAR KVPKTVAVVSPKWPPMGDHVWVRQRTHDYDAAFESIGADCYLSGFWQSEKYFATIAPQIRESFRFKEALTG ANAAIASRMKEAPSAAIHIRRGDYVTDKGAHAFHGLCAWDYYDAAIDHISRHEPDARFFVFSDDVVAAQER FANRQRAEVVAVNSGRHSYRDMMLMAQCKHQIIANSTFSWWAAWLNQNPDKIVVAPGTWFSGNDGQI KDIYCKDWIVI	108
Flavobacterium sp. ACAM 123	WP_01699 1189.1	515558304	protein [Flavobacterium sp. ACAM 123]	26.14		MDVVIIFNGLGNQMSQYAFYSQKKKINNSTYFVPFCKDHNGLELETVFSLNTKETLIQKSLYILFRILLTDRLKIV SDPLKWILNLFKCKIVKESFNYNYNPEYLKPSKGITFYYGGWHAEKYFAKENQQIKSVFEFTGDLGKINKEHVK DIASTNAVSLHVRRGDFMNEAANIGLFGGVSTKAYFEGAIKLIATKVDHPHFFVFSNDMDWVKENLSMDTVT YVTCNSGKDSWKDMCLMSLCQHNIIPNSTFSWWGAWLNKNPHKIVVCPSRFLNNDTYTDIYPDSWVKISD Y	109
Bacteroides fragilis;Bacteroides fragilis 3_1_12	WP_00577 9407.1	492219620	glycosyl transferase family 11 [Bacteroides fragilis]	26.1		MMKLVRMTGGLGNQMFIYAFYIQMKTIFPELRIDMSEMKKYKLHNGYELEDVFSIRPQTISAHKWLKRVIV YAFFSIIREKSEEELSIHKYTQHKRWPIVYYKGFFQSELFFKESSDTIRDIFSFNTENANFRTKEWAKIIKEQRSSV SIHIRRGDYTSAKNKIKYGNICTEEYYQKAISIILKKEPKAFFHIFSDDVEWTKAHLKIIHHLPHQYISWNKGPDS WQDMMLMSLCRHNIIANSSFSWWGAWLNAYKDKTVIAPSRWSNVKKTPHILPESWISIDI	110
Spirosoma panaciterrae	WP_02059 8002.1	522086793	protein [Spirosoma panaciterrae]	26.09		MIISRVTSGLGNQLFQYAAARSLSLRNKTAFYVDLSYYLYEYPDDTSRSFKLGFFSVPYRILQESPVEYLSKSTKL FPNRSLRPFFLFLKEKQFHFDPTILQAHAGCVIMEGFWQSECYFRDHAEIIRRELQLSKSPSSEFEGYHQQIQA TPVPVSVHVRRGDYVNHPEFSKTFGFIGLDYYKTAIRHLTKTIKNPHFYVFSDDKEWARANLPLPTDSVFVTN TGPSGDVADLVLMSTCHHHIIANSSFSWWGAWLNPNPDKLVITPKLWYKNQPTWNTKDLLPPTWVSL	111
uncultured bacterium	EKE06672.1	406985982	glycosyl transferase family 11 [uncultured bacterium]	26.09		MIITKLTGGLGNQLFQYAIGRNLIYINGSDLKLDVSEYDVSNKGNFRHYALDKFNTIQNFASKKETNNFKFGVF KKWLYKSGIVKNKNYFLEKKFNFDKEILKIKDNAFLQGYWQSEKYFIGIRDIILQEFSLKENIELKFGEILKEINES NSVSIHVRRGDYVKNPKNLSFHGVCSPKYYSESTSKIASLIEKPVFFVFSDDIEWVKENLNITFPVVYLSGIKNIK SYEELVLMSKCKHNIIANSSFSWWGAWLNTNQKKIVIAPKRWFNDVKLDTTDLIPENWIRI	112

Thermosynechococcus elongatus;Thermosynechococcus elongatus BP-1	NP_681784.1	22298537	alpha-1,2-fucosyltransferase [Thermosynechococcus elongatus BP-1]	26.07		MIIVHLCGGLGNQMFQYAAGLAAAHRIGSEVKFDTHWFDATCLHQGLELRRVFGLELPFSSKDLRKVLGA CVHPAVRRLLAGHFLHGLRPKSLVIQPHFHYWTGFEHLPDNVYLEGYWQSERYFSNIADIRQQFRFVEPLDP HNAALMDEMQSGVSVSLHIRRGDYFNNPQMRRVHGVDLSEYYPAAVATMIEKTNAERFYVFSDDPQWVL EHLKLPVSYTVVDHNRGAASYRDMQLMSACRHHIIANSTFSWWGAWLNPRPDKVVIAPRHWFNVDVFD TRDLYCPGWIVL	113
Colwellia piezophila	WP_019028421.1	517858213	protein [Colwellia piezophila]	26.03		MKIVKIAGGFGNQLFQYAFYLALDKKYAEQVCLDSLDMAKYRLHNGYELEGIFKLDARYCTEEQRIIVRKDNN IFTKLLSSLKKKLGNNKNYILEPKQEHFTFHEKSFGQANTPTYYKGYWQDVKYLENIEEELKSSLVFPEFELGKNI ELANFISSNSSVSLHVRRGDYVQHKAFGGICDLSYYQRAVEQINTLVKDPIFIVFSDDIQWCKDNLNLEKAKFV DWNIGENSFRDMQLMTLCKHNIIANSSFSWWGAWLNANDDKNVICPDKWVHYTSATGVLPSEWIKIKAS V	114
Prevotella maculosa	WP_019966794.1	518810840	protein [Prevotella maculosa]	26		MKIVKIIGGLGNQMFQYALAIALQERWKDEEIKLDLHGFNGYHKHQGYQLDMLFGHRFEAATLTDVAQLA WPYPHYQLWRVGSRLLPKRRSMLCEPSKGLLPSDVLKQKGSLYYDGYWQDERYFRAIRPQIMAAFKFPDFT DRRNLETEKRLKASEAVSIHVRRGDYLDDVLFQGTCNIAYYQRAIARLCQLKTPVFCIFSNDMAWCKVHIEPL LHGKEILYVDWNRGKESYRDLQLMTLCRHHIIANSSFSWWGAWLSKAEDGITIAPRHWYAHDAKPSPAAE RWIKV	115
Salmonella enterica;Salmonella enterica subsp. enterica serovar Worthington str. ATCC 9607;Salmonella enterica subsp. enterica serovar Cubana str. CFSAN001083;Salmonella enterica subsp. enterica serovar Cubana str. CFSAN002050;Salmonella enterica subsp. enterica serovar Cubana str. CVM42234	YP_008261369.1	525860034	fucosyltransferase [Salmonella enterica subsp. enterica serovar Cubana str. CFSAN002050]	25.99		MYSCLSGGLGNQMFQYAAAYILKQYFQSTTLVLDDSYYYSQPKRDTVRSLELNQFNISYDRFSFADEKEKIKLL RKFKRNPFPKQISEILSIALFGKYALSDRAFYTFETIKNIDKACLFSFYQDADLLNKHKQLILPLFELRDDLLDICKN LELYSLIQRSNNTTALHIRRGDYVTNQHAAKYHGVLDISYYNHAMEYVERERGKQNFIIFSDDVRWAQKAFL ENDNCYVINNSDYDFSAIDMYLMSLCKNNIIANSTYSWWGAWLNKYEDKLVISPKQWFLGNNETSLRNAS WITL	116

Bacteroides sp. 3_2_5	WP_00865 9600.1	495935021	protein [Bacteroides sp. 3_2_5]	25.94		MKKVIFSGGLGNQMFQYAFYLFLKKKGIKAVIDNSLYSEFKMHNGFELIKVFDIKESIYRTYFLKVHLIFIKLLMK IPPVRKLSCKDDVIPIGDIIEFDPPYARFYLGYWQSKKIVNYVIEELRAQFIFRNIPQMTIEKGDFLSSINSVSIHIR RGDYMGIPAYQGICNEIYYERAISFMKEHFLNPRFYVFSNDSIWAKLFLEKFDIDMEIIVTPPIYSYWDMYLMS RCRNHIIANSTFSWWAAVLNINKDKIVISPTIFKKDECIDIIFDDWVKISNI	117
Clostridium sp. CAG:510	WP_02212 4550.1	547662453	protein [Clostridium sp. CAG:510]	25.86		MIMLQMTGGMGNQMFTYALYRSLRQKGKEVCIEDFTHYDTPEKNCLQTVFHLDYRKADREVYQRLTDSEP DFLHKVKRKLTGRKEKIYQEKDAIIFEPEVFQTDDVYMIGYFQSGRYFEKAVFDLRKDFTFAWNTFPEKAKKLR EQMQAESSVSLHIRRGDYMNGKFASIYGNICTDAYYEAARRYMKEHFGDCRFYLFTDDAEWGRQQESEDT VYVDASEGAGAYVDMALMSCCRHHIIANSSFSWWGAWLDENPDKTVIAPAKWLNISEGKDIYAGLCNCLI DANGSVQGE	118
Rhodopirellula europaea;Rhodopirellula europaea SH398	WP_00866 5459.1	495940880	glycosyl transferase family protein [Rhodopirellula europaea]	25.86		MIVTRLIGGLGNQLFQYAFGHSLARSTYQTLLIDDSAFIDYRLHPLAIDHFTISASRLSDADRSRVPGKFLRTPV GRALDKVSRFVPGYQGVLPVRREKPFGFRESLLARESDLYLDGYWQSEKFFPGLRGSLREEFQLREQPSETTR RLSAQMKSENSVAIHVRRGDYVTSAKAKQIYRTLDADYYRRCLLDLAAHETDLKLYLFSNDVPWCESNLDVGI PFTPVQHTDGATAHEDLHLIAQCRHVVIANSTFSWWGAYLGQLHPTRRVYYPEPWFHPGTLDGSAMGCD DWISEASLEFQSSLKSSRRAA	119
uncultured bacterium	EKD23702.1	406873590	glycosyl transferase family protein [uncultured bacterium]	25.82		MIIVKLKGGMGNQMFQYAIGRNLATKLGTQLRLDLTFLLDRSPRKDFVFRDYDLDIFALDVAFAGPTDLKPFT QFRISHLTKIYNIFPRLLGRPYVISEPHFHFSEAILKSSDNVYLDGYWQSEKYFKEIENSIRDDFKFRQPLEGRAA EMAAQIKNEDRAVCLNVRRADFVTSKKAQEFHGFIGLDYYQKAVDLLVSKVGPLHLFIFSDDVDWCAANLK FNYPTTFVTKDYSGKKYEAYLQLMTLCRHYIIPNSTFAWWGAWLNSDPNKIVIAPKQWFKEASIDTTDIIPST WIRL	120
Bacillus cereus;Bacillus cereus AH1271	WP_00058 7678.1	446510160	protein [Bacillus cereus]	25.74		MIIVKLKGGLGNQMFQYALGKSLALYYDKPLKIDADYIKNNEGYVPRDFSLSKFNIELDLYQEADKERVGFILK NNFLAKKLRNYFLKKGKYKGKYIIENPDNLGLFKKELFENHNESMYIDGYWQSYLYFNNIRECLIKEFNLKPEYT KEMTEIMQRINETNSVAVHIRRGDYVKLGWTLDTTYYKKAIAEIVKNVDNPKFYVFSDDTDWVRSNLQELD NAVFIGECNLFDYQELWLMSTCKHNIISNSTFSWWGAWLNQNDHQVVVSPSAWINGMSVETTSLIPDSW KRV	121
Firmicutes bacterium CAG:95	WP_02249 9937.1	548309386	protein [Firmicutes bacterium CAG:95]	25.74		MDIIRMEGGLGNQLFQYALYRQLQFMGRTVKMDVTTEYGREHDRQQMLWAFDVHYEEATQEEINRLTD GFMDLPSRIRRKLTGRRTKKYAEADSNFDPQVLLKTPVYLTGYFQSEKYFKDVEGILHTELGFSDRIYDGISEVF ADQIRNYQKQIRETESVSLHVRRGDYLEHPEIYGMSCTMEYYQAGVRYIRERHPDAEIFVFTNDPVFTEKWL QENFLGDFTLIQGTSEETGYLDLMLMSQCKHQIMANSSFSWWGAWLNPNKDKIVVAPEPWFGDRNFHDI YTEEMIRISPRGEVKKHG	122

Prevotella oris;Prevotella oris F0302	WP_004374901.1	490508875	alpha-1,2-fucosyltransferase [Prevotella oris]	25.74		MIAATLFGGLGNQMFIYATVKALSLHYQVPMAFNLNHGFANDYKYHRKLELCKFNCQLPTAKWITFDYRGE LNIKRISRRIGRNLLCPNYQFVIEEPFHYEKRLFEFTNKNIFLEGYWQSPCYFENYSKEIRADFQLKVPLSKEML EEIYALKATGKTLVMLGIRRYQEVEGRDICTYKLCDKEYYIKAITYIQERIPNALFVVFTQDKEWATTHLPKGAE FYFVKDKQDEYATVADMFLMTQCTHAIISNSTFYWWGAWLQCTTKNHIVIAPDSFINSDCVCKEWIILKRNS LC	123
Escherichia coli	AAO37719.1	37528734	fucosyltransferase [Escherichia coli]	25.73		MYSCLSGGLGNQMFQYAAAYILQRKLKQRSLVLDDSYFLDCSNRDTRRRFELNQFNICYDRLTTSKEKKEISII RHVNRYRLPLFVTNSIFGVLLKKNYLPEAKFYEFLNNCKLQVKNGYCLFSYFQDATLIDSHRDMILPLFQINEDL LHLCNDLHIYKKVICENANTTSLHIRRGDYITNPHASKFHGVLPMDYYEKAIRYIEDVQGEQVIIVFSDDVKWA ENTFANQPNYYVVNNSECEYSAIDMFLMSKCKNNIIANSTYSWWGAWLNTFEDKIVVSPRKWFAGNNKSK LTMDSWINL	124

Leeia oryzae	WP_01815 0480.1	516890767	protein [Leeia oryzae]	25.71		125
					MIIVKIIGGLGNQMMQYAFAHACAKRLGVPFKLDITAFESYKLWPYGLHNFEITAPIASLEEIEHAKSMGVITE	
					TSFRFDDSLVSAVKDGMYLDGYWADYRYSESVVWGELKPVFTLMDPLTPEQQALAMNLSAPNAVALHVRR	
					GDYVTNPNCFLLPQQYYRDAIKLVLDQQPDAVFYCFSDDPDWVEAHLDIPAPKVVVRGQGIDNGFVDMIL	
					MSKARHRIVANSTFSIWASRLADQDGLTIVPSQFFRKDDPWLLQVYGEVLQPCYPPQWRVVDVTGDGKKE	
					AENTSTALLQIAGGDVRGRKLRIGVWGFYEEFYQNNYIFLNKNAPIGHELLKPFNQLYQYGQAHNLEFVTLDL	
					VADLSTLDAVLFFDAPNMRSPLVSSVMQLDIKKYLCLLECELIKPDNWQQSLHELFTRIFTWHDGLVDNHRYI	
					KVNYVTDLMPWIESAQSLTAPFEETARKGYLQKKLICNISGNKLVSHPFELYSKRIEVIRWFESHHPEHFDLYG	
					MGWSASDYPSYKGKIDDKLEVLKGYRFSLCYENAKELPGYITEKIIDCFKAGVVPVYSGAPNIADWIPDNCFID	
					SGKFPDTDALYTYLISMTEEVHADYLENIRQFFLGGKAYPFSADAFINTITRTIVQDCLFPHERTDVSVVVPNY	
					NHGNFVVSAITSALNQNVSVELLVLDNASTDDSWSQLQFFADYPQVRLIRNRWNIGVQHNWNHATWLAT	
					GRYVVMLSADDLLLPGHLEQAVKRLDENPASSLYYTPCLWINEHDQPLGTLNHPGHLESDYVGGRDEISDLL	
					KFDSYITPSAAVIRRETLNRIGSMNLHLKGAIDWDLWIRIAEISPAFIFRKQPGVCYRQHSGNNSVDFYASTAP	
					LEDHIRIVESIIDRKVAVKYLLKAKEEIIAHLDNRASSYPENQIQHLLSRINNIKDYLRKGAGPVISVIIPTKNRPGL	
					LANAIESLTYQTFKDFEVVIHNDGSCDIGGIVDFFSDQLQISYVRSSQSGGAAASRNRALKLAKGRIIAYLDDD	
					DVYLDSHLEKLVDAYKGRSEKFIYTNCEYLIQERKEGRLIEEGRERRYAGISYSRAQLLVSNFIPTPTWSHITKELI	
					DTIGDFDESLEILEDWDFLLRASKVTEFYQVNATTVEVRSDRSRDDHTLRANADKLLAYHQKIYAKHPVENESI	
					LANRQSLINSLSNRQDVTPKNENSYQGWVNARQPNELAVQILAERMMLQWSKQYQFMIVMVVKQSQQ	
					NLLANTIDSFCQQLYSGWKLIVISDFEAPDESFINNEVLGWLTLETVEDENLLTQAFNGVLAEVPSDWVTILPV	
					GTRLTSTALLKVGDRILLNGGACVIYTDHDYVSDDGMIKDPVLKPAFNLDMLRSQDYIGSSIFFRTDSLAAVG	
					GFASFPGARTYEACFRMLDNYGPQTIEHLPEPVMTFPENQPENSLRVAAMQLALEEEHLHRNNISASIEEGYV	
					TGTFLVQYHHSEQPFVSIIIPNKDKHEFLAPCIETLMKVTQYPAFEVIIVDNQSTDPDTLSYYEEIESRFANNVK	
					VICYDNPFNFSAQCNLGAESARGDFILFLNNDTEIVQANWLERMMMQHAQRNDVGVVGARLVFPETVTIQH	
					AGIVLGGKYPDEVFQFPYMNFPVDKDVSLNRTKVVQNYSAVTGACLLVRKSLYQQVGGMNEQNLAVLYGD	
					VDLCLRIRQLHKSVVWTPFSTLVHHTGKTLNSNSDHEKHLMMVIQTRQEREYMLSHWLDIIANDPYYHRLL	
					DKSECNGTIDCTHTPLWDDIPSARPRLQGMALVGGSGEYRVNNMPFRTLERSALAEIVLSNMTSKARLPSITEL	
					ARNAPDVFVVQNALADEFIRMLEMYKKYLPSVFRIQMLDDLLTEIPDASSFKRHFQKNWRDAKARLRKSLKF	
					CDRLIVSTEPLRTFAEDMIDDIIVPNMLERSVWGDLVSKRRAGKKPRVGWVGAQQHAGDLALMTDVVKA	
					TGHEVDWVFQGGMCPDDIRPYVAEVNTFEWLTYDKYPQGIAALNLDLAIAPLEINAFNEAKSNLRLLEYGALG	
					WPVICTDIYPYQTNNAPVCRVPNDASAWIEAIRSHIADLDATAQKGDQLRQWVHDHYMIEDHAQEWLSA	
					LTRPAGK	

Desulfovibrio africanus;Desulfovibrio africanus PCS	WP_005984173.1	492830219	Glycosyl transferase family 11 [Desulfovibrio africanus]	25.68		MFQYAAARALSLRHSASLAADLTWFSQQFDVQTTPREYALPAFRLNLPEADKRIVATFRLNPTELRIVSFLRH RICFPSRFLPRHTELSFDYWDGFRDILPPAYLDGYWQSERYFSDYPDIIRADFSMLSISEQAAWMSAKIASVQ DSISLHIRRGDYVNSLATRKAHGIDTERYYAKALEWIADRIGAATIFAFSDDPRWVRANFDFGKHKGIVVDGS WTAHEDMHLMSLCSHIIANSSFSWWGAWLSTSQGITIAPKSWFSNPHIWTPDVCPATWERIPC	126
Akkermansia muciniphila CAG:154	WP_022196965.1	547786341	glycosyl transferase family 11 [Akkermansia muciniphila CAG:154]	25.66		MAKGKIIVMRLFGGLGNQLFQYAFLFALSRQGGKARLETSSYEHDDKRVCELHHFRVSLPIEGGPPPWAFRK SRIPACLRSLFAAPKYPHFREEKRHGFDPGLAAPPRRHTYFKGYFQTEQYFLHCREQLCREFRLKTPLTPENARI LEDIRSCCSISLHIRRTDYLSNPYLSPPPLEYYLRSMAEMEGRLRAADAPQESLRYFIFSDDIEWARQNLRPALP HVHVDINDGGTGYFDLELMRNCRHHIIANSTFSWWAAWLNEHAEKIVIAPRIWFNREEGDRYHTDDALIP GSWLRI	127
Dysgonomonas mossii;Dysgonomonas mossii DSM 22836	WP_006843524.1	493897667	protein [Dysgonomonas mossii]	25.66		MKIVKLQGGLGNQMFQYAIARTLETNKKKDIFLDLSFLRMNNVSTDCFTARDFELSIFPHLRAKKLNSLQEKF LLSDRVRYKFIRKIANINFHKINQLENEIVGIPFGIKNVYLDGFFQSESYFKHIRFDLIKDFEFPELDTRNEALKKTI VNNNSVSIHIRRGDYVHLKNANTYHGVLSLEYYLNCIKRIGEETKEQLSFFIFSDDPEYASKSLSFLPNMQIVD WNLGKNSWKDMALMLACKHHIIANSSFSWWGAWLSERNGITYAPVKWFNNESQYNINNIIPSDWVII	128
Prevotella oris;Prevotella oris F0302	WP_004372410.1	490506359	glycosyl transferase, family 11 [Prevotella oris]	25.66		MDIVLIFNGLGNQMSQYAFYMSKKKFVPQSKCMYYKGASNNHNGSELDKLFDIKYSETFFCKLILLFKLYENI PRLRKYFHILGINIVSEPQNYDYNESILKKKTRFGITLYKGGWHSEKYFLANKQDVLNTFSFKIAKEDKNFIDLAK SIEEDTNSVSLHVRRGDYLNISPTDHYQFGGVATTNYYKNAVSYMLKRNKQAHFYIFSDDITWCKAEYKDLM PTFIECNKKNKSWRDMLLMSLCTNHINANSTFSWWGAWLSTKNGITICPTEFIHNVVTRDIYPETWVQL	129
Pseudogulbenkiania ferrooxidans 2002;Pseudogulbenkiania ferrooxidans	WP_008952440.1	496239055	glycosyl transferase family 11 [Pseudogulbenkiania ferrooxidans]	25.66		MIIVRLMGGMGNQLFQYATAFALSKRKSEPLVLDTRFFDHYTLHGGYKLDHFNISARILSKEEESLYPNWQA NLLLRYPIIDRAFKKWHVERQFTYQDRIYRMKRGQALLGYWQSELYFQEYRKEISAEFTLKEQSSVTAQQISV AMQGGNSVAVHIRRGDYLSNPSALRTHGICSLGYYNHAMSLLNERINDAQFYIFSDDIAWAKENIKIGKTSK NLIFIEGESVETDFWLMTQSKHHIIANSTFSWWGAWLANNTDEQLVICPSPWFDDKNLSETDLIPKSWIRLN KDLPV	130
Salmonella enterica	WP_000286641.1	446208786	protein [Salmonella enterica]	25.66		MYSCLSGGLGNQMFQYAAAYILKQYFQSTTLVLDDSYYYSQPKRDTVRSLELNQFNISYDRFSFTDEKEKIKLL RKFKRNPFPKKISEILSIALFGKYALSDSAFYAVETIKNIDKACLFSFYQDADLLNKHKQLILPLFELRDDLLDICKN LDVYPLILRNNNTTALHIRRGDYLTNQHAAKYHGVLDTSYYNNAMEYVERERGKQNFIIFSDDVKWAQKAFL GNENCYIVNNGDYDYSAIDMYLMSLCKNNIIANSTYSWWGAWLNKSEDKLVISPKQWFLGNNETSLRNAS WIIL	131
Carnobacterium sp. WN1359	YP_008718688.1	554649642	glycosyl transferase family 11 [Carnobacterium sp. WN1359]	25.59		MLIVKVYGGIGNQMFQYSFYKYLQKNNDDVFLDISDYKVHNHHNGFELIDVFNIEVKQADMSKFKGHVSSK NSIFYRLTSKLFKRNILGYSEFMDSNGISIVRNEKILTDHYFIGFWQDVLYLQSVEEEIKEAFNFKNVAIGKQNLE LISLSESVESVSVHIRKGDYANNSDLSDICDLEYYEEAMKIIDSKVSEPLYFIFSDDIEWCKQKFGKRDNLIYVD WNIAKKSYIDMLLMSKCKHNIIANSTFSWWGAWLNNNSKKIVICPKTWDRKKNENHLLLNDWIAI	132

Prevotella sp. CAG:1185	WP_02196 4668.1	547227670	protein [Prevotella sp. CAG:1185]	25.58		MMKIIVNMACGLANRMFQYSYYLFLMHKGYNVKVDFYNSAKLAHEKVAWNDIFPKARIEQASFSDILKSG GGSDVISKIRRKYLPFLSSVVNMPTAFDANLPVENKKLQYIIGVFQNANMVEAVEEDVKRCFKFQPFTDERNL KLQNEMQSCESVAIHVRKGKDYAQRIWYQNTCPIEYYQNAIRLISEKVNNPKLYVFTDNPEWVKEHFKDFPY TLYEGNPASGWGSHFDMQLMSVCKHNIISNSTYSWWSAFLNVHNEKIVIGPKVWFNPDSCSEFTSERILCK DWIAV	133
Selenomonas sp. CM52	WP_00964 5343.1	497331130	glycosyltransfer ase, family 11 [Selenomonas sp. CM52]	25.58		MFQYAMASSVARRAGEILKLDLSWIRQMEKKLSADDIYGIGIFSFDEKFSTSNEVQKFLPSGKFSAKIYRAVN RRMPFSWRRVLEEGGMGWHPQIMEIRRSVYFYMGYWQSEKYFSDFIQEIRKDFTFREEVRQSIEERRPIVE KIRKSDAVSLHIRRGDYAQNPALGEIFLSFTMQYYIDAARYISERVKTPVFFIFSDDIPWAKENLPLPYEVCYIDD NIQTNEREIGHKSKGYEDMYLMTQCQHNIIANSSFSWWGAWLNHNPNKIVVAPKKWCNGSFNYADIVPE CWVKL	134
Bacteroides nordii;Bacteroides nordii CL02T12C05	WP_00748 6621.1	494751213	protein [Bacteroides nordii]	25.57		MENVFIFNGLGNQMSQYALYLSKRNLGCKVRYAYNIRSLSDHNGFELDRVFGITYPNNLFNKCINIIYRLLFAN KYLFLVQKMIYVLRQMNVYSIKEKDNYDYDYKILTRHKGIVLYYGGWHSEKYFLSNADIIKDKFRFNISKLNSES LVLYHRLSSLNAVALHVRRGDYMAPEHYNVFGCVCGIEYYKAAIQYIQSQILNPVFIVFSNDIEWVKENITGIQ MIFVDFNKKENSWMDMCLMSCCEHNIISNSTFSWWGAWLNNNKNKIVVCPKYFMSNIDTKDIYPESWIKI	135
Parabacteroides merdae;Parabacteroides merdae ATCC 43184;Parabacteroides merdae CL09T00C40	WP_00563 5503.1	491855386	protein [Parabacteroides merdae]	25.54		MKKKDIILRVWGGVGNQLFIYAFAKVLSLITDCKVTLDIRTGFANDGYKRVVRLGDFSISLLPALRFYTLLSFAQ RKIMPYIRHLLAYKFDFFEEDQKYPLETLDSFFKIYSDKNLYLQGYWQYFDFSSYRDVLLKDLRFEVEINNTYLYY SDLIEKSNAVAIHFRRIQYEPVISIDYYKKAIKYISENVENPTFFIFSDDINWCRENLSINGICFFVENFKDELYELK LMSQCNHFIIANSTFSWWGAWLSVNADKKVIMPDGYTDVSMNGSIVHI	136
Butyrivibrio sp. NC2007	WP_02276 8139.1	551024004	protein [Butyrivibrio sp. NC2007]	25.51		MIIIQLKGGLGNQMFQYALYKELKHRGRDVKIDDESGFIGDKLRVPVLDRFGVEYDRATKDEVIALTDSKMDI FSRIRRKLTGRKTFRIDEMEGIFDPKILETENAYLVGYWQSEKYFTSPEVIEQIQEAFGKRPQEIMHDSVSWST LQQIECCESVSIHVRRTDYMDAEHIKIHNLCSEKYYKNAISKIREEHPNAVFFIFTDDKEWCKEHFKGPKFITVE LQEGEFTDVADMLLMSRCKHHIIANSSFSWWSAWLNDSPEKIVIAPSKWINNKKMDDIYTERMTKVAI	137
Bacteroides ovatus;Bacteroides ovatus ATCC 8483;Bacteroides ovatus CL02T12C04	WP_00430 2233.1	490430100	protein [Bacteroides ovatus]	25.5		MIVVYSNAGLANRMFHYALYKALEVKGIDVYFDEKSYYPEWSFETTTLMDVFPNIQYRESLQFKRASKKTFL DKIVIHCSNLFGGRYYVNYRFKYDDKLFTKLETNQDLCLIGLWQSEKYFMDVRQEIQKCFQYRSFVDDKNVKT AQQMLSENSVAIHVRKGADYQQNRIWKNTCTIDYYRLAIDYIRMHVQNPVFYVFTDNKDWVIENFTDLDY TLCDWNPTSGKQNYLDMQLMSCAKHNVIANSTYSWWGAWLNENSDKIVIAPKRWFNKIVTPDILPEQWI KI	138

			glycosyl transferase family protein [Mesotoga prima MesG1.Ag.4.2]			
Mesotoga prima MesG1.Ag.4.2;Mesotoga prima	YP_006346113.1	389844033		25.5	MRVVWFGGGLGNQMFQYGLYCFLKKNNQEVKADCTQYSTTPMNNGFELERLFNLDIAHANLDVISKLTG GNRLSPRKVIWKLFRKPKVYFEEKIPFSFDPDVLKGNNRYLKGYWQNMNYLEPCAKELRDVFTFPAFSSDNN KRLADEIAKVEAVGVHFRRGDFLKSSNLGLFGGICSDQYYLRAIQTMENTVVEPVFYVFCDDPQWAKNSFSD ARFTVIDWNIGSNSYRDMQLMSLCKHNIIANSTFSWWAAWLNRNPNRTVIAPERMVNRDLDFSGIFPND WIRLQG	139
Clostridium sp. KLE 1755	WP_021639228.1	545399562	glycosyltransferase, family 11 [Clostridium sp. KLE 1755]	25.49	MIVLKLQGGLGNQMFEYAFARTIQEQKKDKKLILDTSDFQYDKQREYSLGHFILNENIEIDSSGKFNLWYDQR KNPLLKVGFKFWPKFQFQTLKLFGIYVWDYAKYIPVDVSKKHKNILLHGLWQSDKYFSQISEIIRKEFAVKDEP SQGNKAWLERISSANAVCVHIRRGDFLAKGSVLLTCSNSYYLKAMEIISKKVNEPEFFIFSDDIEDVKKIFEPG YQITLVNQSNPDVEELRLMSKCKHFIIANSTFSWWSSLLSENEDKVIVAPRLWYSDGRDTSALMRDEWIIIDN E	140
Bacteroides plebeius CAG:211	WP_022052991.1	547321746	glycosyltransferase family 11 [Bacteroides plebeius CAG:211]	25.42	MDLVTLSGGLGNQMFQFAFYWALKKRGKKVFLYKNKLAAKEHNGYELQTLFGVEEKCVDGLWMTRLLGC PLLGKILKHILFPHKIRERVLYNYSIYLPLFERNGLHWVGYWQSEKYFQDVADDIRRIFCFDHLSLNPATSAALK CMSEQVAVSVHIRRGDYYLPCNVATYGGLCTVEYYENAIRYVKERYPQAVFYVFSDDLDWVRENIPSAGKM VFVDWNRGKDSWQDMFLMSKCKHHNILANSSFSWWGAWLNTHPEKLVIAPERWANCPAPDALPDGWV RIEGVSRR	141
Treponema lecithinolyticum;Treponema lecithinolyticum ATCC 700332	WP_021686002.1	545448980	glycosyltransferase, family 11 [Treponema lecithinolyticum]	25.4	MAIKIVKISGGLGNQMFCYAFACALQKCGHKVYYDTSIYRKATVHSGIDFCHNGLETERLFGIKFDEADTAD VRRLSTSAEGLLNRIRRKYFTKKTHYIDTVFKYTPELLSDKNDCYLEGYWQTEKYFLPIEKDIRRLFTFRPTLSEKS AAVQSALQAQQAAVLSASIHVRRGDFLNTKTLNVCTETYYNNAIKYAVKKHIAVSRFYIFSDDIPWCREHLCFC NAHAVFIDWNTGNDSWQDMALMSMCRCNIIANSSFSWWAAWLNNASDKTVLAPAIWNRRQLEYVDRY YGYDYSDIVPESWIRIPID	142
Bacteroides eggerthii;Bacteroides eggerthii DSM 20697	WP_004291980.1	490419682	glycosyl transferase [Bacteroides eggerthii]	25.34	MRLIKMTGGLGNQMFIYAFYLRMKKRHTNTRIDLSDMMHYNVHHGYEMHRVFNLPKTEFCINQPLKKVIE FLFFKKIYERKQDPSSLLPFDKKYLWPLLYFKGFYQSERFFADMENDIRIAFTFNSDLFNEKTQAMLTQIKHNE HAVSLHIRRGDYLEPKHWKTTGSVCQLPYYLNAITEMNKRIEQPSYYVFSDDIAWVKENLPLPQAVFIDWNK GAESWQDMMLMSHCRHHIICNSTFSWWGAWLNPRENKTVIMPERWFQHCDTPNIYPDGWIKVPVN	143
Bacteroides stercoris;Bacteroides stercoris ATCC 43183	WP_005656005.1	491891563	glycosyl transferase [Bacteroides stercoris]	25.34	MRFIKMTGGLGNQMFIYAFYMRMKKHYSNTRIDLSDMVHYKAHNGYEMHRVFNLPPIEFRINQPLKKVIEF LFFKKIYERKQVPSSLVPYDKKYFWPLLYFKGFYQSERFFADMADDIRKAFTFNPRLSNRKTKEMSEQIDHDE NAVSIHVRRGDYLEPKYWKTTGCVCQLPYYLNAIAEMNKRISQPSYYVFSDDIAWVKENLPLPKAFFIDWNK GAESWQDMMLMSRCRHHIICNSTFSWWGAWLNPRENKTVIMPERWFRHCETPDICPDKWIKVPINQPD SIQ	144

Butyrivibrio proteoclasticus;Butyrivibrio proteoclasticus B316	YP_003831842.1	302671882	glycosyl transferase 11 [Butyrivibrio proteoclasticus B316]	25.34		MIIIQLKGGMGNQMFQYALYRQLKKLGREVKIDDETGFVDDELRIPVLQRFGISYDKATREEIVKLTDSKMDI FSRIRRKLTGRKTFRIDEESGIFDPRILEVEDAYLVGVWQSDKYFANEEVEKEIREAFEKRPQEVMQDSVSWTI LQQIECCESVSLHIRRTDYIDEEHIHHIHNICTEKYYKSAIDEVRNQYPSAVFFIFTDDKDWCRQHFRGPNFFVVD LDEDTNTDIAEMTLMSRCKHHILANSSFSWWAAWLNDNPGKIVIAPSKWINNRKMDDIYTARMKKIAI	145
Roseobacter sp. GAI101	WP_008228724.1	495504071	alpha-1,2-fucosyltransferase [Roseobacter sp. GAI101]	25.34		MSPIVHFPSDRLLRYEHLNSLWKTAMIYTRLLARLGNQMFQYAAGRGLAARLGVDFTVDSRRAVHKGDGV LTRVFDLDWAAPENMPPAQHERPLAYYAWRGLRRDPKIYRENGLGYNAAFETLPDNTYLHGYWQCERYFA HIADDIRAAFVPRHPMSAQNADMARRIASGPSVSLHVRRGDYLTVGAHGICDQTYYDAALAAVMQGLPSP TVYVFSDDPQWAKDNLPLTFEKVVVDFNGPDSDYEDMRLMSLCQHNVIANSSFSWWGAWLNANPQKRV AGPANWFSNPKLSNPDILPSRWIRI	146
Thalassobacter arenae;Thalassobacter arenae DSM 19593	WP_021099615.1	544666256	alpha-1,2-fucosyltransferase, [Thalassobacter arenae]	25.34		MGQDMIYSRIFGGLGNQLFQYATARAVSLRQGVELVLDTRLAPPGSHWAFGLDHFNISARIAEPSELPPSKD NFFKYVMWRAFGHDPAFMRERGLGYQSRIAQAPDGTYLHGYFQSERYFADVLDHLENELRIVTPPDTRNA EYADRIASAGHTVSLHVRRGDYVETSKSNSTHATCDEAYYLRALARLSEGKSDLKVFVFSDDPEWVRDNLKLP YDTTPVGIHNGPDDKPHEDLRLMSCCSDHVIANSTFSWWGGWLDRRPEARVVGPAKWFNNPKILVNPDILPE RWIAI	147
Prevotella oris;Prevotella oris C735	WP_004377401.1	490511493	protein [Prevotella oris]	25.33		MKIIKIIGGLGNQMFQYALAVALQKKWKDEEIKLDLHGFNGYHKHQGYQLDEIFGHREKAASLKEVAQLAW PYPHYQLWRVGSRLLPKRKTMVCESADCRFQSDLLNLEGSLYYDGYWQDERYFKAFRTEIIEAFKFTPLVGDS NRKVENMLKEGRFASLHVRRGDYLKEPLFQSTCDIAYYQRAISRLNQMADPYCYLIFSNDIAWCKTHIEPLCD GRRTHYVDWNHGKESYRDMQLMTFCKHHIIANSSFSWWGAWLSTANDGITIAPHQWYANDRKPSPAAE AWLKL	148
Prevotella oulorum;Prevotella oulorum F0390	WP_004380180.1	490514606	protein [Prevotella oulorum]	25.33		MKIVRIIGGLGNQMFQYALALALKQQQENEEVKLDLSAFRGYKKHGGFQLVQCFGTTLPAATWQEVAQLA WYYPHYQLWRLGHRVLPVCKTMLKEPDNGAFLPEVLQRKGDAYYEGCWQDERYFSHYRPAILQAFTFPTF TNPRNLAMQQQQNTTESVAIHVRRGDYLHDALFRNTCGLAYFQRAITCILQHVAHPVFYVFSDDMAWCRQ HIQPLLQTNEAVFVDWNHGKASICDLHLMTLCRHHIIANSSFSWWGAWLSPHQAGWIIAPKQWYAHEEK MSPAAERWLKL	149
Spirosoma panaciterrae	WP_020596174.1	522084965	protein [Spirosoma panaciterrae]	25.33		MNRRVAVQLKGGLGNQLFQYALGRRLSLQLEAELLFDCSVLENRIPVTNFTFRSFDLDMFRIAGRVATPSDL PLFPKSASIRSPWPHLVQLARLWKQGYSYVYERGFAYNPKMLRQLSDRVYLNGYWQSYRYFEDIAATLRAD CSFPDPLPDSAVGLAGQINATNSICLHIRRTDFLQVPLHQVSNADYVGRAIAYMAERVNDPHFFVFSDDIAW CQTNLRLSYPVVFVPNELAGPKNSLHFRLMRYCKHFITANSTFSWWAAWLSEPSDGKVIVTPQTWFSDSRSI DDLIPANWIRL	150

Butyrivibrio proteoclasticus;Butyrivibrio proteoclasticus B316	YP_003829826.1	302669866	glycosyl transferase 11 [Butyrivibrio proteoclasticus B316]	25.26		MNYVEVKGGLGNQLFQYTFYKYLEKKSGHKVLLHTDFFKNIDSFEEATKRKLGLDRFDCDFVAVSGFISCEKL VKESDYKDSMLSQDEVFYSGYWQNKRFFLEVMDDIRKDLLLKDENIQDEVKELAKELRAVDSVAIHFRRGDY LSEQNKKIFTSLSVDYYQKAIAQLAERNGADLKGYIFTDEPEYVSGIIDQLGSIDIKLMPVREDYEDLYLMSCAR HHIIANSSFSWWGAALGDTESGITIAPAKWYVDGRTPDLYLRNWISI	151
Butyrivibrio sp. XPD2006	WP_022765786.1	551021623	protein [Butyrivibrio sp. XPD2006]	25.26		MIIIQLKGGLGNQMFQYALYKELKHRGREVKIDDVSGFVNDKLRVPVLDRFGVEYERATREEVVELTDSRMD IFSRIRRKLTGRKTYRIDEMEGIFDPAILETENAYLVGYWQSEKYFTSPEVIEQIQEAFGKRPQEIMHDSVSWST LQQIECCESVSIHVRRTDYVDAEHIKIHNLCSEKYYKNAIGKIREKHPNAVFFIFTDDKEWCKDHFKGPNFITVE LQEGEFTDVADMLLMSRCKHHIIANSSFSWWSAWLNDSPEKMVIAPSKWINNKKMDDIYTERMTRVAI	152
Bacteroides sp. 1_1_6	WP_008766093.1	496041586	protein [Bacteroides sp. 1_1_6]	25.24		MKIVNITGGLGNQMFQYAFAMALKYRNPQEEVFVDIQHYNTIFEKKFKFKGINLHNGYEIDKVFPKAKLPVAGV RQLMKFSYWIPNYILSRLGRKFLPIRKKEYIPPYSMNYSYDEKALNWKGDGYFEGYWQSYNHFGDIKEELQK VYAHPKPNQYNAALISNLESCNSVGIHVRRGDYLAEPEFRGICGLDYYEKGIKEILSDEKKYVFFIFSNDMQWC QENIAPLVGDNRIVFISGNKGKDSCWDMFLMTHCKDLIIANSSFSWWGAFLNKKVDRVICPKPWLNRDCNI DIYNPSWILVPCYSEDW	153
Bacteroides fragilis;Bacteroides fragilis YCH46	YP_099857.1	53713865	alpha-1,2-fucosyltransferase [Bacteroides fragilis YCH46]	25.17		MKIVTFQGGLGNQLFQYVFYLWLDMRCDKDNIYGYYPKKGLRAHNGLEIEKVFEVKLPNSSLSTDLIVKSIKLI NKIFKNRQYISTDGRLDVNGVLFEGFWQDKYFWEDVDIVLNFRWPLKLDVTNSFIMTKIQANNSISIHIRRG DYLLPKYRNIYGDICNEEYYQKAIEYILKCVDDPFFVFSDDIDWAKSIINVSNVTFVNNNKGKDSYIDMFLMS LCHHNIIANSTFSWWAAQLNKHSDKIMIAPIRWFKSLFKDPNIFTESWIRI	154
Bacteroides sp. 9_1_42FAA	WP_008671843.1	495947264	alpha-1,2-fucosyltransferase [Bacteroides sp. 9_1_42FAA]	25.17		MIKIVSFSGGLGNQLFQYLLVVYLRECGHQVYGYYNRKWLIGHNGLEVNNVFDIYLPKTNFIVNALVKVIRVL RCLGFKKYVATDTYNNPIAIFYDGYWQDQKYFNIIDSKLSFKKFDLSAENKSILSKISNISVALHIRCGDYLSSS NVEIYGGVCTKEYYEKALELVCKIKNVMFFVFSDDIEYAKLLLNLPNAIYVNANVGNSSFIDMYLMANCKVNV IANSTFSYWAARLNQDNILTIYPKKWYNSKYAVPDIFPSEWVGV	155
Coraliomargarita sp. CAG:312	WP_022477844.1	548260617	glycosyl transferase family 11 [Coraliomargarita sp. CAG:312]	25.17		MIIVKVQGGLGNQMFQYAFGRALSEKHSQDLYLDCSEYLRPSCKREYGLDHFNIRAKKASCGDVKSMVTPH FALRKKLKKIFAVPYSLSPTHILERNFNFQPSILEFNCGYFDGFWQTQKYFSGISDIVRKDLTFKDAVKYSGGET FAKITSLNSVSLHIRRGDYVKVKRTRKRFSVIRAGYFKRAVEYMRSKLDTPHFFIFTDDPKWVSENFPAGEDYT LVSSSGMYEDLFLMAQCRHNIIFNSSFSWWGAWLNGNPGKIVVAPDMWFTPHYKLDYSDVVPEEWIKLN TGYFESKEF	156

Pseudorhodobacter ferrugineus	WP_022705649.1	550957292	alpha-1,2-fucosyltransferase [Pseudorhodobacter ferrugineus]	25.17		MIVMQIKGGLGNQMFQYAAGRALSLQTGMPLHLDLRYYRREREHGYGLGAFNIEASPLDESLLPPLPRESPL AWLIWRLGRRGPNLVRENGMGFNPTLSNVTKPAWITGYFQSERYFAAHAATIRAELTPVAAPDLVNARWL AEIAAEPRAVSLHVRRGDYVRDAKAAAKHGSCTPAYYERAIAHITARMGTAPVVYAFSDDPAWVRENLRLP AEIRVVPGHNDTAGNVEDLRLMSACRHHIVANSSFSWWGAWLNPRADKIVASPARWFADPAFTNPDIWPE AWARIEG	157
Escherichia coli;Escherichia coli O127:H6 str. E2348/69	YP_002329683.1	215487252	fucosyltransferase [Escherichia coli O127:H6 str. E2348/69]	25.16		MRMYCCLSGGLGNQMFQYAAAYILKQHFPDTILVLDDSYYFNQPQKDTIRHLELDQFKIIFDRFSSKDEKVKIN RLRKHKKIPLLNSFLQFTAIKLCNKYSLNDASYYNPESIKNIDVACLFSFYQDSKILNEHRDLILPLFEIRDDLRVL CHNLQIYSLITDSKNITSIHVRRGDYVNNKHAAKFHGTLSMDYYISAMEYIESECGSQTFIIFTDDVIWAKEKFS KYSNCLVADADENKFSVIDMYLMSLCNNNIIANSTYSWWGAWLNRSEDKLVIAPKQWYISGNECSLKNEN WIAM	158
Lachnospiraceae bacterium 3_1_57FAA_CT1	WP_016359991.1	511537894	protein [Lachnospiraceae bacterium 3_1_57FAA_CT1]	25.16		MIIIKVMGGLGNQMQQYALYEKFKSIGKNVKLDISWFEDSSVQEKVFARRSLELRQFKDLQFDTCSAEEKEA LLGKSGILGKLERKLIPARNKHFYESDIYHSEVFNMSDAYLEGHWACEKYYHDIMPLLQEKIQFPESANSQNIT VKKRMKAENSVSIHIRRGDYLDPENEAMFGGICTNSYYKAAEEYIKSRVPDTHFYLFSDDTAYLRENYHGDEY TIVDWNKGEDSFYDMELMSCCRHNICANSTFSFWGARLNRTPDKIVIRPAKHKNSQEIEPQLLHELWDNW VIIDGDGRIV	159
Butyrivibrio fibrisolvens	WP_022755397.1	551010878	glycosyl transferase [Butyrivibrio fibrisolvens]	25.09		MKPLVSLIVPVLNVEKYLEQCLTSISSQTYDNFEVILVVGKCIDNSENICKKWCEKDHRFRIEPQLKSCLGYARN VGIDAAKGEYIAFCDSDDCITSDFLSCFVDTALKNSSDIVETQFTLCDQNLSPIYDYDRNILGHILGHGFLEYTSA PSVWKYFVKRDIFTSNNLHYPEIRFGEDISMYSLLFSYCNKIDYVEKPTYLYRQVPSSLMNNPQGKRKRYESLF DIIDFVTNEFKTRLLFQKSWLKLLFQLEMHSASIISDSATSDDEAISMRQEISGYLKKVFPVKNTIFEVTALGWG GEIVSSIASKFNTLHGVSSSNMFNRYFFELLEDSTRKKLEEMIINFSPDIFLIDLISEADYLSSVKGNLGTFVKNW KIGFSIFMKMIQTHSNNSSIFLLENYMQQAPDHVDNTNEILKMLYDDIKINHPDIICISPAPDILNRSSEPELPCI YQLKLVSDKLHTMYSPVINCVETKGGLGNQIFQYVFSKYIEKMTGYRPLLHIGFFDYVKAIPGGTKRIFSLDKLF PDIETTSGKIPCSHVVEEKSFISNPGSDIFYRGYWQDIRYFSDVKDEVLESFNVDTSSMSKDVIDFADTIRNANS IAMHIRRQDYLNENNVSLFEQLSIDYYKSAVDMIRKEYADDLVLFIFSDDPEYANSIADSFDIEGFVMPLHKDY EDLYLITLAHHHIIANSTFSLWGALLSARKDGIRIAPRNWFKGTPATNLYPDKWLIL	160
Anaeromusa acidaminophila	WP_018702959.1	517532751	protein [Anaeromusa acidaminophila]	25.08		MFCVRIYGGLGNQMFQYALGRAMAKHYSETAAFDLSWYEQKIKPGFEASVCQYNIELSRKDRPKAWYEPIL KRISRHTDKLEMWFGLFFEKKYHYDSTVFERGLCKKNITLDGYWQSYKYFSAIEDDLRRELTIPKEREELIAISRS LPENSVSIHVRRGDYVSNPKANAMHGTCSWEYYQAAIEKMTGLVKEPQYVVFSDDITWTKENLPLPNAMYI GRELGLFDYEELILMSRCKHNIMANSTFSWWGAWLNSNPNKVVIAPRKWFRHKKIKVNDLFPSSWVVL	161

Bacteroides sp. 2_1_16	WP_00876 8986.1	496044479	glycosyl transferase family 11 [Bacteroides sp. 2_1_16]	25.08		MDIVVIFNGLGNQMSQYAFYLAKKKDNLNCHVIFDPKSTNVHNGAELKRVFGIELNRNYLDKIISYFYGYIFN KRIVNKLFSLVGIRMIYEPKNYDYREELLKPSSNFISFYWGGWHSEKYFKDIELEVKKVFKFPEVTNSPYFTEWF NKIFLDNNSVSIHIRRGDYLDKPSDPYYQFNGVCTIDYYEKAILYLKERILEPNFYIFSNDINWCMKTFGTENMY YVDCNKGKDSWRDMYLMSECRHHINANSTFSWWAAWLSPYSNGIVLHPKYFIKDIETKDYYPQKWIMIE	162
Chlorobium phaeobacteroides;Chlorobium phaeobacteroides BS1	YP_001960 319.1	189500849	glycosyl transferase family protein [Chlorobium phaeobacteroides BS1]	25.08		MDKVVVHLTGGLGNQMFQYALGRSISNRNCPLLLNTSFYDTYDKFSCGLSRYNVKAEFIKKNSYYNNKYYRY VIRLLSRYGVACYFGSYYEKKIFSYDEKVYKRSCVSYYGTWQSYGYFDSIRDILLRDYEMVGCLEEEVEKYVSDIK RVDSVSLHIRRGDYFDNKRLQSIHGILTMEYYYKAMSLFPDSSVFVYVFSDDIEWVRENLITNTNIVYVVLESDN PENEIYLMSLCKNNIISNSTFSWWGAWLNKNKYKKVIAPRMWYKDNQSSSDLMPSDWCLI	163
Treponema bryantii	WP_02293 2606.1	551312724	protein [Treponema bryantii]	25.08		MIVISMGGGLGNQMFEYAFYTQLKHLYPKSEIKVDTKYAFPYSHNGIEVFKIFGLNPPEANWKEVHSLVKTYP IEGNKAHFIKFFLYRILRKANLVEREPTSFCKQKDFTEFYNSFFELPQNKSFYLYGPFVNYNYFAAIHNEIMDLYT FPEITDVTNIEYKRKIESSHSISHIRRGDYITEGVPLVPDAYYREALVYINKKIEDPHFFVFTDDKDYCKSLFSDN QNFTIVEGNTGANSFRDMQLMSLCKHNIIANSTFSFWGAFLNKNSEKIVIAPNIAFKDCSCPYICPDWIIL	164
Bacteroides fragilis;Bacteroides fragilis 638R	YP_005110 943.1	375358171	LPS biosynthesis alpha-1,2-fucosyltransferase [Bacteroides fragilis 638R]	25		MVIAKLFGGLGNQMFIYAAAKGIAQISNQKLTFDIYTGFEDDSRFRRVYELKQFNLSVQESRRWMSFRYPLG RILRKISRKIGFCIPLVNFKFIVEKKPYHFQNEIMRIASFSSIYLEGYFQSYKYFSKIEAQIREDFKFTKEVIGSVEKE ASFITNSRYTPVAIGVRRYSEMKGEFGELAVVEHDYYDAAIKYIANKVPNLIFIVFSEDIDWVKKNLKLDYPVYF VTSKKGELAAIQDMYLMSLCNHHIISNSSFYWWGAYLASTNNHIVIAPSVFLNKDCTPIDWVII	165
Firmicutes bacterium CAG:534	WP_02235 2105.1	547951298	protein [Firmicutes bacterium CAG:534]	25		MSGGLGNQMFQYALYMKLTAMGREVKFDDINEYRGEKAWPIMLAVFGIEYPRATWDEIVAFTDGSMDFS KRLKRLFRGRHPIEYVEQGFYDPKVLSFENMYLKGSFQSQRYFEDILEEVQETFRFPELKDMNLPAPLYETTEK YLLRIEGCNAVGLHMYRGDSRSNEELYDGICTEKYYEGAVRFIQDKCPDAKFFIFSNEPKWVKGNVVISLMKS QIREDMSREEIRALEDHFVLIENNTEYTGYLDMFLMSRCRHNIISNSSFSWWAAFINENPDKLVTAPSRWVN GVPSEDVYVKGMTLIDEKGRVERTIKE	166
Firmicutes bacterium CAG:882	WP_02236 8748.1	547971670	glycosyl transferase family 11 [Firmicutes bacterium CAG:882]	25		MVIVKIGDGIGNQMFNYVCGYSVAKHDNDTLLLDTSDVDNSTLRTYDLDKFNIDFTDRESFTNKGFFHKVYK RLRRSLKYNVIYESRTENCPCVLDVYRRKFIRDKYLHGYFQNLCYFKTCKEDIMRQFTPKEPFSAKADELIHRFA TENTCSVHVRGGDIKPLSIKYYKDALDKIGEAKKDMRFIVFSNVRNLAEEYIKELGVDAEFIWDLGEFTDIEELF LMKACRRHILSDSTFSRWAALLDEKSEEVFVPFSPDADKIYMPEWIMEEYDGNEEKR	167

Vibrio parahaemolyticus;Vibrio parahaemolyticus 10329;Vibrio parahaemolyticus 10296;Vibrio parahaemolyticus 12310;Vibrio parahaemolyticus 10290	WP_00549 6882.1	491639353	glycosyl transferase family 11 [Vibrio parahaemolyticus]	25		MVVKVSGGLGNQLFQYAIGCAISNRLSCELLLDTSFYPKQSLRKYELDKFNIKAKVATQKEVFSCGGGDDLLS RFLRKLNLSSLFFPNYIKEKESLVYLAEISHCKSGSFLDGYWQNPQYFSDIKDELVKQIVPIMPLSSPALEWQNII INTKNCVSLHVRRGDYVNNAHTNSVHGVCDLSYYREAITNIHETVEKPKFFVFSDDISWCKDNLGSLGHFTYV DNTLSAIDDLMLMSFCEHHIIANSTFSWWGAWLNDHGITIAPKRWFSSVERNNKDLFPEKWLIL	168
Herbaspirillum frisingense;Herbaspirillum frisingense GSF30	WP_00646 3714.1	493509348	glycosyl transferase family protein [Herbaspirillum frisingense]	24.92		MIVSRLIGGLGNQMFQYAAGRALALRRGVPFAIDSRAFADYKTHAFGMQCFCADQTEAPSRLLPNPPAEGR LQRLLRRFLPNPLRVYTEKTFTFDEAVLSLPDGIYLDGYWQSEKYFADFADDIRKDFAVKAAPSAPNQAWLEL IGRTHSVSLHIRRGDYVSNAAAAAVHGTCDLGYYERAVAHLHQYTGQAPELFVFSDDLDWVATNLQLPYT MHLVRDNDAATNFEDLRLMTACRHHIVANSSFSWWGAWLDGRSESITIAPARWFVADTPDARDLVPQR WVRL	169
Rhizobium sp. CF080	WP_00775 9661.1	495034125	Glycosyl transferase family 11 [Rhizobium sp. CF080]	24.92		MIITRILGGLGNQMFQYAAGRALAIANEAELKLDLIEMGAYKLRPFALDQFNIKAAIAQPDEVPAKPKRGLLR KFTSAFKPDRSSCERIVENGLTFDSRVPALRGSLHLSGYWQSEQYFASSADAIRSDFSLKSPLGPARQDVLARI GAATTPVSIHVRRGDYVTNPSANAVHGTCEPPWYHEAMRRMLDRAGDASFFVFSDEPQWARDNLQSSRP MVFIEPQNNGRDGEDMHLMAACHAHIIANSSFSWWGAWLNPRPNKHVIAPRQWFRAPDKDDRDIVPA TWERL	170
Verrucomicrobium spinosum	WP_00995 9380.1	497645196	glycosyl transferase, family 11 [Verrucomicrobium spinosum]	24.85		MVISHISEGLGNQMFQYAAGRRLSYHLGTTLKLDDYHYRLHPFRSFQLDRFLITSPIATDAEISHLCPLEGLAR AIRARLPGKLRGATLRLLGNLGLGSPYQPRLHSFKEETPKQPLLIGKVVSERHFHFDPDVLECPDNVCLVGYW QDERYFGEIRDILLRELTLKSPPAGATKAVLERIQRSSSVSLHVRRGDKTKSSSYHCTSLEYCLAAMSEMRARL QAPTFFVFSDDWDWVREQIPCSSSVIHVDHNRAEDVSEDFRLMKSCDHHIIASSSLSWWAAWLGTNENSF VFSPPADRWLNFSNHFTADVLPPHWIQLDGSSLLPAQ	171
Fibrella aestuarina;Fibrella aestuarina BUZ 2	YP_007319 049.1	436833833	glycosyl transferase family 11 [Fibrella aestuarina BUZ 2]	24.83		MTANRVLVNSPMVIAKITSGLGNQLFQYALGRHLALQGNTSLWFDLRYFHQEYATDTPRKFKLDRFNVRYN LLDSSPWLIYASKATRLLPGRSLRPLIDTREADFHFDPTVIRPAAPLTILWGFWQSEKYFAQSTPQIRQELTFN RPLSDTFVGYQQQIEQAEVPISVHVRRGDYVTHPEFSQSFGFVGLAYYQKALAHLQDLFPNATLFFFSDDPD WVRANIVTEQPHVFVQNSGPDADVDDLQLMSLCHHHVIANSSFSWWGAWLNPRPDKVVIGPQRWFAN KPWDTKDLLPSGWLRL	172

Rhodobacter sp. CACIA14H1	WP_02366 5745.1	563380195	alpha-1,2-fucosyltransfera se [Rhodobacter sp. CACIA14H1]	24.83	MIHMRLVGGLGNQLFQYACGRAVALRHGTELVLDTRELSRGAAHAVFGLDHFAIRARMGASADLPPPRSR VLAYGLWRAGFMAPRFLRERGLGVNPAVLAAGDGTYLHGYFQSEAYFRDVVPQIRPELEIVTPPSDDNLRW ASRIAGDDRAVSLHVRRGDYVASAKGQQVHGTCDADYYARAVAAIRARAGIDPRLYVFSDDPHWARDNLA LDAETVVLDHNPPGAAVEDMRLMGVCRHHIIANSSFSWWGAWRNPSAGKVVVAPVRWFADPKLHNPDI CPPEWLRV	173
Rhodopirellula baltica SH 1;Rhodopirellula baltica	NP_868779 .1	32475785	fucosyl transferase [Rhodopirellula baltica SH 1]	24.83	MATSAHLHLSDEKQTLDSKASDRDCATTEASASDKTCTISISGGLGNQMLQYAAGRALSIHHDCSLQLDLKF YSSKRHRSYELDAFPIQAHRSIKPSFFSQILSKIQSESKHVPTYQEQSKRFDPAFFNTEPPVKIRGYFFSEKYFSPY ADQIRTELTPPIPPDQPARDMAIRLKECVSTSLHVRRGDYVTNANARQRFWCCTSEYFEAAIERLPTDSTVFV FSDDIEWAKQNIRSSRTTVYVNDELKKAGSPETGLRDLWLMTHAKSHIIANSSFSWWGAWLANSEANLTIA PKKWFNDPEIDDSDIVPSSWHRI	174
Spirosoma spitsbergense	WP_02060 4054.1	522092845	protein [Spirosoma spitsbergense]	24.83	MVVVELMGGLGNQMFQYAFGMQLAHQRQDTLTVSTFLLSNKLLANLRNYTYRPFELCIFGIDKPKASPFNL LRALLPFDLNTSLLRETDDPEAVIPAASARIVCVGYWQSEHYFEEVTVHVREKFIFRQPFNSFTSRLANNLNGI PNSVFVHIRRGDYVTNKGANAHHGLCDRTYYERAVTFMREHLENPLFFIFSDDLEWVSQELGPILEPATYVG GNQKNDSWQDMYLMSLCRHAIVANSSFSWWGAWLSPHASKIVVAPKEWFGKPLLPVKTNDLIPNSWIRI	175
uncultured bacterium	EKD71402. 1	406938106	protein ACD_46C00193 G0003 [uncultured bacterium]	24.76	MNAIIPRETGGIGNQLFIYAAARRMAIANSMNLVIDDTSGFKYDVLYKRFYQLEKFNITSRMATPTERLEPFSK IRRYLKRKINKTYPFAQRAYITQEKSGFDPRLLVFRPKGNWYLDGYWQSENYFKDIEGIIRQDLIIKSPSDSLNIA TAERIKNTLAIAVHVRFFDMVDISDSSNCQSNYYHTAIAKMEEKIPNAHYFIFSDKPVLARLAMPLPDDRITIID HNIGDMNAYADLWLMSLCKHFVIANSTFSWWGAWLSDNKEKIVIAPDIKITSGVTQWGFDGLIPDEWIKL	176
Prevotella micans;Prevotella micans F0438	WP_00695 0883.1	494008437	protein [Prevotella micans]	24.75	MDVIVIFNGLGNQMSQYAFYLEKRLRNRQTTYFVLNPRSTYELERLFGIPYRSNLMCRMIYKLLDKAYFSNHI RLKKILRTALNAVGIRLIVEPITRNYSLSNFTHHPGLTFYRGGWHSELNFTSVVTELRRKFIFPPSDDEEFKRISAL IIRTQSISLHIRRGDYLDYSEYQGVCTEEYYERAIEYIRSHVENPVFFVFSDDKEYAINKFSGDDSFRIVDFNTGE NSWRDMQLMSLCRHHILANSTFSWWGAWLDSAPEKIVLHPIYHMRDVPTRDFYPHNWIGISGE	177
Thermosynechococcus sp. NK55	AHB87954. 1	564737556	alpha-1,2-fucosyltransfera se [Thermosynechococcus sp. NK55]	24.75	MIIVRLYGGLGNQMFQYAAGLALSIRHAVPLRFDLDWFDGVRLHQGLELHRVFDLDLPRAAPSEMRQVLG SFSHPLVRRLLVRRRLRWLLPQGYALEPHFHYWPGFEALGPKAYLDGYWQSERYFSEYQDAVRAAFRFAQP LDERNRQIVEEMAACESVSLHVRRGDFVQDPVVRRVHGVDLSAYYPRAVALIMERMREPRFYVFSDDPD WVRANLKLPAPMIVIDHNRGEHSFRDMQLMSACRHHILANSSFSWWGAWLNSQPHKLVIAPKRWFNVD DFDTRDLYCSGWTVL	178

Coleofasciculus chthonoplastes;Coleofasciculus chthonoplastes PCC 7420	WP_006100814.1	493031416	Glycosyl transferase family 11 [Coleofasciculus chthonoplastes]	24.73		MLSLNKNFLFVHIPKSCILKEVYIYMISFPNLGKGVRLGNQMFQYAFLRSTARRLGVKFYCPAWSGDSLFTLN DQEERVSQPEGITKQYRQGLNPGFSENALSIQDGTEISGYFQSDKYYDNPDLVRQWFSLKEEKIASIRDRFSRL NFANSVGMHLRFGDVVGQLKRPPMRRSYYKKALSYIPNQELILVFSDEPERTKKMLDGLSGNFLFLSGHKNY EDLYLMTKCQHFICSYSTFSWWGAWLGGERERTVIYPKEGQYRPGYGRKAEGVSCESWIEVQSLRGFLDDY RLVSRLEKRLPKSLMNFFY	179
Bacteroides gallinarum	WP_018666797.1	517496220	glycosyl transferase [Bacteroides gallinarum]	24.66		MRLIKMTGGLGNQMFIYAFYLRMKKRHTNTRIDLSDMMHYNVHHGYEMHHVFNLPKTEFCINQPLKKVIE FLFKKIYERKQDSSNLLPFDKKYFWPLLYFKGFYQSERFFADMENDIRKAFTFNSGLFNEKTQTMLKQIEHNE HAVSLHVRRGDYLEPKHWKTTGSVCQLPYYINAIAEMNRRIEQPFYVVFSDDIAWVKENLPLPQAVFIDWN KGVESWQDMMLMSHCRHHIICNSTFSWWGAWLNPKENKTVIMPERWFQHCETPNIYPAGWIKVPIN	180
Firmicutes bacterium CAG:882	WP_022367483.1	547967507	glycosyl transferase GT11 family [Firmicutes bacterium CAG:882]	24.66		MNNVEIMGGLGKQLFQYAFSRYLQKLGVKNVVLRKDFFTIQFPENNGITKREFVLDKYNTRYVAAAGEKTYR DYCDENDYRDDYAIGSDEVLYEGYWQNIDFYNVVRKEMQEELKLKPEFIDNSMAAVEKDMSSCNSVALHIR RSDYLTQVNAQIFEQLTQDYYASAVSIIEQYTHEKPVLIYIFSDDPEYAAENMKDFMGCRTVIMPPCEPYQDM YLMTRAKHNIIANSTFSWWGATLNANPDNITVAPSRWMKGRTVNLYHKDWITL	181
Bacteroides xylanisolvens;Bacteroides xylanisolvens CL03T12C04	WP_008021494.1	495296741	protein [Bacteroides xylanisolvens]	24.6		MIAVNVNAGLANQMFHYAFGRGLMAKGLDVCFDQSNFKPRSQWAFELVRLQDAFPSIDIKVMPEGHFK WVFPSLPRNGLERRFQEFMKKWHNFIGDEVYIDEPMYGYVPDMEKCATRNCIYKGFWQSEKYFRHCEDDI RKQFTFLPFDELKNIEVAAKMSQENSVAIHLRKGDDYMQSELMGKGLCTVDYYMKAIDYMRKHINNPHFY VFTDNPCWVKDNLPFEFYILVDWNEVSGKRNFRDMQLMSCAKHNIIGNSTYSWWAAWLNANQDKIVVG PKRFFNPINSFFSTCDIMCEDWISL	182
Geobacter sp. M18	YP_004197726.1	322418503	glycosyl transferase family protein [Geobacter sp. M18]	24.58		MIGMVIFRAYNGLGNQMFQYAVLGRHLALLNEAELKIDTTAFADDPLREYELHRLKVQGSIATPDEIAFFREM ENTHPQAYLRLTQKSRLFDPAILSARGNIYLHGFWQTEKYFADIREILLDEFEPIVPAGEDSIKVLSHMKATNA VALHVRRSDYVSNPMTLRHHGVLPLDYYREAVRRIAGMVPDPVFFIFSDDPQWAKDNIRLEYPAFCVDAHD ASNGHEDLRLMRNCKHFIIANSSFSWWGAWLSQNTGKKVVAPLKWFAKPEIDTRDIVPLQWIRI	183
Ruegeria pomeroyi DSS-3	YP_168587.1	56698215	alpha-1,2-fucosyltransferase, [Ruegeria pomeroyi DSS-3]	24.57		MITTRLHGRLGNQMFQYAAARGLAARIGTQVALDTRLAESRGEGVLTRVFDLDLAQPDQLPPLKGDGLLR HGAWRLLGLAPRFRREHGLGYNAAIETWDDGTYLHGYWQSERYFAHIAARIRADFAFPAFSNSQNAEMAA RIGDTDAISLHVRRGDYVALAAHTLCDQRYYAAALTRLLEGVAGDPVVYLFSDDPAWARDNLALPVQKVVV DFNGPETDFEDMRLMSLCRHNIIGNSSFSWWAAWLNAHPGKRIAAPASWFGDAKLHNPDLLPPDWVLKIE V	184

Lachnospiraceae bacterium 28-4	WP_01629 1997.1	511037973	protein [Lachnospiracea e bacterium 28-4]	24.52		MIIQLAGGLGNQMQQYAMYQKLLSLGKKVKLDISWFEEKNRQKNVYARRELELNYFKKAEYEACTEEERKA	185
						LVGEGGFAGKIKGKLFPGTRKIFRETEMYHPEIFDFEDRYLYGYFACEKYYADIMEILQEQFVFPPSGNPENQK	
						MAERIADGESVSLHIRRGDYLDAENMAMFGNICTEEYYAGAIREMKKIYPSAHFFVFSDDIPYAKETYSGEEF	
						TVVDINRGKDSFFDIWLMSGCRHNICANSTFSFWGARLNRNKGKVVMRPFIHKNSQKFEPELMHELWKG	
						WYFIDNRGNIC	
Prevotella sp. CAG:1092	WP_02198 9703.1	547254188	glycosyl transferase family 11 [Prevotella sp. CAG:1092]	24.49		MRILVFTGGLGNQMFEYAFYKHLKSCFPKESFYGHYGVKLKEHYGLEINKWFDVTLPPAKWWTLPVVGLFYL	186
						YKKLVPNSKWLDLFQREWKHKDAKVFFPFKFTKQYFPKENGWLKWKVDEASLCEKNKKLLQVIHDEETCFV	
						HVRRGDYLASNFKSIFEGCCTLDYYKRALEYMNKNNPKVRFICFSDDLEWMRKNLPMDDSAIYVDWNTGT	
						DSPLDMYMMSQCDNGIIANSSFSYWGAYLGGKKTTVIYPQKWWNMEGGNPNIFMDEWLGM	
Spirosoma luteum	WP_01861 8567.1	517447743	protein [Spirosoma luteum]	24.41		MVISVLSGGLGNQLFQYAFGLKLAAQLQTELRLERHLLESKAIARLRQYTPRTYELDTFGVEAPAASLMDTVS	187
						CLSRVALSDKTALLLRESTLTPNAINNLNNRVRDVVCLGYWQSEEYFRPATEQLRKHLVFRKNPAQSRSMAD	
						TILSCQNAAFVHIRRGDYVTNTHANQHHGLCDVSYYRRACEYVKECIPDVQFFVFSDDPDWAKRELGIHLQP	
						ARFIDHNRGADSWQDMYLMSLCRHAIVANSSFSWWGAWLNPVAERLVVAPGQWFVNQPVLSQQIIPPH	
						WHCL	
Marinomonas posidonica IVIA-Po-181;Marinomonas posidonica	YP_004480 472.1	333906886	glycosyl transferase family protein [Marinomonas posidonica IVIA-Po-181]	24.34		MIIVDLSGGLGNQMFQYACARSLSIELNLPLKVVYGSLASQTVHNGYELNRVFGLDLEFATENDMQKNLGFF	188
						LSKPILRKIFSKKPLNNLKFQNFFPENSFNYNSSLFSYIKDSGFLQGYWQTEKYFLNHKSQILKDFCFVNMDDE	
						TNISIANDIQSGHSISIHVRRGDYLTNLKAKAIHGHCSLDYYLKAIEFLQEKIGESRLFIFSDDPEWVSENIATRFS	
						DVSVIQHNRGVKSFNDMRLMSMCDHHIIANSSFSWWGAWLNPSQNKKIIAPKNWFVTDKMNTIDLIPSS	
						WILK	
Bacteroides;Bacteroides sp. 4_3_47FAA;Bacteroides sp. 3_1_40A;Bacteroides dorei 5_1_36/D4;Bacteroides vulgatus PC510;Bacteroides dorei CL03T12C01;Bacteroides vulgatus dnLKV7	WP_00583 9979.1	492425792	glycosyl transferase family 11 [Bacteroides]	24.32		MKIVVFKGGLGNQLFQYAFYKYLSRKDETFYFYNDAWYNVSHNGFELDKYFKTDDLKKCSRFWIILFKTILSKL	189
						YHWKIYVVGSVEYQYPNHLFQAGYFLDKKYYDENTIDFKHLLLSEKNQSLLKDIQNSNSVGVHIRRGDYMTK	
						QNLVIFGNICTQKYYHDAIRIITEKVNDAVFYVFSDDISWVQTHLDIPNAVYVNWNTGESSIYDMYLMSSCKY	
						NIIANSTFSYWAARLNKKTNMVIYPSKWYNTFTPDIFPESWCGI	

			protein [Candidatus Pelagibacter			
Candidatus Pelagibacter ubique	WP_02016 9431.1	519013556	ubique]	24.32	MTIRIKLTGGLGNQMFQFATGFAIAKKNVRLSLDLKYINKRKLFNGFELQKIFNIYSKVSFLNKTLSFKSINFTE ILNRIDTTFYNFKEPHFHYTSNILNLPKHSFLDGYWQSELYFNEFATEIKRIFNFSGKLDKSNLLVADDINRNNSI SIHIRRGDFLLKQNNNHHTDLKEYYLKAINETSKIFKNPKYFIFSDDTSWTVDNFVIDIHPYIIVDINFGARSFLD MYLMSLCKSNIIANSSFSWWSAWLNNNKDKIIYAPKNWFNDKSICTDDLIPESWNIIL	190
Bacteroides sp. CAG:875	WP_02235 3174.1	547952428	uncharacterized protein [Bacteroides sp. CAG:875]	24.29	MSVIINMACGLANRMFQYAFYLYLQKEGYDAYVDYFTRADLVHENVDWLRIFPEATFRRATARDIRKMGG GHDCFSRLRRKLLPMTTKVLETSGAFEIILPPKNRDSYLLGAFQSAKMVESVDAEVRRIFTFPEFESGKNQYFQ TRLAQENSVGLHIRKGKDYQERIWYKNTCGVEYYRKAVDLMKEKVDSPSFYVFTDNPAWVKENLSWLEYKL VDGNPGSGWGSHCDMQLMSLCKHNIISNSSYSWWGAYLNNNTLNKIVVCPRIWFNPESTKDFSSNPLLAEG WISL	191
Butyrivibrio fibrisolvens	WP_02275 6327.1	551011911	protein [Butyrivibrio fibrisolvens]	24.29	MIIIKLQGGLGNQLFLYGLYKNLKHLKRDVKMDIESGFEGDELRKPCLDCMNLEYAIATRDEVTDIRDSYMDI FSRIRRKITGRKTFDYYEPEDGNYDPKVLEMTKAYLNGYFQSEKYFGDEESVKALKDELTKGKEDILTSTDLITKI YHDIKNSESVSLHIRRGDYLTPGIIETYGGICTDEYYDKAIAMIRETFPEARFFIFSNDIEWCKEKFAGDKNILFV NTIGINLDSEDNIKIGKSDKDISEYRDLAELYLMSACKHHILANSSFSWWGAWLSDHEGMTIAPSKWLNNKN MTDIYTKDMLLI	192
Roseburia hominis;Roseburia hominis A2-183	YP_004839 455.1	347532692	glycosyl transferase family protein [Roseburia hominis A2-183]	24.22	MVTVKIGDGMGNQMYNYACGYAAAKRSGEKLRLDISECDNSTLRDYELDHFRVVYDEKESFPNRTFWQKL YKRLRRDIRYHVIRERDMYAVDARVFVPARRGRYIHGYWQCIGYFFFYLDDIREMFTPAYEQTDAVRELM CQFTQTPTCALHVRGGDLGGPNRAYFQQAIARMQKEKPDVTFIVFTNDLPKAKECLDDGEARMRYIAEFGE ALSDIDEFFLMSACQNQIISNSTYSTWAAYLNTLPGRIVIVPKFHGVEQMALPDWIVLDGGACQKGEIDAV	193
Rhodopirellula europaea;Rhodopirellula europaea 6C	WP_00865 9200.1	495934621	alpha-1,2-fucosyltransferase [Rhodopirellula europaea]	24.16	MATSVHPHLSDGKQALDSKAAQQVCSTQAASASDRACTISISGGLGNQMLQYAAGRALSIHHDCPLQLDLK FYSSKRHRSYELDAFPIQAQRWIKPSFFSQVLDKIQGESKSAPTYEEQSKRFDRAFFDIELPARIRGYFFSEKYFL PYADQIRTELTPPVPLDQPARDMAQRLSEGMSTSLHVRRGDYVSNANARQRFWSCTSEYFEAAIEQMPAD STVFVFSDDIEWAKQNIRSSRPTVYVNDELKLAGSPETGLRDLWLMTHAKSHIIANSSFSWWGAWLSGSEA NLTIAPKKWFNDPEIDDSDIVPTSWRRI	194
Rudanella lutea	WP_01998 8573.1	518832653	protein [Rudanella lutea]	24.16	MVIAKITSGLGNQLFQYALGRHLAIQNQTRLWFDLRYYHRTYETDTPRQFKLDRFSIDYDLLDYSPWLYVSKA TRLLPGRSLRPLFDTRKEPHFHLDPAVPNAKGAFITLDGFWQSEGYFASNAATIRRELTFTRQPGPMYAEYR QQIEQTQTPVSVHIRRGDYVSHPEFSQSFGALDDTYYQTALAQINGQFPDATLLVFSDDPEWVRQHMRFER PHVLVENTGPDADVDDLQLMSLCHHHIIANSSFSWWGAWLNPRPDKRVIAPKQWFRNKPWNTADLIPAG WVRL	195

Bacteroidetes;Capnocytophaga sp. oral taxon 329 str. F0087;Paraprevotella clara YIT 11840	WP_00861 8094.1	495893515	glycosyl transferase [Bacteroidetes]	24.15		MRLIKMTGGLGNQMFIYAMYLKMKTIFPDVRIDLSDMVHYQVHYGYEMNKVFHLPRTEFCINRSLKKIIEFL LFKTILERKQGGSLVPYTRKYHWPWIYFKGFYQSEKYFAGIEKEVREAFVFDIRRASRRSLRAMQEIKADPHA VSIHVRRGDYLLEKHWKALGCICQSSYYLNALAELEKRVKHPHYYVFSEDLNWVRQNLPLIKAEFIDWNKGE DSWQDMMLMSHCRHHIICNSTFSWWGAWLNPLPDKIVIAPERWTQITDSADVVPESWLKVSIG	196
Smaragdicoccus niigatensis	WP_01815 9152.1	516906936	protein [Smaragdicoccus niigatensis]	24.08		MADVVTLAGGLGNQLFQTAYAKNLEARGHRVTLDGTVVRWTRGLHIDPQICGLKILNATPPAPVPGRLAA TVLRRALATRLRFGPDGRIVRTQRTLEFDEQYLNLNSPGRYRVEGYWQCERYFSDVGQTVRKVFLDMLGRH VSYNGLSRLPAMADPSSISLHVRRGDYVTANFIDPLALEYYERALEELAVPSPRIFVFSDDLDWATRELGRICD VIPVEPDWTSHPGGEIFLMSQCSHHIIANSSFSWWGAWLDGRTSSRVVAPRQWFSLETYSARDIVPDRWT KV	197
Bacteroides fragilis CAG:558	WP_02201 2576.1	547279005	family 11 glycosyltransferase [Bacteroides fragilis CAG:558]	24.05		MIHLILGGGIGNQMFQYAFARSLALQYNENISFNTILYKELKNEERSFSLGHLNINTMCIVETPDENKRIWELF NKQIFHQKIARKILPASIRWVWMSNRNIYANVCGPYKYYHPRHRSQNTTIIHGGFQSWKYFKEHQSMIKAE LKVITPISEPNKKILKEIQNSNSICVHIRRGDFLSAQFSPHLEVCNKDYYEKAIKMISSQIENPTFFIFSNTHEDLV WIRKNYNIPQNSVYVDLNNPDYEELRLMYNCKHFILSNSSFSWWAQYLSESKNKIIIAPKIWDKRKGIDFSDIY MPEWIIIK	198
Desulfovibrio desulfuricans	WP_02265 7592.1	550904402	protein [Desulfovibrio desulfuricans]	24.05		MSFSIDVAAIQRMALVKVDGGLGSQMWQYALSLAVGKSSSFTVKHDLSWFRHYAKDIRGIENRFFILNSVFT NINLRLASENERLFFHIALNKYPDSICNFDPDILALKQPTYLGGYYVNAQYVTSAEKEIREAYVFAPAVEESNQA MLQTIHAAPMPVAVHVRRGDYIGSMHEVLTPRYFERAFKILAAALQPKPTFFVFSNGMEWTKKAFAGLPYD FVYVDANDNDNVAGDLFLMTQCKHFTISNSSLSWWGAWLSQRAENKTVIMPSKWRGGKSPIPGECMRV EGWHIMCPVE	199
Hoeflea phototrophica;Hoeflea phototrophica DFL-43	WP_00719 9917.1	494373839	alpha-1,2-fucosyltransferase, [Hoeflea phototrophica]	24.05		MHGGLGNQLFQYAVGRAVALRTGSELLLDTREFTSSNPFQYDLGHFSIQAKVANSSELPPGKNRPLAYAW WRKFGRSPRFVREQDLGYNARIETIEADCYIHGYFQSQKYFEDIASILWKDLSFRQAISGENASMAERIQSAP SVSMHIRRGDYLTSAKARSTHGAPDLGYYGRALGEIRARSGSDPVWYLFSDDPDWVRNNMRMDANLVTVA INDGKTAFEDLRLMSLCDHNIIVNSTFSWWGAWLNPSLDKIVVAPKRWFADPKLSNPDITPPGWLRLGD	200
Vibrio cholerae;Vibrio cholerae O1 str. 87395	WP_00203 0616.1	487957217	glycosyl transferase, family 11 [Vibrio cholerae]	24.04		MKIISFSGGLGNQLFQYAFYLYLKDNSDFGNIFLDFSFYESQNKRDAVIRNFYGVDSLDIIKQSSYVRGKFLILKL INKFRFFNNLLEVDKENGLDETLLSTNKVFFDGYWQSYRYKDYKSNIKELFSFYDFKGNILEVRKKICQSNSV CMHVRRGDYVAEKNTKLVHGVCSLQYYRDALNNIKNVDNSIDHIFIFSDDIDWVKNNISFDIPVTVVDFVGQ SVFDYAEMLLFSCGKHKVIANSTFSWWGAFLSDRNGVIVSPKKWFAKEEKNYDEIFIEGSLRL	201

Lachnospiraceae bacterium NK4A179	WP_02278 4718.1	551041074	protein [Lachnospiracea e bacterium NK4A179]	24.03		MIIVRFRGGMGNQMFQYAFLRYLEMKGATLKADLSEFKCMKTHAGYELDKAFDLHPAEASYKEIRAVADYI PVMHIRFPFSRKVFEILYKKETKRVEAEGPKKSHISEEKYFDMSEDERLHLASSSEDLYMDGFWIKPDMYDDE VLKCFTFSKTLDEKYKGTIEDEHSCSVHVRCGDYTGTGLDILGKEYYEKAAEKILSEDADVKFYVFSDDREKAEK LLSPFMKKMVFCDTPASHAYDDMYLMSRCRHHIIANSTFSFWGARLSADKSGITICPKYEDKNNTANRLVHE GWQML	202
Cecembia lonarensis;Cecembia lonarensis LW9	WP_00918 5692.1	496476931	Glycosyl transferase family 11 [Cecembia lonarensis]	24.01		MIIMKFMGGLGNQIYQYALGRKLSELHNSFLASDIHIYKNDPDREFVLDKFNIKVKHLPWKVIKLLNSDYALKF DKVFHTEFYHELVLEKALESKDIPRKNNLYLRGSWGNRKYYEDYIDKISDEITLKEKFKTKDFNTVNKKVKNSDS VGIHIRRGDYEKVAHFKNFYGLLPPSYYSAAVDFIGNRIEKSNFFIFSDDTDWVKENLPFLKDSFFVSDIIGSVD YLEFELLKNCKHQIIANSTFSWWAARLNSNPAKIVIKPKRWFADDRQQAVYEIEDSYYIKEAIKL	203
Bacteroides ovatus;Bacteroides ovatus ATCC 8483	WP_00429 5547.1	490423336	protein [Bacteroides ovatus]	24		MKIVNILGGLGNQMFVYAMYLALKEAHPEEFIILCRRSYKGYPLHNGYELERIFGVEAPEAALSQLARVAYPF FNYKSWQLMRHFLPLRKSMASGTTQIPFEDYSEVTRNDNVYYDGYWQNEKNFLSIRDKVIKAFTFPEFRDEK NKALSDKLKSVKTASCHIRRGDYLKDPIYGVCNSDYYTRAITELNQSVNPDMYCIFSDDIGWCKENFKFLIGDK EVVFVDWNKGQESFYDMQLMSLCHYNIIANSSFSWWGAWLNNNDDKVVVAPERWMNKTLENDPICDN WKRIKVE	204
Bacteroides coprocola CAG:162	WP_02212 5287.1	547668508	glycosyltransfer ase family 11 [Bacteroides coprocola CAG:162]	23.99		MRLIKMTGGLGNQMFIYAFYLKMKKLFPHTKIDLSDMMHYHVHHGYEMNRVFALPHTEFCINRTLKKLME FLLCKVVYERKQKNGSMEAFEKKYAWPLIYFKGFYQSERFFADIEDDVRKTFCFNMELINSRSREMMKIIDAD EHAVSIHIRRGDYLLPKFWANAGCVCQLPYYKNAITELEKHESTPSFYVFSDDIEWVKQNLSLPNAHYIDWN QGNDSWQDMMLMSHCRNHIICNSTFSWWGAWLNPRKNKTVIVPSRWFMKEETPYIYPVSWIKVPIN	205
Bacteroides dorei;Bacteroides dorei DSM 17855;Bacteroides dorei CL03T12C01	WP_00783 5585.1	495110765	glycosyl transferase [Bacteroides dorei]	23.99		MRLIKVTGGLGNQMFIYAFYLRMKKYYPKVRIDLSDMMHYKVHYGYEMHRVFKLPHTEFCINQPLKKIEFLF FKKIYERKQAPNSLRAFEKKYFWPLLYFKGFYQSERFFADIKDEVREAFTFDRSKANSRSLDMLDILDKDENAV SLHIRRGDYLQPKHWATTGSVCQLPYYQNAIAEMSKRVTSPSYYIFSDDIVWVRENLPLQNAVYIDWNTGE DSWQDMMLMSHCKHHIICNSTFSWWGAWLNPSIDKTVIVPSRWFQYSETPDIYPTGWIKVPVD	206
Bacteroides;Bacteroides intestinalis DSM 17393;Bacteroides intestinalis CAG:564	WP_00766 2951.1	494936920	protein [Bacteroides]	23.97		MIIVRLWGGLGNQLFQYSFGQYLEIETDKKVFYDVASFGTSDQLRKLELCSFIPDIPLYNAYFTRYTGVKNRLF KALFQWSNTYLSESMFDICLEKARGKIFLQGYWQEEKYATYFPMQKVLSEWKNPNVLSEIEENIRSAKISVS LHVRRGDYFSPKNINVYGVCTEKYYEQAIDRANSEIEEDKQFFVFSDDILWVKNHVSLPESTVFVPNHEISQFA YIYLMSLCKVNIISNSTFSWWGAYLNQHKNQLVIAPSRWTFTSNKTLALDSWTKI	207

Lachnospiraceae bacterium A4	WP_016283022.1	511028838	protein [Lachnospiraceae bacterium A4]	23.95		MIVIHVMGGLGNQLYQYALYEKLRALGREVKLDVYAYRQAEGAEREWRALELEWLEGIRYEVCTAAERQQL LDNSMRLADRVRRRLTGRRDKTVRECAAYMPEIFEMDDVVLYGFWGCEKYYEDIIPLLQEKIVFPESSNPKN ADVLRAMAGENAVSVHIRRKDYLTVADGKRYMGICTDAYYKGAFRYITERVERPVFYIFSDDPAFAKTQFCE ENMHVVDWNTGRESLQDMALMSRCRHNICANSTFSWGARLNNRHPDKIMIRPLHHDNYEALDARTVHEY WKGWVLIDADGKV	208
Phaeobacter gallaeciensis;Phaeobacter gallaeciensis DSM 17395 = CIP 105210	YP_006574665.1	399994425	protein PGA1_c33070 [Phaeobacter gallaeciensis DSM 17395 = CIP 105210]	23.91		MIITRLHGRLGNQMFQYAAGRALADRLGVSVALDSRGAELRGEGVLTRVFDLDLATPDILPPLRQRAPLGYA LWRGLGQHLGTGPKLRREVGLGYNPDFVDWSDNSYLHGYWQSERYFAQSAERIRRDFTFPEYSNQQNAE MAARIGETNAISLHVRRGDYLTLAAHVLCDQAYYEAALAQVLDGLEGQPTVYVFSDDPQWAKENLPLPCDK VVVDFNGADTDYEDMRLMSLCKHNIIGNSSFSWWAAWLNQTPDRRVAGPTKWFGDPKLNNPDILPPDW LRISV	209
Firmicutes bacterium CAG:791	WP_021849028.1	546362318	protein [Firmicutes bacterium CAG:791]	23.88		MSGGLGNQMFQYALYLKLRSLGREVCFDDKSQYDEETFRNSSQKRRPKHLDIFGITYPSAGKEELEKLTDGA MDLPSRIRRKILGRKSLEKNDRDFMFDPSFLEETEGYFCGGFQSPRYFAGAEEEVRKAFTFPEELLCPKEGCSR QEQKMLEQSASYAERIRKANCEAADRGVPGGGSASIHLRFGDYVDKGDIYGGICTDAYYDTAIRCLKERDPG MIFFVFSNDEEKAGEWIRYQAERSENLGRGHFVLVKGCDEDHGYLDLYLMTLCRNHVIANSSFSWWAASFM CDAPDKMVFAPSIWNNQKDGSELARTDIYADFMQRISPRGTRLSDRPLISVIVTAYNVAPYIGRALDSVCGQ TWKNLEIIAVDDGSSDETGAILDRYAAGDSRIQVVHTENRGVSAARNEGIAHARGEYIGFVDGDDRAHPAM YEAMIRGILSSGADMAVVRYREVSAEETLTDAEEQVASFDPVLRASVLLQQRDAVQCFIRAGMAEEEGKIVL RSAVWNKLFHRRLLRDNRFPEGTSAEDIPFTTRALCLSKKVLCVPEILYDYVVNRQESIMNTGRAERTLTQEIP AWRTHLEILKESGLSDLAEESEYWFYRRMLSYEEEYRRCSETAKEAKELQERILKHRDRILELAEEHSFGRRGD RERLKLYVNSPRQYFELLSDLYEKTVVNWKNRPDKT	210
Butyrivibrio proteoclasticus;Butyrivibrio proteoclasticus B316	YP_003829733.1	302669773	glycosyl transferase 11 [Butyrivibrio proteoclasticus B316]	23.84		MRKRIIALNGGLGNQMFQYAFARMLEDRKHCLIEFDTGFYSTVNDRKLAIQNYNIHKYDFCNHEYYNKIRLLF QKIPFVAWLAGTYKEYSEYQLDPRVFLFNYRFYYGYWQNKQYFENISNDIRNELSYIGNVSEKENALLNMLEA HNAIAIHVRRGDYTQEGYNKIYISLSKEYYKRAVSIACKELGDNNIPLYVFSDDIDWCKANLADIGNVTFVDNT ISSSADIDMLMMKKSRCLITANSTFSWWSAWLSDRDDKIVLVPDKWLQDEEKNTKLMKAFICDKWKIVPV	211
Bacteroides sp. 2_1_16	WP_008768245.1	496043738	protein [Bacteroides sp. 2_1_16]	23.81		>gi|496043738|ref|WP_008768245.1| protein [Bacteroides sp. 2_1_16]MQVVARIIGGLGNQMFIYATARALALRIDADLILDTQSGYKNDLFKRNFLLDSFCISYRKANCFQKY DYYLGEKVKSLGKKTHFSVIPFMKYISENTSCDFVDGLLKKHILSVYLDGYWQNEAYFKDYASIIKKDFQFCQV NDLRTLSEAEIIKKSITPVAIGVRRYQELNSHQNTKVTDLDFYQKAINYIESKVDNPTFFIFSEDQEWVKNNLEQ KSNFIMISPKEGNYSALNDMYLISLCKHHIVSNSSFYWWGAWLANNKNKIVVASDCFLNPQSIPDSWIKF	212

Desulfomicrobium baculatum;Desulfomicrobium baculatum DSM 4028	YP_003159045.1	256830317	glycosyl transferase family protein [Desulfomicrobium baculatum DSM 4028]	23.76		>gi|256830317|ref|YP_003159045.1| glycosyl transferase family protein [Desulfomicrobium baculatum DSM 4028] MAKIVTRIMGGIGNQLFCYAAARRLALVNHAELVIDDVTGFSRDRVYRRRYMLDHFNISARKATNYE RMEPFERYRRGLAKYISKKLPFFEREYIEQERIEFDPRFLEYRTYNNIYIDGLWQSENYFKDVEDIIRDDLKIIPPT DLENINIAKKIKNIQNTIAMHVRWFDLPGINLGNNVSTYYYHRAIAMMEQRINAPHYFLFSDNLEAVHSKLD LPEGRVTFVSNNDGDDNAYADLWLMSQCKHFITANSTFSWWGAWLGESRDSVVLVPRFSPDGGVTSWC FTGLIPERWEQVSSIR	213
Prevotella pleuritidis;Prevotella pleuritidis F0068	WP_02158 4236.1	545304945	galactoside 2-alpha-L-fucosyltransferase [Prevotella pleuritidis]	23.76		MDIVLIFNGLGNQMSQYAFYLAKRQRNNHTVYCVFGPRTQYSLDKLFDIPYRHNAVLVLYRALDKAHFSN HRWLRRLIRPTLQLLGVKMIVEPLSRDFDMRHFTHQKGIVFYRGGWHSELNFTAVADAVKRRFRFPEIQDA AVLAVIDRIKSCQSVSLHLRRGDYLGLSEFQGVCTEAYYEHAIAYFESQIESPEYFVFSDDPTYAREQFGADPNF HIIDLNHGEDAWCDILLMMTQCRYNIIANSTFSWWGAWLNDNPSKIVVHPRYHLNGVETRDFYPRNWICIE	214
Bacteroides sp. 1_1_14	WP_00876 3191.1	496038684	glycosyl transferase, family 11 [Bacteroides sp. 1_1_14]	23.75		MKVIWFNGNLGNQVFYCKYKEFLHNKYPNETIKYYSNSRSPKICVEQYFRLSLPDRIDSFKVRFVFEFLGKFFR RIPLKFVPKWYCTRKSLNYEASYFEHYLQDKSFFEKEDSSWLKAKKPDNFSEKYLIFENLICNTNSVAVHIRRGD YIKPGSDYEDLSATDYYEQAIKKATEVYLDSQFFFSDDLEFVKNNFKGDNIYYVDCNRGADSYLDILLMSQAK INIIANSTFSYWGAYMNHEKKKVMYSDLWFRNESGRQMPNIMLDSWICIETKRK	215
Agromyces subbeticus	WP_02289 3737.1	551273588	protein [Agromyces subbeticus]	23.65		MVGRVGIARRQAADVSCTDGEGLVAWRIRTGEIYLGLQGGIGNQLFEWAFAMALRSIGRRVLFDAVRCRG DRPLMIGPLLPASDWLAAPVGLALAGATKAGLLSDRSWPRLVRQRRSGYDPSVLERLGGTSYLLGTFQSARY FDGVEHEVRAAVRALLEGMLTPSGRRFADELRADPHRVAVHVRRGDYVSDPNAAVRHGVLGAGYYDQAL EHAAALGHVRRVWFSDDLDWVREHLARDDDLLCPADATRHDGGEIALIASCATRIIANSSFSWWGGWLG APSSPAHPVIAPSTWFADGHSDAAELVPRDWVRL	216
Prevotella salivae;Prevotella salivae DSM 15606	WP_00713 3870.1	494220705	alpha-1,2-fucosyltransferase [Prevotella salivae]	23.59		MIATTLFGGLGNQMFIYATAKALSLHYRTPMAFNLRQGFEQDYKYQRHLELNHFKCQLPTAKWITFNYKGE LNIKRISRRIGRNILCPHYQFIKEKEPFHYEKRLFEFTNKNIFLEGYWQSPRYFENYSDEIRRDFQLKSILPHTITD ELQMLKGTGKPLVMLGIRRYQEVKDKKDSPYPLCNKDYYAKAISHVQEQLPAPLFVVFTQEQAWAMNNLP TNANLYFVKEKDNAWATIADMYLMTQCQHAIISNSTFYWWGAWLQHPIENHIVVAPNNFFINRDCVCDN WIILD	217

Carnobacterium sp. WN1359	YP_008718687.1	554649641	glycosyl transferase [Carnobacterium sp. WN1359]	23.57		MIFVDLSEGLGNQMFQYAYSRYLQELYGGTLYLNTSSFKRKNSTRSYSLNNFYLYENVKLPSKFRRVIYNFYSK TIRMFIKKVIRMNPYSDKYYFSMIPYGFYVSSQVFKYLTVPTTKRHNIFVMGTWQTNKYFQSINDKIKDELKV KTEPNELNKKLITEINSNQSVCVHIRLGDYTNPEFDYLHVCTSDYYLKGMDYIVSKVKEPNFYIFSNSSSDIEWI KNNYNFKYKVKYIDLNNPDFEDFRLMYNCKHFIISNSTFSWWAQFLSNNDKKIIVAPSKWQKSNENEAKDIY LDHWKLIEIE	218
Butyrivibrio sp. AD3002	WP_022762290.1	551018062	glycosyl transferase [Butyrivibrio sp. AD3002]	23.55		MLIIQIAGGLGNQMQQYAMYRKLLKAGADRNIKLDTKWFDEDKQSGVLAKRKLELEYFTGLPLPVCSESERA RFTDRSVARKVVEKLVPGMGSRFTESCMYHPEIFELKDKYIEGYFACQKYYDDIMGELQELFVFPTHPDEEINI KNMINLMNEMEMVPSVSVHIRRGDYLDPENAALFGNIATDAYYDSAMEYFKAIDPDTHFYIFTNDPEYAREK YADPGRYTIVDHNTGKYSLLDIQLMSHCRGNICANSTFSFWGARLNRRKDKIPVRTLVMRNNQPVTPELMIH EYWPGWVLVDKDGKVR	219
Clostridium sp. KLE 1755	WP_021636935.1	545396682	glycosyltransferase, family 11 [Clostridium sp. KLE 1755]	23.55		MIVIRVMGGLGNQMQQYALYEKFKALGKETRLDTSWFDNASMQENVLARRSLELRFFDNLTYEACTPQER EALLGKEGFFNKLERKLFPSKNKHFYESEMFHPEIFKLDNVYLEGHWACEKYYHDIMPLLQSKIIFPKTDNIQN NMLKNKMNSENSVSIHIRRGDYLDPENAAMFGGICTDSYYKSAEGYIRNRVTNPHFYLFSDDPAYLREHYKG EEYTVVDWNHGADSFYDMELMSCCKHNVCANSTFSFWGARLNRTEKKIVIRPAKHKNSQQAEPERMHEL WENWVIIDEEGRIV	220
Bacteroides;Bacteroides vulgatus ATCC 8482;Bacteroides dorei DSM 17855;Bacteroides massiliensis dnLKV3	YP_001300694.1	150005950	glycosyl transferase family protein [Bacteroides vulgatus ATCC 8482]	23.47		MKFFVFGGGLGNQLFQYSYYRYLKKKYPSERILGIYPDSLKAHNGIEIDKWFDIELPPTSYLYNKLGILLYRVNRF LYNHGYRLLFCNRVYPQSMKHFFQWGDWQDYSIIKQINIFEFRSELPIGKENMEFLKKMETCNSISVHIRRG DYLKTDLIHIYGGICTSKYYREAIKFMEQEVEEPFFFFSDDCLYVETEFADIRNKIIISHNRDDRSFFDMYLMAH AKNMILANSTFSCWAAYLNRTAKIIITPDRWVNTDFSKLEALPNEWIKIRV	221
Paraprevotella xylaniphila;Paraprevotella xylaniphila YIT 11841	WP_008626629.1	495902050	glycosyl transferase [Paraprevotella xylaniphila]	23.47		MRLIKMTGGLGNQMFIYAMYLKMRAVFPDTRIDLSDMVHYRVHYGYEMNKVFNLPRTEFRINRSLKKIIEFL LFKTILERKQGGSLVPYIRKYHWPWIYFKGFYQSEEYFAGVEKEVREAFVFDVRRVNRKSLCAMQEIMADPD AVSIHVRRGDYLQGKHWKSLGCICQRSYYLNALSELEKRIVHPHYYVFSEDLDWVRQYLPLENAVFIDWNKG EDSWQDMMLMSHCRHHIICNSTFSWWGAWLNPSPDKIVIAPERWTQTTNSADVVPESWLKVSIG	222
Thauera sp. 28	WP_002930798.1	489020296	glycosyl transferase family protein [Thauera sp. 28]	23.47		MTDRALIAIVKGGLGNQLFIYAAARAMALRTGRQLYLDAVRGYLADDYGRSFRLNRFPIEAELMPEQWRVA STLRHPRAKLVRALNKYLPEAWRFYVAERGDTRPGALWNHGRNVKRVTLMGYWQDEAYFLDYAELLRREL GPPMPDAPEVRARGERFAGTESVFLHVRRCRYSPLLDAGYYQKAVDLACAELNKPVFMIFGDDIEWVVNNI DFRGAGYERQDYDESDELADLWLMTRCRHAIIANSSFSWWAAWLGGAAGSGRHVWAPGQSGLALKCAK SWEAVDAQPE	223

Subdoligranulum variabile;Subdoligranulum variabile DSM 15176	WP_00704 8308.1	494107522	alpha-1,2-fucosyltransferase [Subdoligranulum variabile]	23.44		MIYAELAGGLGNQMFIYAFARALGLRCGEAVTLLDRQDWRDGAPAHTACALEGLNLVPEVKILAEPGFAKR HLPRQNTAKALMIKYEQRQGLMARDWHDWERRCAPVLNLLGLHFATDGYTPVRRGPARDFLAWGVFQS EAYFADFAPTIRAELRAKQAPAGVVAEKIRAAACPVALHLRRGDYCRPENEILQVCSPAYYARAAAAAAAAY PFATLFVFSDDIDWAKEHLDTAGLPAVVWMPRGDAVGDLNLMALCRGFILSNSTYSWWAAQYLAGEGRTV WAPDRWFAHTKQTALYQPGWHLIETR	224
Firmicutes bacterium CAG:24	WP_02191 6223.1	547127527	protein [Firmicutes bacterium CAG:24]	23.4		MIIVEVMGGLGNQMQQYALYRKLESLGKDARLDVSWFLDKERQTKVLASRKLELSWFENLPAKYCTQEEK QAILGKNNLIGKLKKKLLGGSNRHFTESDMYHPEIFDLEDAYLSGFWACEAYYADILPMLRSQIHFPDPEKGE GWDLEAAAKNKETMERMKQETSVSIHIRRGDYLDAKNAEMFGGICTDAYYEAAISYIKEQTPDAHFYVFSD DSAYVKNAYPGKEFTVVDWNTGKNSLFDMQLMSCCNHNICANSTFSFWGARLNPSPDKVMIRPSKHKNS QNIVPEEMKRLWDGWVLIDGKGRII	225
Prevotella sp. CAG:474	WP_02231 0139.1	547906803	glycosyl transferase family 11 [Prevotella sp. CAG:474]	23.39		MIITKLNGGLGNQLFEYACARNLQLKYNDVLYLDIEGFKRSPRHYSLEKFKLSSDVRMLPEKDSKSLILLQAISK LNRNLAFKLGPLFGTYIWKSSNYRPLKIKNTRGKKLYLYGYWQSYEYFKENEAIIKQELNVKTEIPIECSELLKEIN KPHSICVHVRRGDYVSCGFLHCDEAYYNRGINHIFDKHPDSNVVVFSDDIKWVKANMNFDHPVAYVEVDV PDYETLRLMYMCKHFVMSNSSFSWWAASYLSDNKEKIVVAPSYWLPANKDNKSMYLDNWTIL	226
Roseburia intestinalis;Roseburia intestinalis L1-82	WP_00685 5899.1	493910390	glycosyl transferase family 11 [Roseburia intestinalis]	23.38		intestinalis]MRGNRGMIAVKIGDGMGNQLFNYACGYAQARRDGDSLVLDISECDNSTLRDFELDKFHLKY DKKESFPNRNLGQKIYKNLRRALKYHVIKEREVYHNRDHRYDVNDIDPRVYKKKGLRNKYLYGYWQHLAYFE DYLDEITAMMTPAYEQSETVKKLQEEFKKTPTCAVHVRGGDIMGPAGAYFKHAMERMEQEKPGVRYIVFT NDMERAEEALAPVLESQKKDAYGQAENRLEFVSEMGEFSDVDEFFLMAACQNQILSNSTFSTWAAYLNQN PDKTVIMPDDLLSERMRQKNWIILK	227
Bacteroides ovatus;Bacteroides ovatus ATCC 8483	WP_00429 6622.1	490424433	protein [Bacteroides ovatus]	23.29		MKIVLFTPGLGNQMFQYLFYLYLRDNYPNQNIYGYYNRNILNKHNGLEVDKVFDIQLPPHTVISDASAFFIRA LGGLGLKYFIGKDQLSPWKVYFDGYWQNKEYFQNNVDKMRFREGFLNKKNDDILSLIRNTNSVSVHVRRG DYCDSCRKDLFLQSCTPQYYESAISVMKEKFQKPVFFVFSDDIPWVKVNLNIPNAYYIDWNKKENSYLDMYL MSLCTASIIANSTFSFWGAMLGNKKELVIKPKKWIGDEIPEIFPPSWLSL	228
Butyrivibrio sp. AE3009	WP_02277 9599.1	551035785	glycosyl transferase [Butyrivibrio sp. AE3009]	23.25		MLIIQIAGGLGNQMQQYALYRKLLKYHPDGVRLDLSWFDSEVQKNMLAKREFELALFKGLPYIECKPEERAA FLDRNAAQKLSGKVLKKLGLRDNANPNVFEESRMFHPEIFELDNKYIIGYFACQKYYDDIMGDLCNLFEFPEH LDPELEKKNLELISKMEKENSVSVHIRRGDYLDPENFKILGNIATDEYYESAMKYFEDRYEKVHFYIFTSDHEYA REHFADESKYTIVDWNTGKDSLQDVRLMNHCLGNICANSTFSFWGARLNQRQDKVMIRTYKMRNNQPV DPDTMHDYWKGWILIDETGREV	229

Butyrivibrio proteoclasticus;Butyrivibrio proteoclasticus B316	YP_003829712.1	302669752	glycosyl transferase 11 [Butyrivibrio proteoclasticus B316]	23.23		MTKNEKKLIVKFQGGLGNQLYEYAFCEWLRQQYSDYEVLADLSYYKIRSAHGELGIWNIFPNINIEVASNWDI IKYSDQIPIMYGGKGADRLNSVRTNVNDRFFSKRKHSYYTEISNTDVSEVINALNNGIRYFDGYWQNIDYFKG NIEDLRNKLKFSEKCDKYITDEMLRDNAVSLHVRRGDYVGSEYEKEVGLSYYKKAVEYYLDRVDQAKFFIFSD DKYYAETAFEWIDNKTVVAGYDNELAHVDMLLMSRMKNNIIANSTFSLWAAYLNDSMNPLIVYPDVESLD KKTFSDWNGIK	230
Prevotella nanceiensis	WP_018362656.1	517173838	protein [Prevotella nanceiensis]	23.23		MDSQFLKHIKLSGGFGNQLFQYFFGEYLKEKYNCSISFFSEPALDINQLQIHRFFPALRISHNTELRPYHYSFTQ QLAYRCMRKLLLFPFLNRKVKIENGSNYQNQSFNDTYCFDGYWQSYRYLSAFTPSLQFEDQLINDISADYIN AIECSEAVFLHIRRGDYLNKENQKVFAECPLNYFENAANRIKEDIKNVHFFVFSNDIQWVKSHLKLNDNEVTF IQNEGNSCDLKDFYLMTRCKHAIISNSTFSWWAAYLINNSDKKVIAPKHWYNDISMNNATKDLIPPTWIRL	231
Ruegeria sp. R11	WP_008562971.1	495838392	alpha-1,2-fucosyltransferase [Ruegeria sp. R11]	23.23		MITRLHGRLGNQMFQYAAGRALADRAGVPLALDSRGAILRGEGVLTRVFDLELADPVHLPPLKQTNPLRYA IWRGIGQKVGAKPYFRRERGLGYNPAFEDWGDNSYLHGYWQSQKYFQNSAERIRSDFTFPAFSNQQNAE MAARIAESTAISLHVRRGDYLTFAAHVLCDQAYYDAALAKVLDGLQGDPIVYVFSDDPQWAKDNLSLPCEK VVVDFNGPETDFEDMRLMSLCQHNIIGNSSFSWWAAWLNQTPGRRVAGPAKWFGDPKLSNPDDIFPHDW LRISV	232
Winogradskyella psychrotolerans RS-3;Winogradskyella psychrotolerans	WP_020895733.1	527072096	alpha-1,2-fucosyltransferase [Winogradskyella psychrotolerans]	23.21		MGNQLYEYATAKAMAVALNKKLVIDPRPILKEAPQRHYDLGLFNIQDEDFGSPFVQWLVRWVASVRLGKFF KTIMPFAWSYQMIRDKEEGFDESLLQQKSRNIVIEGYWQSFKYFESIRPTLLKELSFKDKPNAINQKYLDEIESV NAVAVHIRRGDYVANPVANAVHGLCDMDYYKKAIAIIKDKVENPYFFIFTDDPDWAEDNFKISEHQKIIKHNI GKQDHEDFRLLTNCKYFIIANSSFSWWGAWLSDYKNKIVISPNKWFNVDAVPITERIPESWIRV	233
Lachnospiraceae bacterium NK4A179	WP_022785342.1	551041720	protein [Lachnospiracea e bacterium NK4A179]	23.2		MITVRIDGGFGNQMFQYAFFLHLKKTITDNKISVDLNCYNPHGSGDIFTRFKLAPEQAAPSEIKRFHRNSIYHL LRPLDSAGITTNPYYREEDIDDLNSVLNKKRVYLRGYWQDKRYPFSVKDQLIDCFDLGKKMDMTGASAENNVI LEQIASEESRSVGVHLRGGDYIGDPVYSGICTPEYYEAAFKHVSEKIKDPVFHIFTNDISMIEKCGLSGKYDLKIT DINDEAHGWADLKLMSACRHHIISNSSFSWWAAFLGEATTEASADVINVIPEYMRQGVSAETLRCPCWTT VTSDGRVYPS	234
Prevotella sp. oral taxon 317 str. F0108;Prevotella sp. oral taxon 317	WP_009230832.1	496522549	alpha-1,2-fucosyltransferase [Prevotella sp. oral taxon 317]	23.13		MKIVCIKGGLGNQLFEYCRYRSLHRHDNRGVYLHYDRRRTKQHGGVWLDKAFHITLPNEPLRVKLLVMVLK TLRRLHLFKRLYREEDPRAVLIDDYSQHKQYITNAAEILNFRPFEQLDYAEEIQTTPFAVSVHVRRGDYLLLANK SNFGVCSVHYYLSAAVAVRERHPESRFFVFSDDMEWAKENLNLPNCVFVEHAQAQPDHADLYLMSLCKGH IIANSTFSFWGAYLSKGSSAIAIYPKQWFAEPTWNVPDIFPAHWMAL	235

Butyrivibrio sp. XPD2006	WP_02276 5796.1	551021633	glycosyl transferase [Butyrivibrio sp. XPD2006]	23.1		MLIIQIAGGLGNQMQQYAVYTKLRGMGKDVRLDLSWFDPSVQKNMLAPREFELSMFEGVDYTECTAEER DSFLKQGMIANVTGKMLKKLGLRDEANPKVFSEKEMYHPEIFELEDRYIKGYFACQKYYDDIMGELWEKYTF PAHSDPDLHTRNMALVERMEKETSVSVHIRRGDYLDPSNVEILGNIATEEYYQGAMDYFSVKDPDTHFYIFT SDHEYARKKFSDESKYTIVDWNSGRNSVQDLVLMLMSHCKGNICANSTFSFWGARLNRRPDKTVIRTYKMRN NQPVNPDIIMHDYWKGWILMDEKGSII	236
Butyrivibrio fibrisolvens	WP_02275 2717.1	551008140	protein [Butyrivibrio fibrisolvens]	23.08		MIIIKLQGGLGNQLFLYGLYKNLKHLKRDVKMDIESGFEEDKLRVPCLKSMGLDYEVATRDEIVAIRDSYMDIF SRIRRKITGRKTFDYYEPEDGNFDPRVLEQTHAYLDGYFQSEKYFGDSDDRKKLKDELLKEKIRVLDSSDTLKDL YNMMSSGSSVSLHIRRGDYLTPGIMETYGGICTDEYYDIAMNRIKNEYPDSKFFIFSNDIDWCKEKYGSRDDV IFVDSCDEHEGLTNVSGIDQDDICVQGDIKEHGNNSLRDAAEELYLMSACKHHILANSSFSWWGAWLSDHEG MTIAPSKWLNNKNMTDIYTKDMLLI	237
Cylindrospermopsis raciborskii;Cylindrospermopsis raciborskii CS-505	WP_00627 8973.1	493321658	Glycosyl transferase family 11 [Cylindrospermopsis raciborskii]	23.05		MKKTVVLLKGGLGNQMFQYAFARSISLKNSSKLVIDNWSGFTFDYKYHRQYELGTFSIVGRPANLTEKFPFW FYELKSKFFPRLPKVFQQQFYGLLINEVGGEYIPEIEETKISQNCWLNGYWQSPLYFQKHSDSITRELMPPEPM EKHFLELGKLLRETESVALGIRLYEESKNPGSHSSSGELKSHFEINQAILKLRELCNGAKFFVFCTHRSPLLQELAL PENTIFVTHDDGYVGSMERMWLLTQCKHHIFTNSTFYWWGAWLSQKFYIQGSQIVFAADNFINSDAIPKH WKPF	238
Prevotella multiformis;Prevotella multiformis DSM 16608	WP_00736 8154.1	494609908	alpha-1,2-fucosyltransferase [Prevotella multiformis]	23.05		MKIVNFQGGLGNQMFIYAFSRYLSRLYPQEKIYGSYWSRSLYVHSAFQLDRIFSLQLPPHNLFTDCISKLARFF ERLRLVPVEETPGSMFYNGYWLDKKYWEGIDLSEMFCFRNPDLSAEAGAVLSMIERSNAVSVHIRRGDYQS EEHIEKFGRFCPPDYYRIATERIRQREDDPLFFVFSDDMMWVKSNMDVPNAVYVDCHHGDDSWKDMFL MAKCRHNIIANSTFSFWAAMLNANPDKVVVYPQRWFCWPSPDIFPEMWLPVTEKEIKSSF	239
Bacteroides sp. CAG:462	WP_02238 4635.1	548151455	protein [Bacteroides sp. CAG:462]	23		MIIVNMACGLANRMFQYAFYLSLKERGYNVKVDFYKSATLPHENVPWNDIFPYAEIDQVSNFRVLILGGGA NLLSKLRRKYLPSLTNVITMSTAFDTDLQIDDDRKDKYIIGVFQSAAMVEGVCKKVKQCFSFLPFTDLRHLQLE KEMQECESVAIHVRKGNDYQQQRIWVQNTCTMDYYRKAIAEIKGKVKDPRFYVFTDNADWVRRNFTDFDYK MVEGNPVYGWGSHFDMQLMSRCKYNISNSTYSWWGAYLNANRNKIVICPNIWFNPESCNEYTSCKLLCK GWIAL	240

Desulfovibrio africanus;Desulfovibrio africanus PCS	WP_00598 4176.1	492830222	Glycosyl transferase family 11/Glycosyltransferase family 6 [Desulfovibrio africanus]	23		MRIGILYICTGKYTVFWNHFFTSCEQHFLREHEKHYYIFTDGEIAHLNCNRVHRIEQQHLGWPDSTLKRFHM FERIADTLRQNSDFIVFFNANMVFLRDVGKEFLPTREQALVFHRHIPGLFRRPAWLLPYERRPESTAYIPYGSG SIYVCGGVNGGYTQPYLDFVAMLRRNIDIDVERGIIARWHDESHINRFVIGRHYKIGHPGVVYPDRRNLPFPR IIRVIDKASVGGHTFLRGQTPEPAPEEQSKTVAKKLRSQLKRPCMPRAAQDEPIILARMMGGLGNQMFIYAA ARVLAERQGAQLHLDTGKLSGDSIRQYDLPAFSIDAPLWHIPCGCDRIVQAWFALRHVAAGCGMPKPTMQ VLRSGFHLDQRFFSIRHSAYLIGYWQSPHYWRGHEDRVRSSFDLTREFRPHLREALAAVSQPNTISVHLRRG DFRAPKNSDKHLLIDGSYYERARKLLLEMTPQSHFYIFSDEPEEAQRLFAHWENTSFQPRRSQEEDLLLMSRC SASIIANSSFSWWGAWLGRPKQHVIAPRMWFTRDVLMHTYTLDLFPEKWILL	241
Roseburia sp. CAG:100	WP_02251 8697.1	548374190	protein [Roseburia sp. CAG:100]	22.98		MILIHVMGGLGNQLYQYALYEKMKSLGKKVKLDTYAYNDAAGEDKEWRSLELDRFPAIEYDKATSEDRTKLL DNSGLLTAKIRRKLLGRKDKTIRESKEYMPEIHHMDDVYLYGFWNCERYYEDIIPLLQDKLQFPISNNPRNQQ CIEQMQKENAVSIHIRRTDYLTVADGARYMGICTEDYYKGAMAYIEERVSNPVYYIFSDDVEYAKQHYHQD NMHVVDWNSKADSIYDMQLMSKCKHNICANSTFSMWAARLNQNKEKIMIRPLHHDNYETTTATQVKQN WKNWILLDQNGQVCE	242
Lachnospiraceae bacterium 10-1	WP_02274 2385.1	550997676	protein [Lachnospiraceae bacterium 10-1]	22.96		MTMNIIRMSGGLGSQMFQYALYLKLKSMGKEVKFDDINEYRGEKARPIMLAVFGIEYPRATWDEITSFTDG SMDLLKRLRRKIFGRKAIEYEEQGFYDPNVLNFDSMYLRGNFQSEKYFQDIKEEVRKLYRFSTLEDMRLPERLY KATKACLDGIESSESVGLHMYRSDSRVDGELYDGICTGNYYKGAVRFIQDKVPDAKFYIFSNEPKWVRGWVV DLIQSQIQEGMSPSQVKEMEKRFVMVEANTEYTGYLDMMMLMSKCKHNIISNSSFSWWSAWMNDHPEKV VVAPDRWSSDKEGNEIYTTGMTLVNEKGRVNYTIHENSTVK	243
Prevotella nigrescens;Prevotella nigrescens F0103	WP_00436 2670.1	490496500	protein [Prevotella nigrescens]	22.96		MILSYITGRLGNQLFEYAYARSLLLKRGKNEELILNFSLVRAAGKEIEGFDDNLRYFNVYSYTELDKDIVLSKGDL LQLFIYILFKLDQKLFRIIKKEKWFSFFRRFGIIFQDYLDNISNLIIPRTKNVFCYGKYENPKYFDDIRSILLKEFTPRI PPLKNNDQLYSVIESTNSVCISIRRGDFLCDKFKDRFLVCDKEYFLEAMEEAKKRISNSTFIFFSDDIEWVRENIH SDVPCYYESGKDPVVWEKLRLMYSCKHFIISNSTFSWWAQYLSRNEEKVVIAPDRWSNVPGEKSFLLSNSFIKI PIGILP	244
Bacteroides sp. CAG:875	WP_02235 3235.1	547952493	fucosyl transferase [Bacteroides sp. CAG:875]	22.95		MIYVEINGRLGNNMFEIAAAKSLTDEVTLWCKGDWQLNCIKMYSDTLFKNYPIVKSLPNNIRIYEEPEFTFHPI PYKENQDLLIKGYFQSYKYLDREKVLKLYPCPMPVKLDIEKRFGDILSQYTVVSINVRRGDYLNLPHIRHPFVGK KFLERAMLWFGDKVHYIISSDDIEWCKAHFKQFDNVHYLTNSYPLLDLYIQTACHHNIISNSSFSWWGAYLN NHPQKIVIAPHRWFGMSTNINTQDLLPPEWMIEQCVYEPKVFLKALPLHAKYLLKRVLK	245

Prevotella sp. oral taxon 299 str. F0039;Prevotella sp. oral taxon 299	YP_008444280.1	532354444	protein HMPREF0669_00176 [Prevotella sp. oral taxon 299 str. F0039]	22.9		MDSQLLKHIKLSGGFGNQLFQYFFGEYLKEKYNCSISFFSEPALDINQLQIHREFPTLRISHNTELRREHYAFTQ QLAYRCMRKLLLLFPFLNRKVKIENGSNYQNQSFNDTYCFDGYWQSYRYLSAFTPSLQFEDQLINDISADYIN AIEQSEAVFLHIRRGDYLNKENQKVFAECPLNYFENAVNKIKEGNKTYHIFFVFSNDIEWVKCHLKLNNNEVTF IQNEGSSCDLKDFYLMTRCKHAIISNSTFSWWAAYLINNNDKKVIAPKRWYNDLSMNNATKDLIPPTWIRL	246
Paraprevotella xylaniphila;Paraprevotella xylaniphila YIT 11841	WP_008628783.1	495904204	alpha-1,2-fucosyltransferase [Paraprevotella xylaniphila]	22.87		MKIYCLKGGLGNQMFEYCRFRDLMDSGNGKVYLFYDRRRLKQHDGLRLSDCFELELPSCPWGIRLVVWGL KICRAIGVLKRLYDDEKPDAVLIDDYSQHRRFIPNARRYFSFRQFLAELQSGFVQMIRAVDYPVSVHVRRGDY LHPSNSSFVLCGVDYFRQAIAYVRKKRPDARFFFFSDDMEWVRENLWMEDAVYVEHTELMPDYMDLYLM TLCRGHIISNSTFSFWGAYLAVDGNGMKIYPRRWFRDPTWITPPIFSEEWVGL	247
Dethiosulfovibrio peptidovorans;Dethiosulfovibrio peptidovorans DSM 11002	WP_005658864.1	491897177	glycosyl transferase family 11 [Dethiosulfovibrio peptidovorans]	22.84		MFQYAFGRALALDLGLDLKLDISNFGSDSRPFSLGIYSLTKNIPFGCYLSTSTRLKVKMTKKLRRWGVWGMD KNMPGVLVEPFPPVLVSLDEVLSEKLSHLFVDGYWQSEKYFSRYSDVIRSDFRVIEESSAFLAWKKRMLSEPG GSISVHVRRGDYVTDSSAANRVHGVLPIEYYLRAKEILNTISDGLVFYVFTDDPVWARNNLCLGDKTIYVSGEDL KDYEELALMSCCDHHVVANSSFSWWGAWLGQDTSTVTIAPGRWFRKMDSSFVIPDNWIKIWT	248
Lachnospiraceae bacterium 10-1	WP_016229292.1	510896192	protein [Lachnospiraceae bacterium 10-1]	22.83		MIIIQVMGGLGNQLQQYALYRKFVRMGKEARLDISWFLDKEKRGEVLAERELELDYFDRLIYETCTPEEKEQLI GSEGVAGKLKRKFLPGRIRWFHIESKIYHPELLQMENMYLSGYFACEKYYADILYDLREKIQFPVNDHPKNIKM AQEMQERESVSVHLRRGDYLDEKNTAMFGNICTDAYYCKAIEYMKTLCSKPHFYIFSDDIPYVRQRFTGEEYT VVDINHGRDSFFDMWLMSRCRHNICANSTFSFWGARLNSNDNKIMIRPTIHKNSQVFVKEEMEQLWPG WKFISPDGGIK	249
Treponema maltophilum;Treponema maltophilum ATCC 51939	WP_016525279.1	513872223	protein [Treponema maltophilum]	22.82		MFCAAFVEALKHAGQKVFVDTSLYNKGTVRSGIDFCHNGLETEHLFGIKFDEADKADVHRLSTSAEGLLNRIR RKYFTKKTHYIDTVERYTPEVLSDKSDRYLEGFWQTEKYFLPIESDIRTLFRFRQPLSEKSAAVQSALQAQEPAS LSASIHVRRGDFLHTKTLNVCTETYYNNAIEYAAKKYAVSAFYVFSDDIQWCREHLNFFGARSVFIDWNIGAD SWQDMVLMSMCRCNIIANSSFSWWAAWLNAASDKIVLAPAIWNRRQLEYADRYYGYDYSDVIPETWIRIP I	250
Bacteroides massiliensis;Bacteroides massiliensis dnLKV3	WP_016276676.1	511022363	protein [Bacteroides massiliensis]	22.79		MKLVSFTAGLGNQLFQYCFYRYLLNKFPNEKIYGYYNKKWLKKHGGIIIEHFFDVKLPRSTRWINLYGQYLRIIY KCFSCGVSKDDDFEMNRTIMFVGYWQDQCFFSGINISYKKNLVISEKNTWLLGEILKCNSYAIHFRRGDYMLP QFKKIFGEVCTVKYYLKSIRKVEEKISEPVFFVFSDDIDWVKQNFTFNKVYFVDWNKGQNSFWDMYLMSQC SANIIANSTFSFWGAYLNKNNPFVIYPQKWVRTNLKQPNIFPKTWMAL	251

Enterococcus faecium;Enterococcus faecium DO;*Enterococcus faecium* EnGen0035	YP_006376560.1	389869137	family 11 glycosyltransferase [Enterococcus faecium DO]	22.71		MIVLTLGGGLGNQMFQYGYARYIQKIHREKFIYINDSEVIKEADRFNSLGNLNTVNIKVLPRIISKPLNETERLV RKIMVRLFGVAGFNESAIFQSLNKFGIYYHPSVYKFYESLKTGFPIKIIEGGFQSWKYLETCPEIKQELRVKYEP MGENLRLLNLISQSESVCVHIRRGDYLSPKYKHLNVCDYQYYFESMNYIISKLNNPTFFIFSNTSDDLDWIKEN YSLPGKIVYVKNDNPDYEELRLMYSCKHFIISNSTFSWWAQYLSNNSGIVIAPEIWNRLNHDGIADLYMPNW ITMKVNR	252
Bacteroides;Bacteroides sp. 2_1_22;Bacteroides sp. 2_2_4;Bacteroides sp. D1;Bacteroides xylanisolvens SD CC 2a;Bacteroides xylanisolvens SD CC 1b;Bacteroides ovatus CAG:22	WP_004313284.1	490442319	glycosyl transferase family 11 [Bacteroides]	22.67		MDVVVIFNGLGNQMSQYAYYLAKKKVNPNTKVIFDIMSKHNHYGYDLERAFGIEVNKTLLIKVLQIIYVLSRK FRLFKSVGVRTIYEPLNYDYTPLLMQKGPWGINYYVGGWHSEKNFMNVPDEVKKAFMFREQPNEDRFNE WLQVIRGDNSSVSVHIRRGDYMNIEPTGYYQLNGVATLDYYHEAIDYIRQYVDTPHFYVFSNDLDWCKEQF GVENFFYIECNQGVNSWRDMYLMSECHYHINANSTFSWWGAWLCKFEDSITVCPERFIRNVVTKDFYPER WHKIKSC	253
Synechococcus phage S-SM2	YP_004322362.1	326781960	glycosyltransferase family 11 [Synechococcus phage S-SM2]	22.6		MIGFNALGRMGRLANQMFQYASLKGIARNTGVDFCVPYHFEAVNDGIGNMLRTEIFDSFDLQVNVGLLNK GHAPVVQERFFHFDEELFRMCPDHVDIRGYFQTEKYFKHIEDEIREDFTFKDEILNPCKEMIAGVDNPLALHV RRTDYVTNSANIHPPCTLEYYEAALKHFDDDRNVIVFSDDPAWCKEQELFSDDRFMISENEDNRIDLCLMSLC DDFIIANSTYSWWGAWLSANKDKKVIAPVQWFGTGYTKDHDTSDLIPDGWTRIATA	254
Geobacter metallireducens;Geobacter metallireducens GS-15;Geobacter metallireducens RCH3	YP_006720295.1	404496189	glycosyltransferase [Geobacter metallireducens GS-15]	22.58	22.58	MDIHVLSYGLGNQLSQYAFFINRRQLMQRAYAFYAFKQHNGYELDRIFGLKEGLPWYLQFVRVVFRLGISRR FYSKRTADFVLSLFRIKVIDEAYNYEFDPSLLKPWFGIRILYGGWHDSRYFHPSEAAVRTAFSFPPLDDVNDAIL QQIDAVYGVSIHVRRGDYLKGINSNLFGGIATLEYYRNAIGWAITYCKHRSLEIKFYVFSDDIDWCKQNIGLR DAVYVSGNSKTDSWKDILLMSHCRANIIANSTFSWWAAWLNQQPNKVVICPTKFINTDSPNQTIYPAAWH QIEG	255
Lachnospiraceae bacterium NK4A136	WP_022780989.1	551037245	protein [Lachnospiraceae bacterium NK4A136]	22.58		MIIVRFHGGLGNQMFEYAFYRYMTNKYGADNVIGDMTWFDRNYSEHQGYELKKVFDIDIPAIDYKTLAKIH EYYPRYHRFAGLRYLSRMYAKYKNKHLKPTGEYIMDFGPSQYIHNDAFDKLDTNKDYYIEGVFCSDAYIKYYE NQIKKDLTFKPNYSQHTKDMILPKIEETNSVAIHVRRGDYYGNVFDIVTPDYYRQAVNYIRERVENPVFFVFSD DMDYIKANFDFLGDFVPVHNCGKDSFQDMYLISRCRHMIIANSSFSYFGALLGEKDSTIVIAPKKYKADEDLA LAREWVLL	256

Bacteroides caprophilus;Bacteroides coprophilus DSM 18228 = JCM 13818	WP_008144634.1	495419937	protein [Bacteroides coprophilus]	22.56		MGFIVNMACGLANRMFQYSYYLFLKKQGYKVTVDFYRSAKLAHEKVAWNSIFPYAEIKQASRLKVFLWGGG SDLCSKVRRRYFPSSTNVRTTTGAFDASLPANTARNEYIIGVFLNASIVEAVDDEIKKCFTFLPFTDEMNLRLKK EIEECESVAIHVRKGKDYQSRIWYQNTCSMEYYRKAILQMKEKLQHSKFYVFTDNVDWVKENFQEIDYTLVE GNPADGYGSHFDMQLMSLCKHNIISNSTYSWWSAFLNRRNPEKVVIAPEIWFNPDSCDEFRSDRALCKGWI VL	257
Bacteroidetes;Capnocytophaga sp. oral taxon 329 str. F0087;Paraprevotella clara YIT 11840	WP_008619736.1	495895157	alpha-1,2-fucosyltransferase [Bacteroidetes]	22.53		MKIVCLKGGLGNQMFEYCRFRDLMESGHDEVYLFYDHRRLKQHNGLRLSDCFELELPSCPWGIKLVVWGLK ICRAVGVLKRLYDDEKPEAVLIDDYSQHRRFIPNARRYFFFRQFLAELQSGFVQMIRAVDVPVSVHVRRGDYL HPSNSSFGLCGVDYFQQAIAYVRKKRPDARFFFFSDDMEWVRENLWMEDAVYVEHTELLPDYVDLYLMTL CRGHIISNSTFSFWGAYLAVDGNGMKIYPRRWFRDPTWTSPPIFSEEWVGL	258
Butyrivibrio sp. NC2007	WP_022770361.1	551026242	glycosyl transferase [Butyrivibrio sp. NC2007]	22.47		MLIIQIAGGLGNQMQQYAVYTKLREMGKDVKLDLSWFDPQVQKNMLAPREFELPIFGGTDYEECSAYERD ALLKQGAFAAIAGKVLKKLGLRDEANPKVFSEKEMYHPEVFEIEDKYIKGYFACQKYYGDIMDKLQEKFIFPE HSDPDLHARNMALVERMEREPSVSVHIRRGDYLDPSNVEILGNIATEQYYQGAMDYFTVKEPDTHFYIFTSD HEYAREKFSDESKYTIVDWNNGKNSVQDLMLMSHCKGNICANSTFSFWGARLNKRPDKTVIRTYKMRNN QPVNPQIMHDYWKGWILMDEKGSII	259
Paraprevotella xylaniphila;Paraprevotella xylaniphila YIT 11841	WP_008628536.1	495903957	glycosyl transferase [Paraprevotella xylaniphila]	22.45		MKILVFTGGLGNQMFAYAFYLYLKRLFPQERFYGLYGKKLSEHYGLEIDKWFKVSLPRQPWVVLPVTGLFYL YKQCVPNSKWLDLNQEICKNPRAIVFFPFKFTKKYIPDDNIWLEWKVDESGLSEKNRLLLSEIRSSDCCFVHVR RGDYLSPTFKSLFEGCCTLSYYQRALKSMKEISPFVKFVCFSDDIQWVKQNLELGNRAVFVDWNSGTDSPLD MYLMSQCRYGIMANSTFSYWGARLGRKKKRIYYPQKWWNHGTGLPDIFPNTWVKI	260
Blautia hydrogenotrophica DSM 10507;Blautia;Blautia hydrogenotrophica CAG:147	WP_005944761.1	492742598	protein [Blautia]	22.44		MEIHVYLTGRLGNQLFQYAFARHLQKEYGGKIICNIYEIEHRSEKAAWVPGKFNYEMSNYKLNDSILIEDIKLP WFADFSNPIIRIVKKVIPRIYFNLMASKGYLLWQKNSYINIPAIRNNEIIVNGWWQDVRFFHDVEAELSNEIVP TTKPISENEYLYNIAERENSVCVSIRGGNYLVPKVKKKLFVCDKEYFYNAIEIIKSKVRNAIFIVFSDDLEWVKSYI KLEEKFPECKFYYESGKDTVEEKLRMMTKCKHFIISNSSFSWWAQYLAKNENKIVIAPDAWFTNGDKNGLYI DDWILIPTQTKDM	261
Geobacter lovleyi;Geobacter lovleyi SZ	YP_001952981.1	189425804	glycoside hydrolase family protein [Geobacter lovleyi SZ]	22.44		MITVLLNGGLGNQLFQYAAGRALAEKHDVELLDLSRLQHPKPGDTPRCFELAPFNIKASLLAEEGRQPLGSY QACMHRLLLKASIPLWGSIILKEQGCGFDPLIFRAPSSCILDGFWQSECYFKQITSLLQQELSLKAPSPALRKAS SVLSDATVAVHVRRGDYVTNPAAASFHGICSQDYYQAAVANILTSYPDSQFLVFSDDPAWCQEHLDLGQPF RLAADFGLNGSAEELVLISRCAHQIIANSSFSWWGAWLNPSPHKLVVAPCRWFTDPAITTNDLLPETWVRLP	262

Lachnospiraceae bacterium NK4A136	WP_02278 1176.1	551037435	protein [Lachnospiracea e bacterium NK4A136]	22.41		MVISHLSGGFGNQLYSYAFAYAVAKARKEELWIDTAIQDAPWFFRNPDILNLNIKYDKRVSYKIGEKKIDKIFN RINFRNAIGWNTKIINESDMPNIDDWFDTCVNQKGNIYIKGNWSYEKLFISVKQEIIDMFTFKNELSKEANDI AQQINSQETSVGIHYRLGDYVKIGIVINPDYFISAMTSMVEKYGNPVFYSFSEDNDWVKKQFEGLPYNIKYVE YSSDDKGLEDFRLYSMCKHQIASNSSYSWWGAYLNNNPNKYIIAPTDYNGGWKSEIYPKHWDVRPFEFLK	263
Bacteroides vulgatus;Bacteroides vulgatus PC510	WP_00584 0359.1	492426440	glycosyl transferase family 11 [Bacteroides vulgatus]	22.37		MFHYKFLLFGGGLGNQIFEYFYLWLRKKYPNIVFLGCYRKASFKAHNGLEISDVFDVDLPNDGGLSGRFISYV LSVLSRIIPSLSMKANTEYSSKYLLINAYQPNLLFYLNEEKIKFRPFKLDEVNRRLLNSIKMESSVSIHVRRGDYLF GQYRDIYSNICTLAYYQKAVDKCKGILESPRFFVFSDDIEWARDVFVGREYEFVSNNIGKNSFIDMFLMSNCKI QIIANSTFSYWAAYLSNSLVKIYPAKWINGIERPNIFPDNWIGL	264
Planctomyces brasiliensis;Planctomyces brasiliensis DSM 5305	YP_004271 766.1	325110698	glycosyl transferase family protein [Planctomyces brasiliensis DSM 5305]	22.37		MIIARIENGLGNQLFKYAAGRALSLKHRTSLYTIPGSVRKPHETFILSKYFNVQAKSVSPFLLQTGFRLRLLKGYE NHSFGFDPRFETTRNNTVVSGNFQSARYFLPFFDQINRELTLKPEVVDGLESVYPHVLESLRTPNSVCVHIRLG DYVSSGYDICGPEYYAKAISRLQQLHGELRAFVFSDTPQAASRFLPADIDAQIMSEFPEVRDAARSLTVERSTI RDYFLMQQCRHFVIPNSSFSYWAALLSSSDGDVIYPNRWYIDIDTSPRDLGLAPAEWTPIPLT	265
Butyrivibrio sp. AE2015	WP_02277 2730.1	551028648	glycosyl transferase [Butyrivibrio sp. AE2015]	22.36		MIILQIAGGLGNQMQQYALYRKLLKCGKTVKLDLSWFGPEIQKNMLAPREFELVLFKDLPFEICTKEEKDALIK QNLFQKIAGKVSQKLGKSASSNAKVFVETKMYHEEIFDLDDVYITGYFACQYYYDDVMAELQDLFVFPSHSIP ELDQRNAVLASKMEKENSVSVHIRRGDYLSPENVGILGNIASDKYYESAMNYFLEKDENTIIFYIFTNDIEYAR EHYSDESRYTIIDWNTGKNSLQDLMLMSHCKGNICANSTFSFWGARLNKRPDRELVRTLKMRNNQEAQPEI MHEYWKNWILIDENGVIV	266
Roseovarius nubinhibens ISM;Roseovarius nubinhibens	WP_00981 3856.1	497499658	alpha-1,2-fucosyltransfera se, [Roseovarius nubinhibens]	22.34		MTDTPPPSQVITSRLFGGAGNQLFQYAAGRALADRLGCDLMIDARYVAGSRDRGDCFTHFAKARLRRDVA LPPAKSDGPLRYALWRKFGRSPRFHRERGLGVDPEFFNLPRGTYLHGYWQSEQYFGPDTDALRRDLTLTTAL DAPNAAMAAQIDAAPCPVSFHVRRGDYIAAGAYAACTPDYYRAAADHLATTLGKPLTCFIFSNDPAWARD NLDLGQDQVIVDLNDEATGHFDMALMARCAHHVIANSTFSWWGAWLNPDPDKLVVAPRNWFATQALH NPDLIPEQWHRL	267
Eubacterium sp. CAG:581	WP_02250 5071.1	548315094	protein [Eubacterium sp. CAG:581]	22.33		MIEVNIVGQLGNQMFEYACARQLQKKYGGEIVLNTYEMRKETPNFKLSILDYKLSENVKIISDKPLSSANANN YLVKIMRQYFPNWYFNFMAKRGTFVWKSARKYKELPELNEQLSKHIVLNGYWQCDKYFNDVVDTIREDFTP KYPLKAENEQLLEKIKSTESVCVTIRRGDFMNEKNKDTFYKCDDDYFNKALSKIKELCPDCTFFGFSDDVEWIKK NVNFPGEVYFESGNDPVWEKLRLMSACKHFVLSNSSFSWWAQYLSDNNNKIVVAPDIWYKTGDPKKTALY QDGWNLIHIGD	268

Providencia alcalifaciens	AFH02807.1	383289327	glycosyltransferase [Providencia alcalifaciens]	22.26		MKINGKESSMKIKQKKISHIIGGLGNQLFQYATSYALAKENNAKIVIDDRLFKKYKLHGGYRLDKLNIIGEKISS IDKLLFPLILCKLSQKENFIFKSTKKFILEKKTSSFKYLTFSDKEHTKMLIGYWQNAIYFQKYFSELKEMFVPLDIS QEQLDLSIQIHAQQSVALHVRRGDYISNKNALAMIHGICSIDYYKNSIQHINAKLEKPFFYIFSNDKLWCEENLT PLFDGNFHIVENNSQEIDLWLISQCQHIIANSTFSWWGAWLANSDSQIVITPDPWFNKEIDIPSPVLSHWL KLKK	269
Salmonella enterica	AFW04804.1	411146173	glycosyltransferase [Salmonella enterica]	22.26		MFSCLSGGLGNQMFQYSAAYILKKNICHAQLIIDDSYFYCQPQKDTPRNFEINQFNIVFDRVTTDEEKRAISKL RKFKKIPLPLFKSNVITEFLFGKSLLTDEDFYKVLKKNQFTVKMNACLFSLYQDSSLINKYRDLILPLFTINDELLQ VCQQLDSYGFICEHTNTTSLHIRRGDYVTNPHAAKFHGTLSMNYYSQAMNYVDHKLGKQLFIIFSDDVQWA AEKFGGRSDCYIVNNVNCQFSAIDMYLMSLCNNNIIANSTYSWWGAWLNKSEEKLVIAPRKWFAEDKESLL AVNDWISI	270
Sulfurospirillum deleyianum;Sulfurospirillum deleyianum DSM 6946	YP_003304829.1	268680398	protein Sdel_1779 [Sulfurospirillum deleyianum DSM 6946]	22.18		MIIIKIMGGLASQLHKYSVGRALSLKYNTELKLDIFWFDNISGSDTIREYHLDKYNVVAKIATEQEIKQFKPNKY LLKINNLFQKFTNWKINYRNTYCNESFISLENFNLLPDNIYVEGEWSGDRYFSHIKEILQKELTLKSEYMDSTNHF LAKQSSDFAHDDNASKLHCTCSLEYYKKALQYISKNLLKMKLLIFSDDLDWLKPNFNFLDNVEFEFVEGFQDY EEFHLMTLSKHNIIANSGFSLFFAWLNINPHNKIIISLSEWVFEEKLNKYIIDNIKDKNILFLENLE	271
Pseudovibrio sp. FO-BEG1	YP_005080114.1	374329930	alpha-1,2-fucosyltransferase [Pseudovibrio sp. FO-BEG1]	22.15		MSVASQVRISGAARRRKLKPTLIVRIRGGIGNQLFQYALGRKIALETGMKLRFDRSEYDQYFNRSYCLNLFKT QGLSATESEMSAVLWPAQSFGQTVKLCRKFYPFYQRRYIREDELLQDSETPVLKQSAYLDGYWQTWEIPFSI MEQLRDEITLKKPMVLERLKLLQRIKSGPSAALHVRYGDYSQAHNLQNFGLCSAGYYKGAMDFLTERVPGLT FYVFSDSPERAREVVPQQENVYFSDPMQDGKDHEDLMVMSSCDHIVTANSTFSWWAAFLNGNEDKHVIA PLKWFKNPNLDDSLIVPPHWQRL	272
Prevotella sp. oral taxon 472 str. F0295;Prevotella sp. oral taxon 472	WP_009236633.1	496529942	alpha-1,2-fucosyltransferase [Prevotella sp. oral taxon 472]	22.11		MKIVCIKGGLGNQLFEYCRYHGLLRQHNNHGVYLHYDRRRTKQHGGVWLDKAFLITLPTEPWRVKLMVM ALKMLRKLHLFKRLYREDDPRAVLIDDYSQHKQFITNAAEILNFRPFAQLDYVDEITSEPFAVSVHVRRGDYLL PANKANFGVCSVHYYLSAAVAVRERHPDARFFVFSDDIEWAKMNILNLPNCVFVEHAQPQPDHADLYLMSL CKGHIIANSTFSFWGAYLSMGSSAIAIYPKQWFAEPTWNAPDIFLGHWIAL	273
Butyrivibrio fibrisolvens	WP_022752732.1	551008155	glycosyl transferase [Butyrivibrio fibrisolvens]	22.08		MLIIRVAGGLGNQMQQYAMYRKLKSLGKEVKLDLSWFDVENQEGQLAPRKCELKYFDGVDFEECTDAERA YFTKRSILTKALNKVFPATCKIFEETEMFHPEIYSFKDKYLEGYFLCNKYYDDILPFIQNEIVFPKHSDPKMQKRN EELMERMDGWHTASIHLRRGDYITEPQNEALFGNIATDAYYDAAIRYVLDKDYQTHFYIFSNDPEYAREHYS DESRYTIVTGNDGDNSLLDMELMSHCRYNICANSTFSFWGARLNKRSDKEMIRTFKMRNNQEVTAREMTD YWKDWILIDEKGNRIF	274

Lewinella persica	WP_02057 1066.1	522059857	protein [Lewinella persica]	22.04		MVISRLHSGLGNQMFQYAFARRIQLQLNVKLRIDLSILLDSRPPDGYIKREYDLDIFKLSPAYHCNPTSLRILYA PGKYRWSQVVRDLARKGYPVYMEKSFSVDNTLLDSPPDNVIYQGYWQSERYFSEVANTIRKDFAFQHSIQP QSESLAREIRKEDSVCLNIRRKDYLASPTHNVTDETYYENCIQQMRERFSGARFFLFSDDLVWCREFFADFHD VVIVGHDHAGPKFGNYLQLMAQCHHYIIPNSTFAWWAAWLGERTGSVIMAPERWFGTDEFDYRDVVPER WLKVPN	275

OTHER EMBODIMENTS

While the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

The patent and scientific literature referred to herein establishes the knowledge that is available to those with skill in the art. All United States patents and published or unpublished United States patent applications cited herein are incorporated by reference. All published foreign patents and patent applications cited herein are hereby incorporated by reference. Genbank and NCBI submissions indicated by accession number cited herein are hereby incorporated by reference. All other published references, documents, manuscripts and scientific literature cited herein are hereby incorporated by reference.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

What is claimed is:

1. A method for producing a fucosylated oligosaccharide in a bacterium comprising providing bacterium comprising an exogenous lactose-utilizing α(1,2) fucosyltransferase enzyme, wherein the amino acid sequence of said enzyme comprises
    (i) the amino acid sequence of SEQ ID NO: 11;
    (ii) the amino acid sequence of SEQ ID NO: 12;
    (iii) the amino acid sequence of SEQ ID NO: 13;
    (iv) the amino acid sequence of SEQ ID NO: 14;
    (v) the amino acid sequence of SEQ ID NO: 17; or
    (vi) the amino acid sequence of SEQ ID NO: 19.

2. The method of claim 1, further comprising retrieving the fucosylated oligosaccharide from said bacterium or from a culture supernatant of said bacterium.

3. The method of claim 1 or 2, wherein said fucosylated oligosaccharide comprises 2'-fucosyllactose (2'-FL), lactodifucotetraose (LDFT), Lacto-N-fucopentaose I (LNF I), or lacto-N-difucohexaose I (LDFH I).

4. The method of any one of the preceding claims, wherein the bacterium further comprises an exogenous lactose-utilizing α(1,3) fucosyltransferase enzyme and/or an exogenous lactose-utilizing α(1,4) fucosyltransferase enzyme.

5. The method of claim 4, wherein the exogenous lactose-utilizing α(1,3) fucosyltransferase enzyme comprises (i) a *Helicobacter pylori* 26695 *futA* gene; or (ii) a Helicobacter pylori UA948 FucTa gene or a Helicobacter pylori strain DMS6709 FucT III gene.

6. The method of claim 5, wherein
    (i) said method further comprises culturing said bacterium in the presence of tryptophan and in the absence of thymidine;

(ii) said bacterium comprises a deleted or inactivated endogenous lacZ gene and/or a deleted or inactivated endogenous lacI gene of said bacterium;

(iii) said bacterium comprises a deleted or inactivated endogenous lacZ gene and/or a deleted or inactivated endogenous lacI gene of said bacterium, wherein said bacterium further comprises an exogenous lacZ gene or variant thereof, wherein said exogenous lacZ gene or variant thereof comprises an β-galactosidase activity level less than wild-type bacterium;

(iv) said bacterium comprises less than 6,000 units of β-galactosidase activity;

(v) said bacterium comprises less than 1,000 units of β-galactosidase activity;

(vi) said bacterium comprises a lacIq gene promoter immediately upstream of a lacY gene;

(vii) said bacterium comprises an inactivation of the wcaJ gene;

(viii) said bacterium comprises a null mutation, inactivating mutation, or deletion of an endogenous lon gene;

(ix) said bacterium comprises an inactivating mutation of an endogenous lon gene, wherein said inactivating mutation of an endogenous lon gene comprises the insertion of a functional E. coli $lacZ^+$ gene;

(x) said bacterium further comprises a functional lactose permease gene;

(xi) said bacterium further comprises a functional lactose permease gene, wherein said bacterium comprises E. coli lacY;

(xii) said bacterium further comprises an exogenous E. coli rcsA or E. coli rcsB gene;

(xiii) said bacterium further comprises a mutation in the thyA gene;

(xiv) said bacterium accumulates intracellular lactose in the presence of exogenous lactose;

(xv) said bacterium accumulates intracellular GDP-fucose; or

(xvi) said bacterium comprises the genotype ΔampC::$P_{trp}{}^{B}$cI, Δ(lacI-lacZ)::FRT,P_{lacIq}lacY$^+$, ΔwcaJ::FRT, thyA::Tn10, Δlon:(npt3, lacZ$^{+)}$.

7. The method of any one of the preceding claims, wherein said bacterium is a member of the *Bacillus, Pantoea, Lactobacillus, Lactococcus, Streptococcus, Proprionibacterium, Enterococcus, Bifidobacterium, Sporolactobacillus, Micromomospora, Micrococcus, Rhodococcus,* or *Pseudomonas* genus, preferably wherein said production strain is selected from the group consisting of *Bacillus*

licheniformis, Bacillus subtilis, Bacillus coagulans, Bacillus thermophilus, Bacillus laterosporus, Bacillus megaterium, Bacillus mycoides, Bacillus pumilus, Bacillus lentus, Bacillus cereus, and Bacillus circulans, Erwinia herbicola (*Pantoea agglomerans*), *Citrobacter freundii, Pantoea citrea, Pectobacterium carotovorum, Xanthomonas campestris Lactobacillus acidophilus, Lactobacillus salivarius, Lactobacillus plantarum, Lactobacillus helveticus, Lactobacillus delbrueckii, Lactobacillus rhamnosus, Lactobacillus bulgaricus, Lactobacillus crispatus, Lactobacillus gasseri, Lactobacillus casei, Lactobacillus reuteri, Lactobacillus jensenii, Lactococcus lactis, Streptococcus thermophiles, Proprionibacterium freudenreichii, Enterococcus faecium, Enterococcus thermophiles*), *Bifidobacterium longum, Bifidobacterium infantis, Bifidobacterium bifidum, Pseudomonas fluorescens* and *Pseudomonas aeruginosa*.

8. The method of any of the preceding claims, wherein said bacterium comprises a nucleic acid construct comprising an isolated nucleic acid encoding said α(1,2) fucosyltransferase enzyme.

9. An isolated bacterium comprising an isolated nucleic acid encoding a lactose-accepting α (1,2) fucosyltransferase enzyme, wherein the amino acid sequence of said enzyme encoded by said nucleic acid comprises
   (i) the amino acid sequence of SEQ ID NO: 11;
   (ii) the amino acid sequence of SEQ ID NO: 12;
   (iii) the amino acid sequence of SEQ ID NO: 13;
   (iv) the amino acid sequence of SEQ ID NO: 14;
   (v) the amino acid sequence of SEQ ID NO: 17; or
   (vi) the amino acid sequence of SEQ ID NO: 19; and
   wherein the isolated bacterium further comprising a α(1,3) fucosyltransferase enzyme and/or a α(1,4) fucosyltransferase enzyme.

10. The method of any one of claims 1 to 8, or the isolated bacterium of claim 9, wherein said bacterium is *Escherichia coli*.

11. The method of any one of claims 1 to 8 or 10, or the isolated bacterium of any one of claims 9 to 10, wherein said bacterium further comprises reduced level of

β-galactosidase activity, a defective colanic acid synthesis pathway, an inactivated adenosine-5'-triphosphate (ATP)-dependent intracellular protease, an inactivated endogenous *lacA* gene, or any combination thereof.

12. The method of any one of claims 1 to 8 or 10 to 11, or the isolated bacterium of any one of claims 9 to 11, wherein the enzyme comprises the amino acid sequence of SEQ ID NO: 13.