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.

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 reduced level of β-galactosidase activity comprises a deleted or inactivated endogenous lacZ gene and/or a deleted or inactivated endogenous lacl gene of said bacterium;

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

(v) said reduced level of β-galactosidase activity comprises an activity level less than wild-type bacterium, wherein said reduced level of β-galactosidase activity comprises less than 6,000 units of β-galactosidase activity;

(vi) said reduced level of β-galactosidase activity comprises an activity level less than wild-type bacterium, wherein said reduced level of β-galactosidase activity comprises less than 1,000 units of β-galactosidase activity;

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

(viii) said defective colonic acid synthesis pathway comprises the inactivation of the wcaJ gene of said bacterium;

(ix) said inactivated ATP-dependent intracellular protease is a null mutation, inactivating mutation, or deletion of an endogenous Ion gene;

(x) said inactivated ATP-dependent intracellular protease is 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;

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

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

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

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

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

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

(xvii) 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 α(1,2) fucosyltransferase enzyme comprises any one selected from *Tannerella sp.* FutS, *Clostridium bolteae* +13 FutP, *Methanosphaerula palustries* FutR, *Akkermansia muciniphilia* FutY, *Prevotellaa sp.* FutW, or *Clostridium bolteae* FutP,.

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

12. The method of any one of claims 1 to 8, 10 to 11, or the isolated bacterium of any one of claims 9 to 11, 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.

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