



US009695223B2

(12) **United States Patent**
Alper et al.

(10) **Patent No.:** **US 9,695,223 B2**

(45) **Date of Patent:** **Jul. 4, 2017**

(54) **ENGINEERED XYLOSE TRANSPORTERS WITH REDUCED GLUCOSE INHIBITION**

(71) Applicant: **Board of Regents, The University of Texas System**, Austin, TX (US)

(72) Inventors: **Hal Alper**, Austin, TX (US); **Eric Young**, Arlington, MA (US); **Sunmi Lee**, Austin, TX (US)

(73) Assignee: **BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM**, Austin, TX (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **14/719,206**

(22) Filed: **May 21, 2015**

(65) **Prior Publication Data**

US 2015/0344532 A1 Dec. 3, 2015

Related U.S. Application Data

(60) Provisional application No. 62/001,495, filed on May 21, 2014.

(51) **Int. Cl.**

C12N 9/00 (2006.01)
C07K 14/40 (2006.01)
C07K 14/33 (2006.01)
A61K 38/00 (2006.01)

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

CPC **C07K 14/40** (2013.01); **C07K 14/33** (2013.01); **A61K 38/00** (2013.01)

(58) **Field of Classification Search**

CPC C07K 14/40; C07K 14/39
USPC 435/183, 252.3; 536/23.2
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

7,910,718 B2 3/2011 Simkin et al.
2010/0017904 A1 1/2010 Abad et al.
2012/0329109 A1 12/2012 Chua et al.
2016/0280745 A1* 9/2016 Alper C07K 14/40

FOREIGN PATENT DOCUMENTS

WO WO-2014/018552 A1 1/2014
WO WO-2015/179701 A1 11/2015

OTHER PUBLICATIONS

Written Opinion published in a related application: PCT/US15/32058, filed on May 21, 2015.*

Bengtsson, O. et al. (Nov. 2008). "Identification of common traits in improved xylose-growing *Saccharomyces cerevisiae* for inverse metabolic engineering," 25(11):835-847.

Curran, K.A. et al. (Jul. 2012). "Expanding the chemical palate of cells by combining systems biology and metabolic engineering," *Metab Eng* 14(4):289-297.

International Search Report mailed on Aug. 26, 2015, for PCT Application No. PCT/US2015/032058, filed May 21, 2015, 4 pages.

Wahlbom, C.F. et al. (Feb. 2003). "Molecular analysis of a *Saccharomyces cerevisiae* mutant with improved ability to utilize xylose shows enhanced expression of proteins involved in transport, initial xylose metabolism, and the pentose phosphate pathway," *Appl Environ Microbiol* 69(2):740-746.

Written Opinion mailed on Aug. 26, 2015, for PCT Application No. PCT/US2015/032058, filed May 21, 2015, 5 pages.

Young, E.M. et al. (Jan. 7, 2004, e-published Dec. 16, 2013). "Rewiring yeast sugar transporter preference through modifying a conserved protein motif," *Proc Natl Acad Sci USA* 111(1):131-136.

* cited by examiner

Primary Examiner — Tekchand Saidha

(74) *Attorney, Agent, or Firm* — Meunier Carlin & Curfman LLC

(57) **ABSTRACT**

Provided herein are compositions and methods useful for reducing glucose inhibition in transporting xylose, arabinose and other monosaccharides, into a yeast cell.

9 Claims, 12 Drawing Sheets

FIG. 1

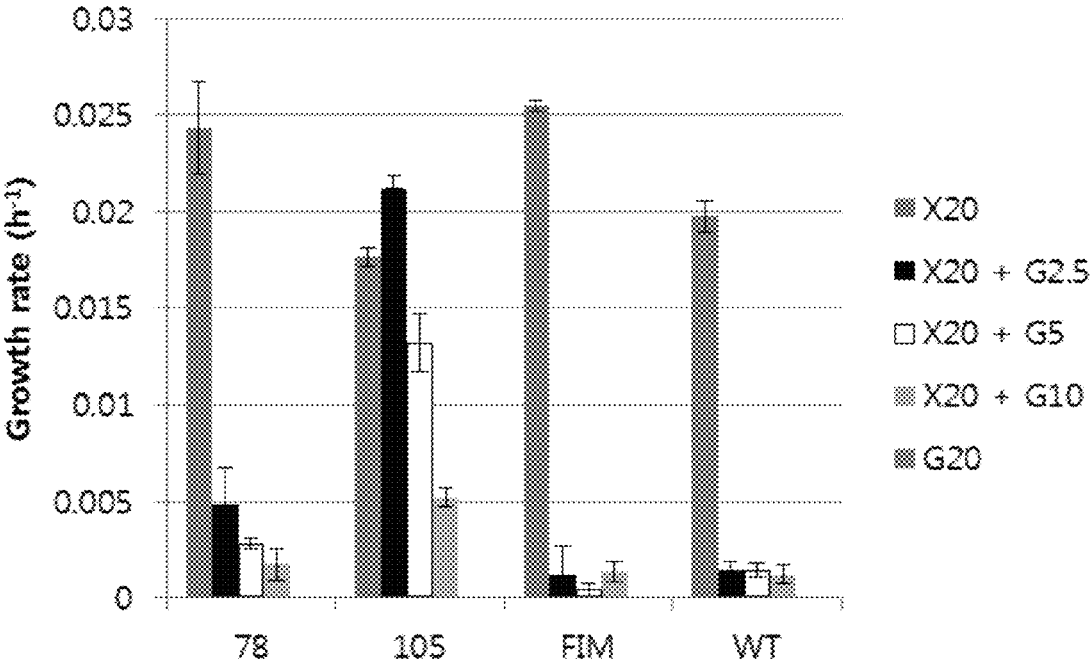


FIG. 2

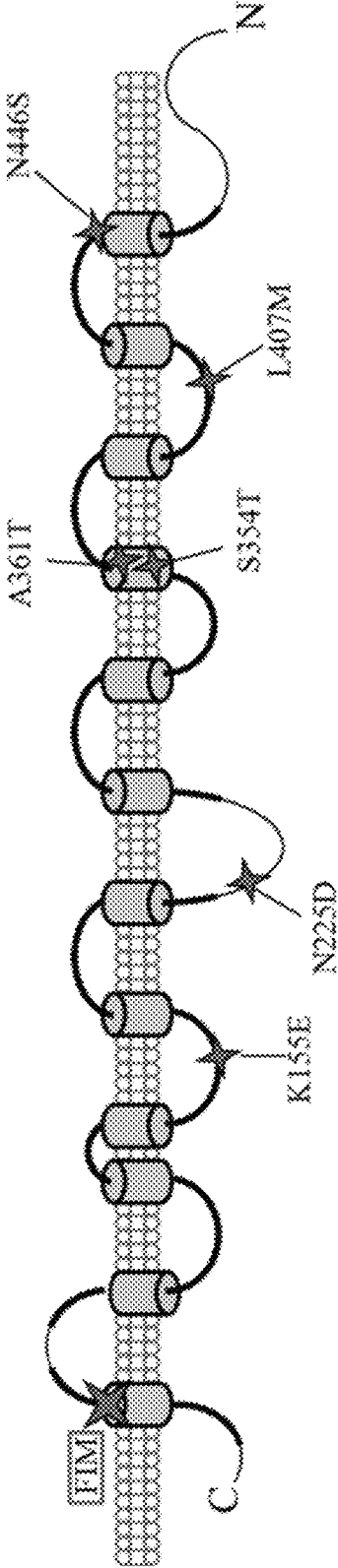


FIG. 3

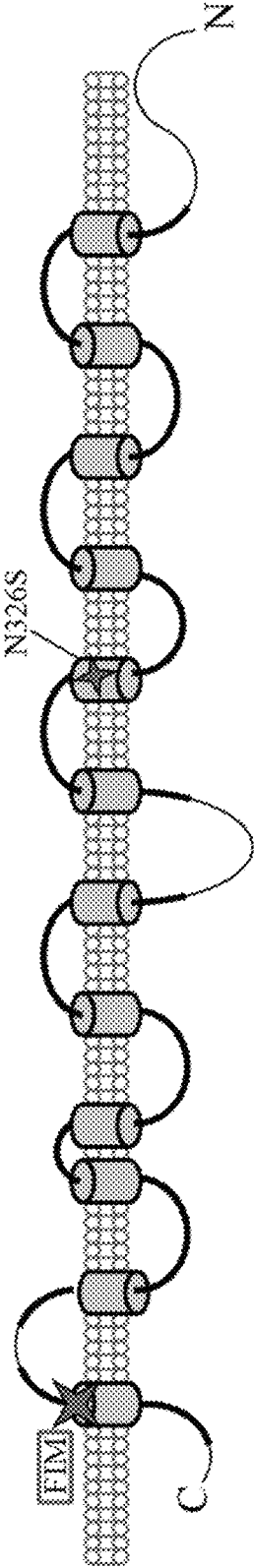


FIG. 4

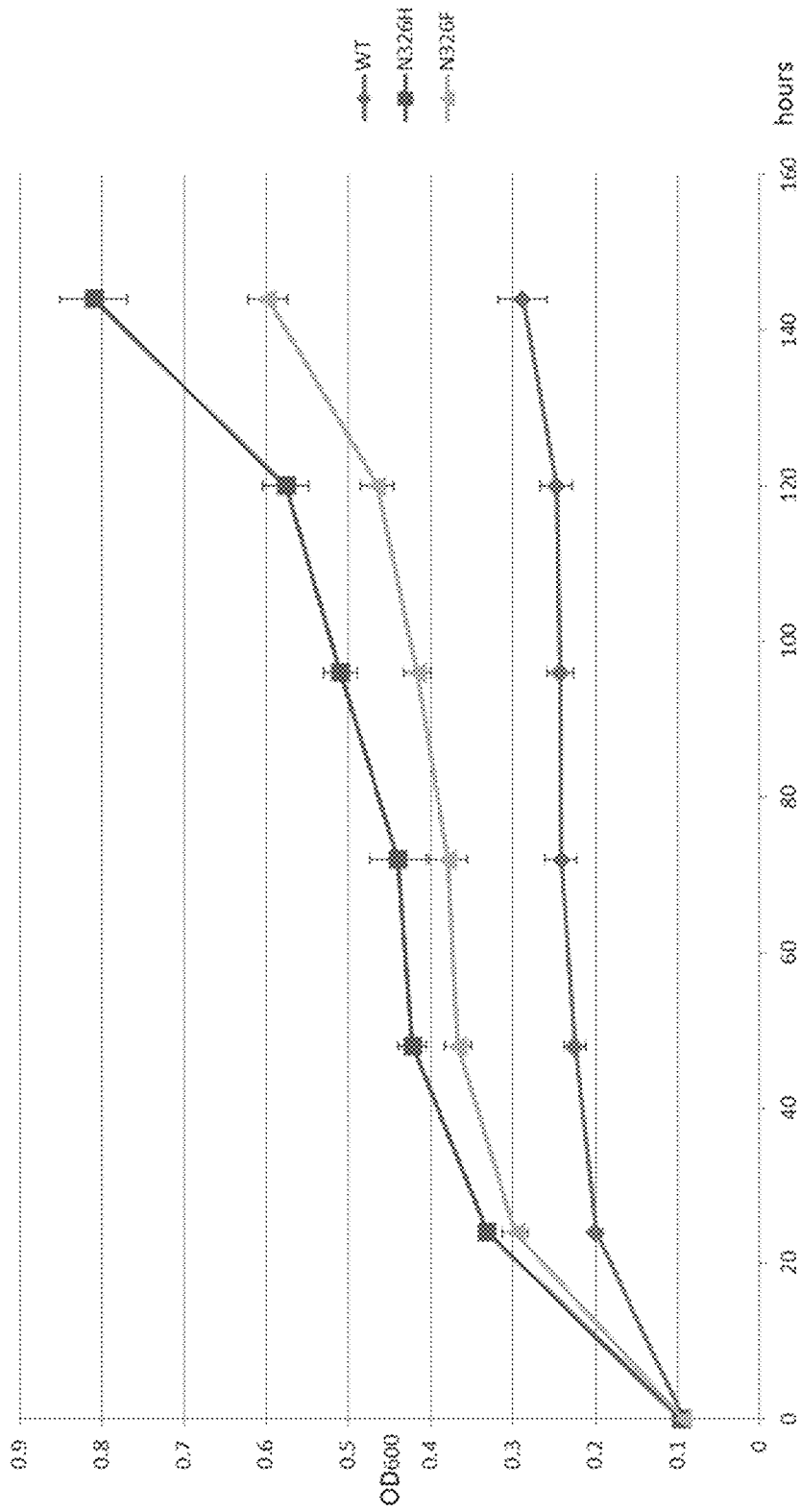


FIG. 5

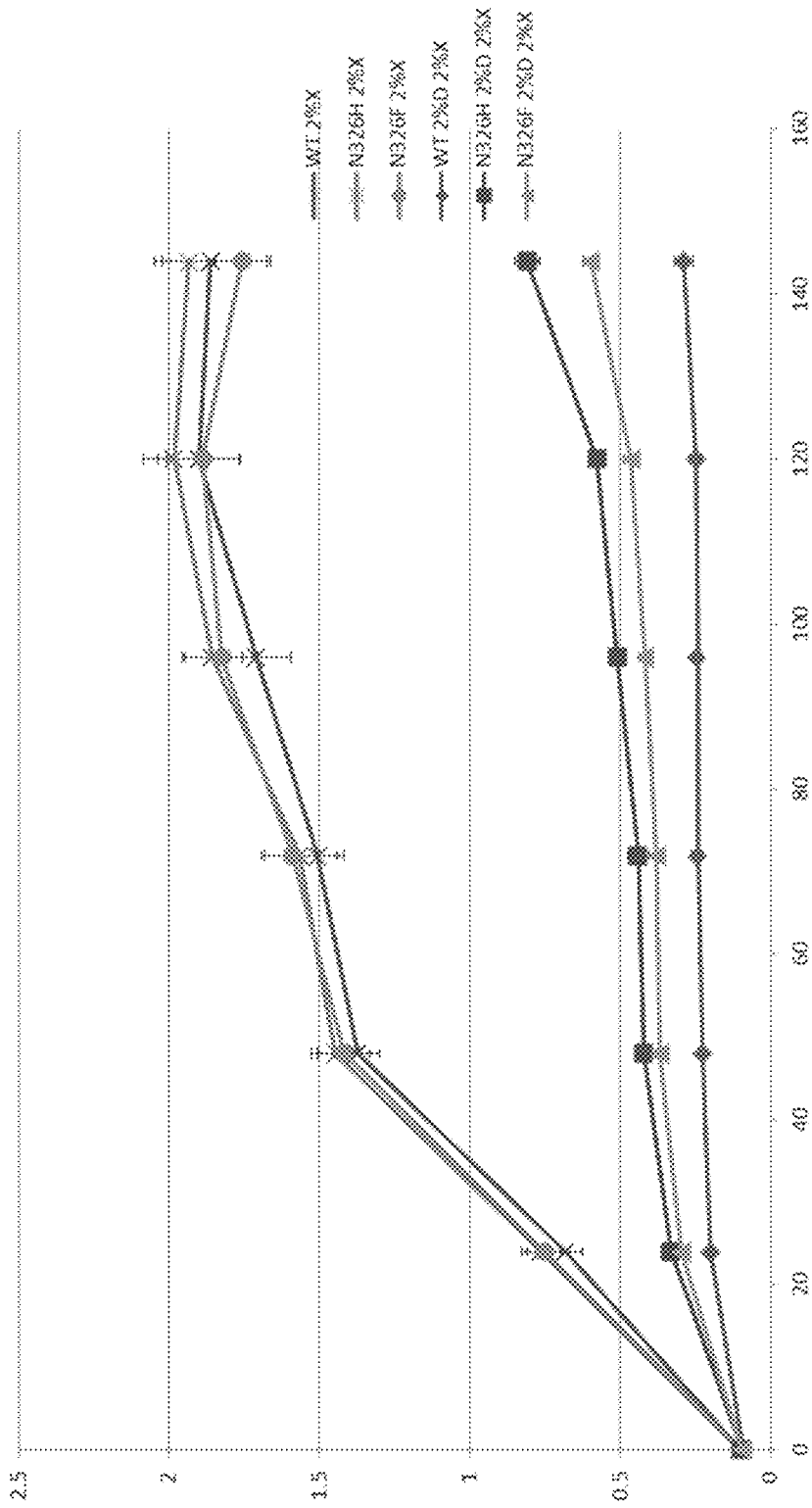


FIG. 6

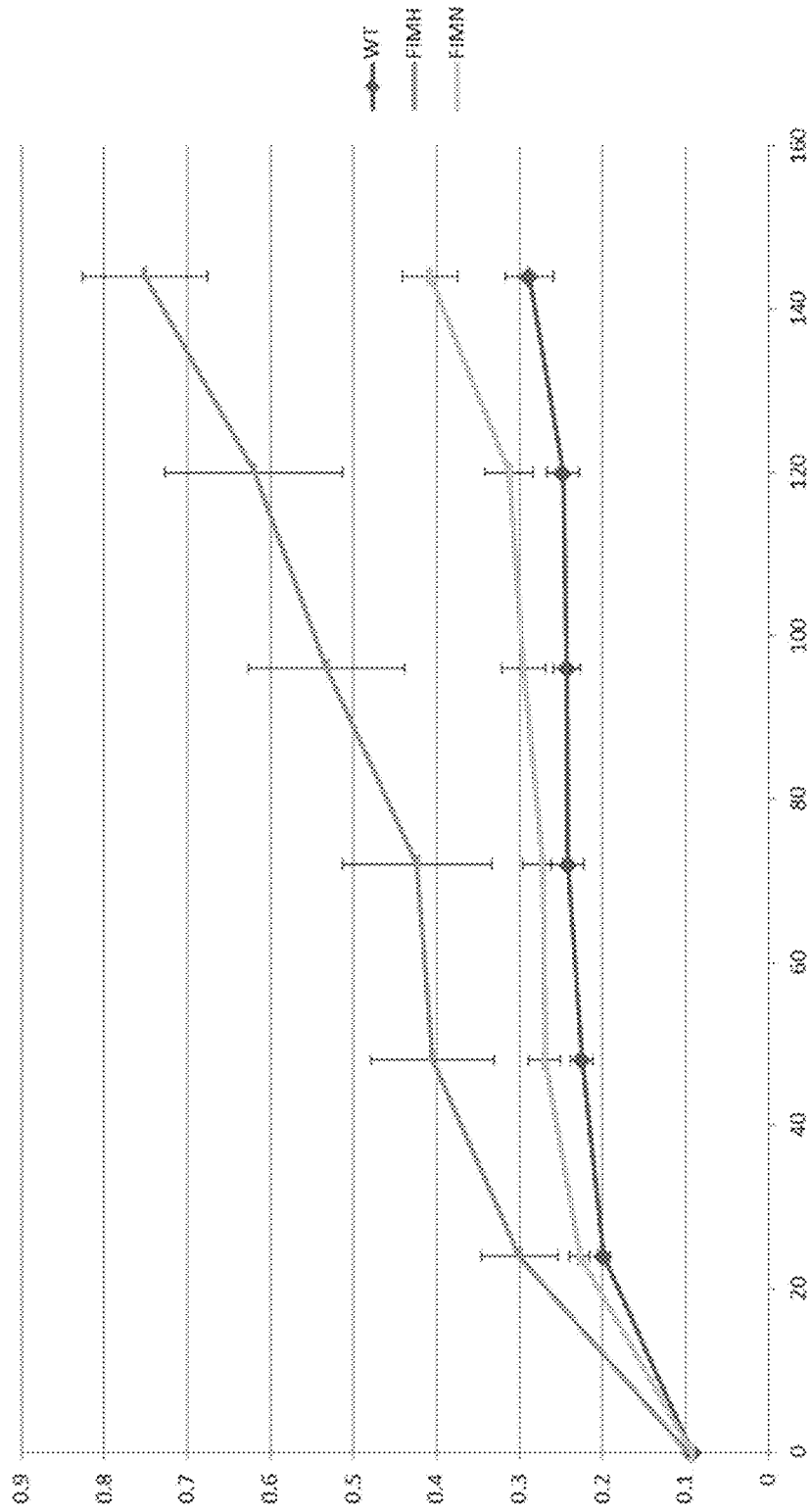


FIG. 7

Extracellular

Cytoplasm

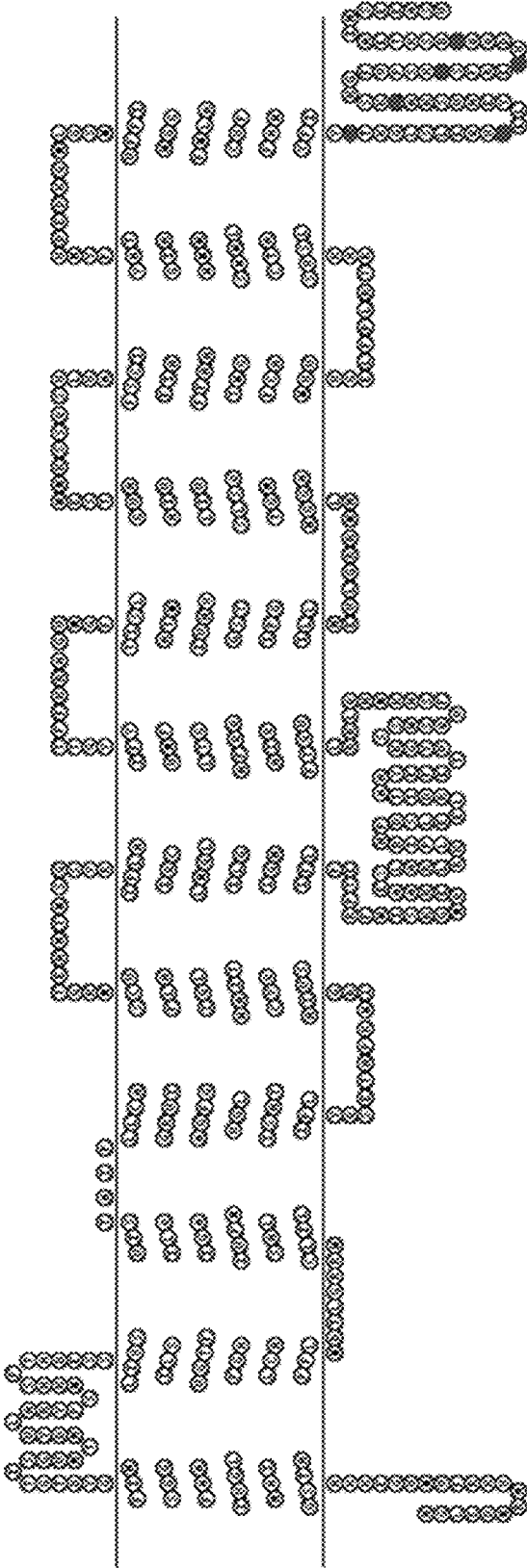


FIG. 8

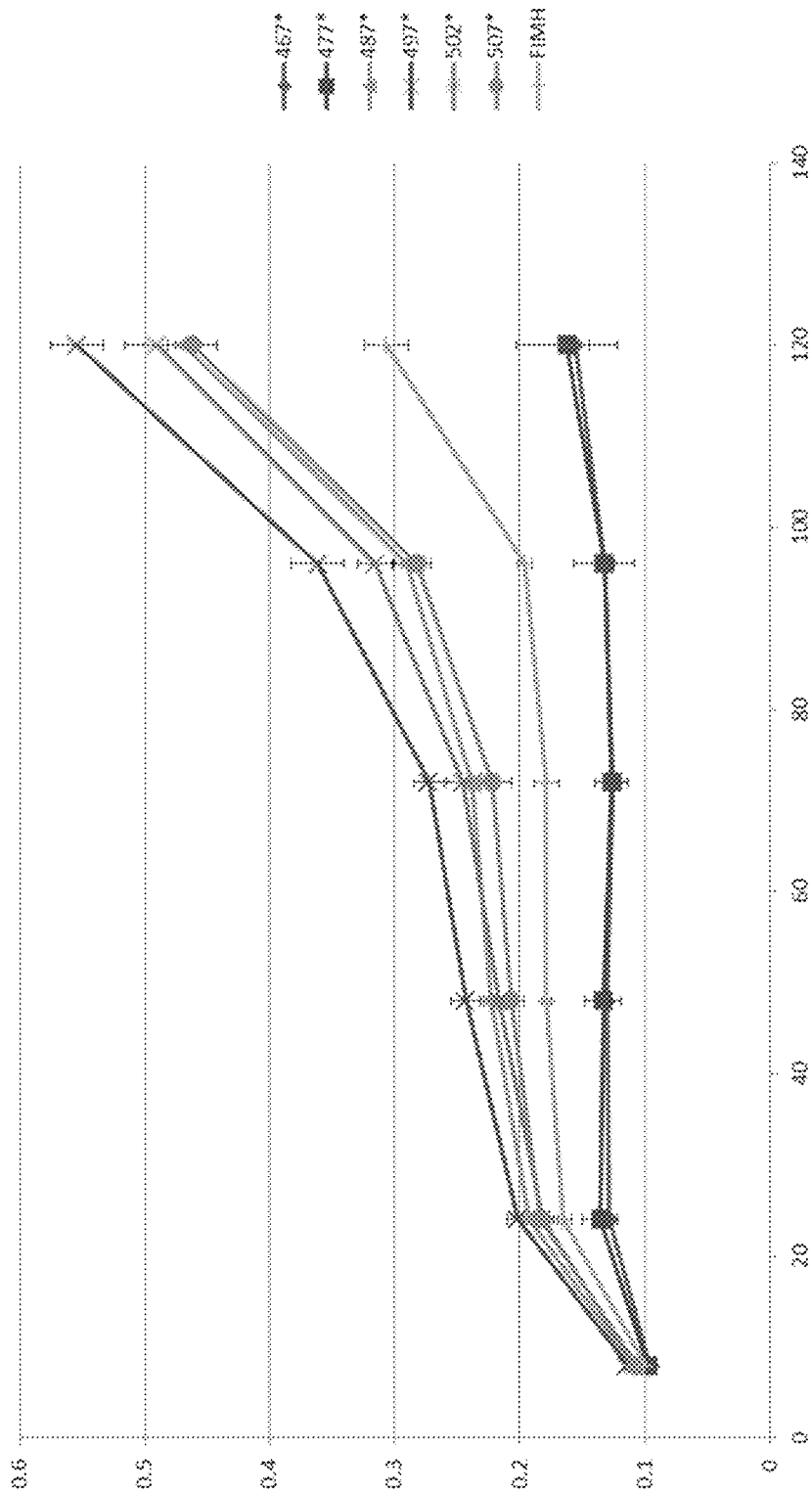


FIG. 9A

WT CIGXS1

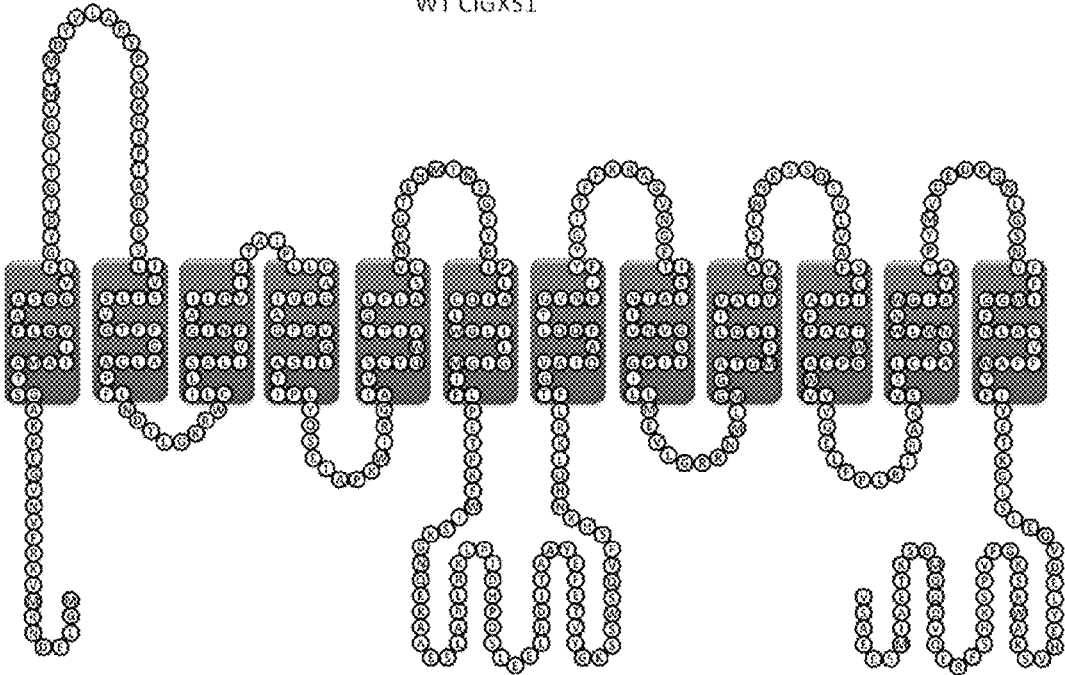


FIG. 9B

CIGXS1 FIM

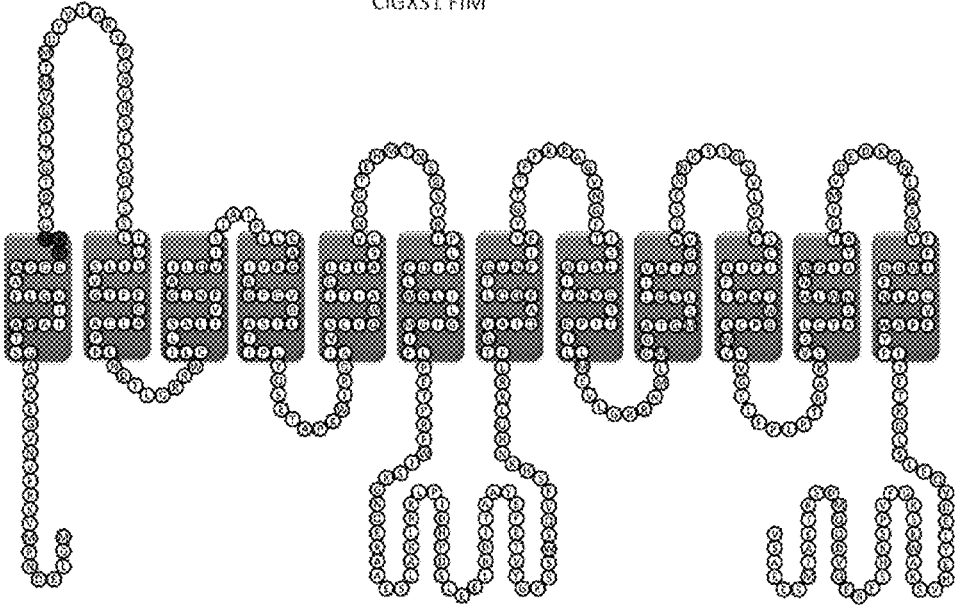


FIG. 9C

CIGXS1 FIM N326H

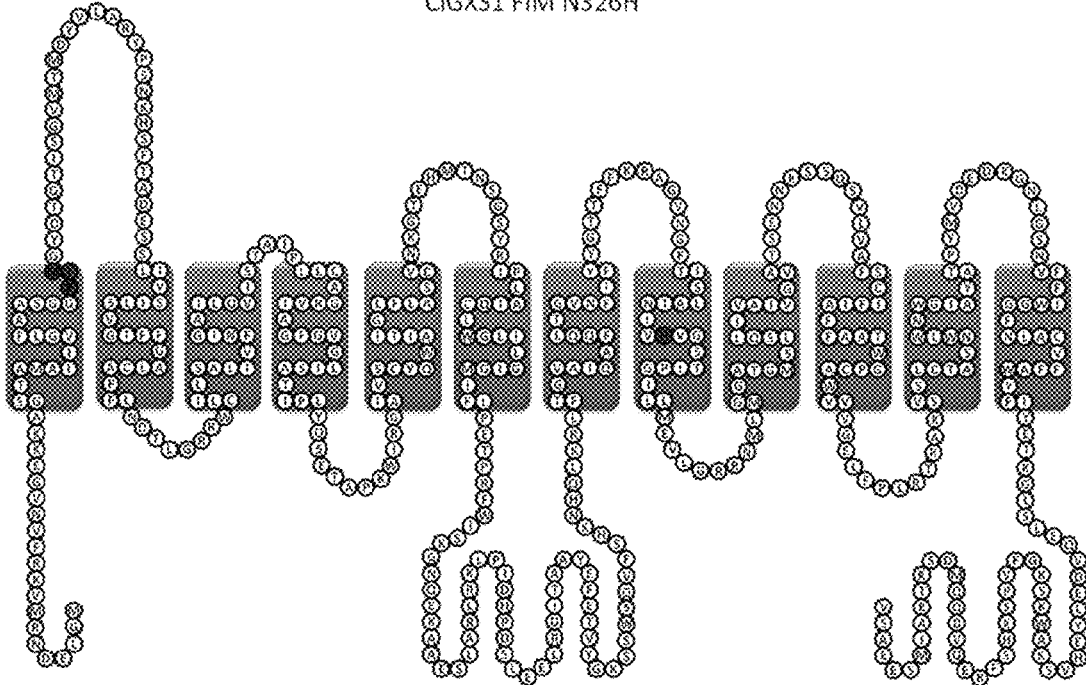


FIG. 9D

CIGXS1 FIM T170N

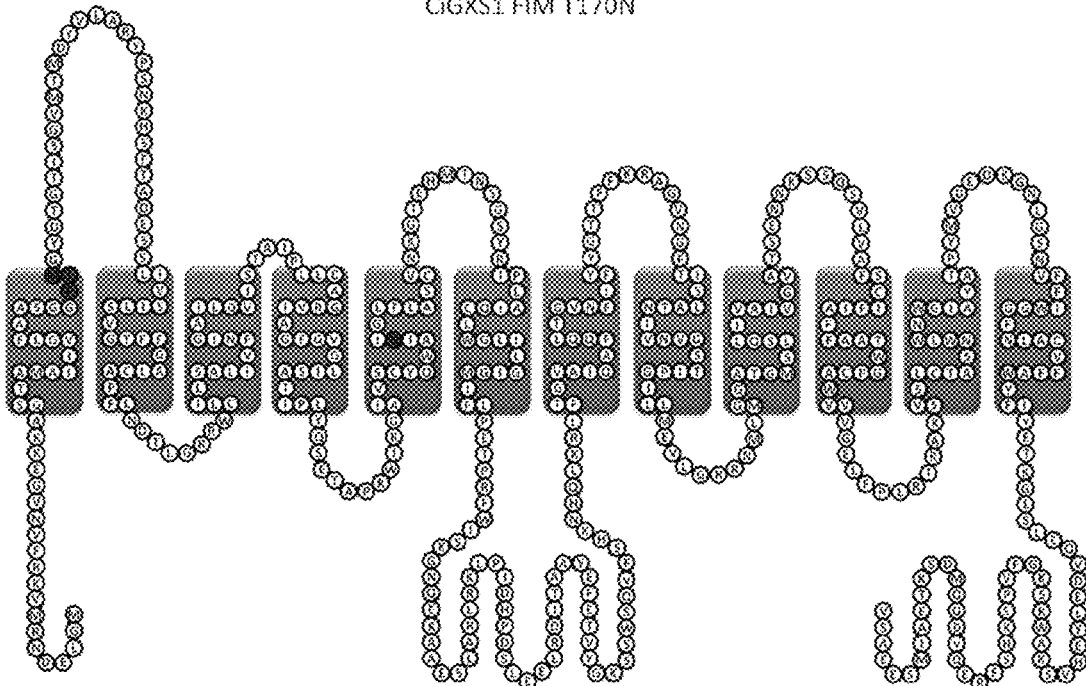


FIG. 9E

CIGXS1 (FIMH) Δ497

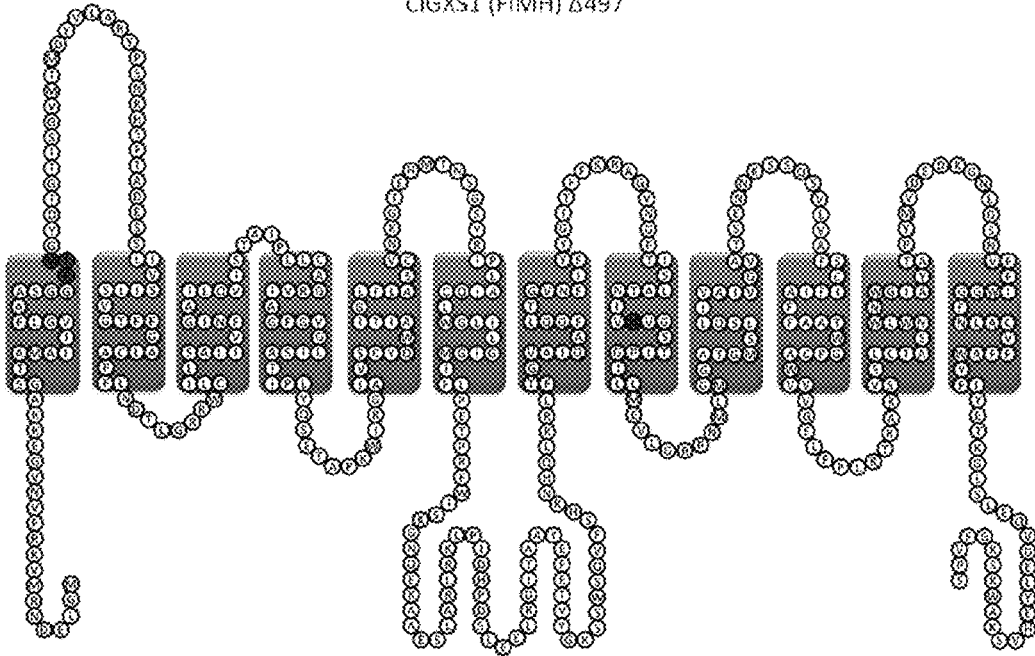


FIG. 9F

GXS1 (FIMH Δ497) with I171F

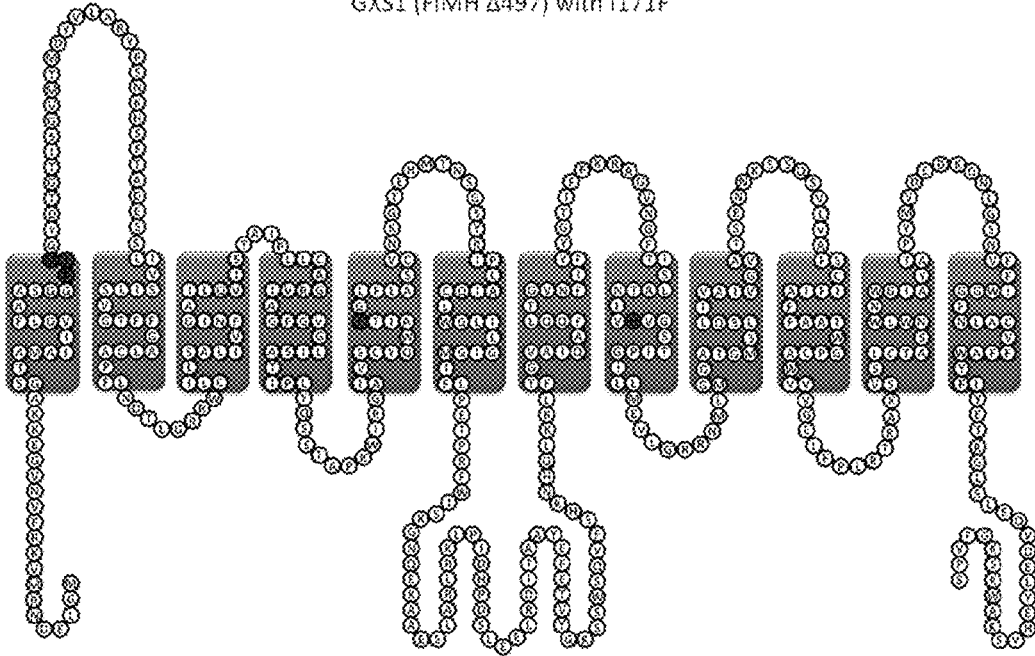
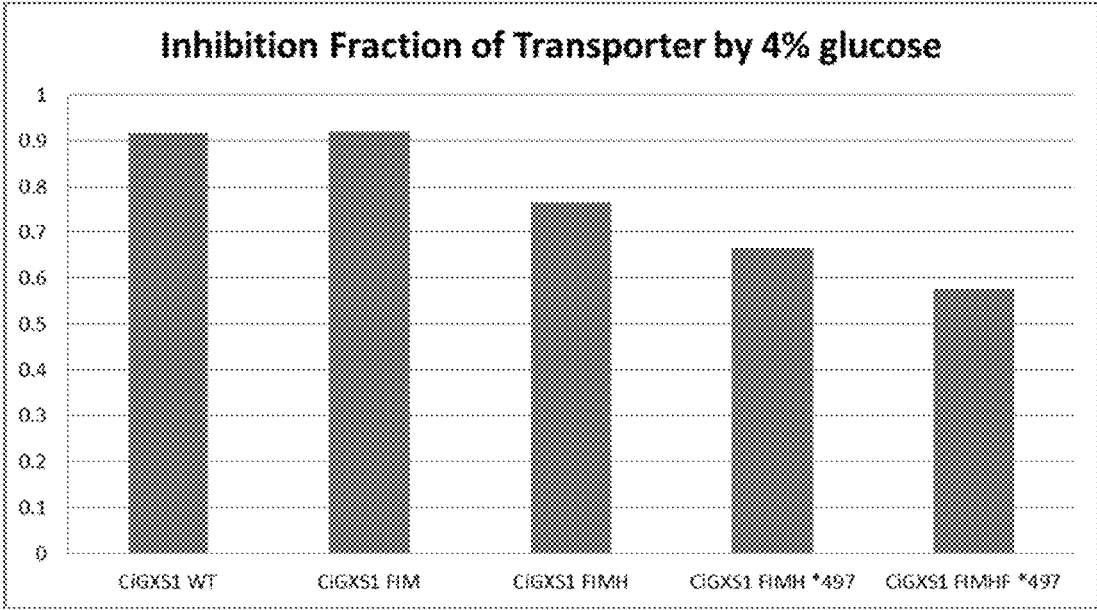


FIG. 10



1

**ENGINEERED XYLOSE TRANSPORTERS
WITH REDUCED GLUCOSE INHIBITION****CROSS-REFERENCES TO RELATED
APPLICATIONS**

This application claims priority to U.S. Provisional Application No. 62/001,495, filed May 21, 2014, the disclosure of which is incorporated herein in its entirety and for all purposes.

**STATEMENT AS TO RIGHTS TO INVENTIONS
MADE UNDER FEDERALLY SPONSORED
RESEARCH AND DEVELOPMENT**

This invention was made with government support under grant no. CBET1067506, awarded by the National Science Foundation. The government has certain rights in the invention.

**REFERENCE TO A "SEQUENCE LISTING," A
TABLE, OR A COMPUTER PROGRAM LISTING
APPENDIX SUBMITTED ON A COMPACT
DISK**

The Sequence Listing written in file 48932-525001US_ST25.TXT, created on May 21, 2015, 171,093 bytes, machine format IBM-PC, MS Windows operating system, is hereby incorporated by reference.

BACKGROUND OF THE INVENTION

The quest for an optimal xylose pathway in yeast is of utmost importance along the way to realizing the potential of lignocellulosic biomass conversion into fuels and chemicals. An often overlooked aspect of this catabolic pathway is the molecular transport of this sugar. Molecular transporter proteins facilitate monosaccharide uptake and serve as the first step in catabolic metabolism. In this capacity, the preferences, regulation, and kinetics of these transporters ultimately dictate total carbon flux (1-3); and optimization of intracellular catabolic pathways only increases the degree to which transport exerts control over metabolic flux (4, 5). Thus, monosaccharide transport profiles and rates are important design criteria and a driving force to enable metabolic engineering advances (6-10). Furthermore, the presence of other hexose sugars, such as glucose, can hamper efficient transport of xylose by inhibiting sugar transporters. There is a need in the art for efficient transport systems for xylose in yeast when glucose is present. Provided herein are solutions to these and other problems in the art.

BRIEF SUMMARY OF THE INVENTION

Provided herein are compositions and methods useful for transporting xylose, arabinose, and galactose into a yeast cell in the presence of glucose.

Recombinant transporter proteins are provided herein that transport hexoses or pentoses. In one aspect, the recombinant transporter is a recombinant xylose transporter protein that includes a xylose transporter motif sequence and a glucose mitigation mutation. In another aspect, the recombinant transporter is a recombinant arabinose transporter protein that includes an arabinose transporter motif sequence and a glucose mitigation mutation. In yet another aspect, the recombinant transporter protein is a recombinant galactose

2

transporter protein that includes a galactose transporter motif sequence and a glucose mitigation mutation.

Also provided herein are nucleic acids that encode recombinant transporter proteins described herein. Thus, in one aspect is a nucleic acid encoding a recombinant xylose transporter as described herein, including embodiments thereof. In another aspect is a nucleic acid encoding a recombinant arabinose transporter as described herein, including embodiments thereof. In yet another aspect is a nucleic acid encoding a recombinant galactose transporter as described herein, including embodiments thereof.

Recombinant yeast cells are described herein which include a recombinant transporter protein as described herein. In one aspect is a recombinant yeast cell that includes a recombinant xylose transporter as described herein, including embodiments thereof. In another aspect is a recombinant yeast cell that includes a recombinant arabinose transporter as described herein, including embodiments thereof. In another aspect is a recombinant yeast cell that includes a recombinant galactose transporter as described herein, including embodiments thereof.

Methods of transporting xylose into a recombinant yeast cell are also described herein. In one aspect, the method includes contacting a recombinant yeast cell with a xylose compound, where the recombinant yeast cell includes a recombinant xylose transporter protein as described herein, including embodiments thereof. The recombinant xylose transporter protein transports the xylose compound into the recombinant yeast cell. In another aspect, the method includes contacting a recombinant yeast cell with a xylose compound, where the xylose compound is the only sugar (i.e. carbon source) in the media, and where the recombinant yeast cell includes a recombinant xylose transporter protein as described herein, including embodiments thereof.

In another aspect is a method of transporting arabinose into a recombinant yeast cell. The method includes contacting a recombinant yeast cell with an arabinose compound, where the recombinant yeast cell includes a recombinant arabinose transporter protein as described herein, including embodiments thereof. The recombinant arabinose transporter protein transports the arabinose compound into the recombinant yeast cell. In another aspect, the method includes contacting a recombinant yeast cell with an arabinose compound, where the arabinose compound is the only sugar (i.e. carbon source) in the media, and where the recombinant yeast cell includes a recombinant arabinose transporter protein as described herein, including embodiments thereof.

In another aspect is a method of transporting galactose into a recombinant yeast cell. The method includes contacting a recombinant yeast cell with a galactose compound, where the recombinant yeast cell includes a recombinant galactose transporter protein as described herein, including embodiments thereof. The recombinant galactose transporter protein transports the galactose compound into the recombinant yeast cell. In another aspect, the method includes contacting a recombinant yeast cell with a galactose compound, where the galactose compound is the only sugar (i.e. carbon source) in the media, and where the recombinant yeast cell includes a recombinant galactose transporter protein as described herein, including embodiments thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1: The growth of *S. cerevisiae* expressing mutants (78 and 105), FIM, wild-type CiGXS1 transporters on xylose medium supplemented with different concentration

of glucose: X20: xylose 20 g/L, X20+G2.5: xylose 20 g/L+glucose 2.5 g/L, X20+G5: xylose 20 g/L+glucose 5 g/L, X20+G10: xylose 20 g/L+glucose 10 g/L, G20: glucose 20 g/L.

FIG. 2: SACS HMMTOP Prediction of the mutant 105 xylose transporter and mutant positions: mutations are indicated with arrows.

FIG. 3: SACS HMMTOP Prediction of the mutant 78 xylose transporter and mutant positions: mutations are indicated with arrows.

FIG. 4: Mutations to residue N326 to convert to a histidine or phenylalanine were explored in the background of the *gxs1* (FIM) mutant containing the xylose transporter motif sequence G-G-F-I-M-G (SEQ ID NO:107). Mutations in residue N326 improve assimilation rates and growth rates in pure xylose over the wild-type GXS1 and N326H is a better mutation than N326F.

FIG. 5: Mutations to residue N326 to convert to a histidine or phenylalanine were explored in the background of the *gxs1* (FIM) mutant containing the xylose transporter motif sequence G-G-F-I-M-G (SEQ ID NO:107). Mutations in residue N326 improve assimilation rates and growth rates in a mixture of 2% glucose and 2% xylose over the wild-type GXS1 and N326H is a better mutation than N326F.

FIG. 6: Mutations to residue N326 to convert to a histidine (N326H) are compared with an additionally discovered mutation T170N in the background of the *gxs1* (FIM) mutant containing the xylose transporter motif sequence G-G-F-I-M-G (SEQ ID NO:107). Both mutants improved xylose growth rates over the wild-type transporter with the mutation of N326H stronger than T170N.

FIG. 7: An additional round of mutagenesis and selection was completed in the background of *gxs1* (FIM) containing N326H (SEQ ID NO:3). These experiments were conducted in the presence of 4% glucose and 2% xylose. Several mutations in the tail region of the transporter were identified including a stop codon (dark dots on schematic of tail with circle indicating location of stop codon). To assess the potential inhibition of this tail region, several truncations of this transporter protein were created (see vertical lines in the sequence on the bottom).

FIG. 8: The *gxs1* (FIM) strain with the N326H mutation was truncated at various amino acid residues (indicated by number) and assessed for growth in the presence of 4% glucose, 2% xylose. Several truncations were superior with respect to growth over the starting transporter with a truncation after amino acid residue 497 giving the highest performance. Truncations before residue 487 were detrimental to performance.

FIG. 9A-9F: Schematics of the transporter structures for the wild-type GXS1 and *gxs1* mutant showing the location of the F-I-M xylose transporter sequence motif. FIG. 9A: WT CiGXS1 (SEQ ID NO:1). FIG. 9B: CiGXS1FIM (SEQ ID NO:2). FIG. 9C: CiGXS1FIM N326H (SEQ ID NO:3). FIG. 9D: CiGXS1FIM T170N (SEQ ID NO:4). FIG. 9E: CiGXS1(FIMH)Δ497 (SEQ ID NO:5). FIG. 9F: CiGXS1 (FIMH-Δ497) with I171F (SEQ ID NO:6).

FIG. 10: The inhibition fraction (the ratio of the transporter capacity in 4% glucose compared to the transporter capacity in pure xylose) for various mutants. The *gxs1* mutant with the F-I-M xylose transporter sequence motif and the N326H mutation along with the tail region after truncating past residue 497 and the T171F mutation performed best and was the least inhibited by high glucose levels.

DETAILED DESCRIPTION OF THE INVENTION

Unless defined otherwise, all technical and scientific terms used herein generally have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Generally, the nomenclature used herein and the laboratory procedures in cell culture, molecular genetics, organic chemistry, and nucleic acid chemistry and hybridization described below are those well-known and commonly employed in the art. Standard techniques are used for nucleic acid and peptide synthesis. The techniques and procedures are generally performed according to conventional methods in the art and various general references (see generally, Sambrook et al. MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed. (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., which is incorporated herein by reference), which are provided throughout this document.

“Nucleic acid” refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form, and complements thereof. The term “polynucleotide” refers to a linear sequence of nucleotides. The term “nucleotide” typically refers to a single unit of a polynucleotide, i.e., a monomer. Nucleotides can be ribonucleotides, deoxyribonucleotides, or modified versions thereof. Examples of polynucleotides contemplated herein include single and double stranded DNA, single and double stranded RNA (including siRNA), and hybrid molecules having mixtures of single and double stranded DNA and RNA. Nucleic acid as used herein also refers nucleic acids that have the same basic chemical structure as a naturally occurring nucleic acids. Such analogues have modified sugars and/or modified ring substituents, but retain the same basic chemical structure as the naturally occurring nucleic acid. A nucleic acid mimetic refers to chemical compounds that have a structure that is different the general chemical structure of a nucleic acid, but that functions in a manner similar to a naturally occurring nucleic acid. Examples of such analogues include, without limitation, phosphorothiolates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, and peptide-nucleic acids (PNAs).

“Synthetic mRNA” as used herein refers to any mRNA derived through non-natural means such as standard oligonucleotide synthesis techniques or cloning techniques. Such mRNA may also include non-proteinogenic derivatives of naturally occurring nucleotides. Additionally, “synthetic mRNA” herein also includes mRNA that has been expressed through recombinant techniques or exogenously, using any expression vehicle, including but not limited to prokaryotic cells, eukaryotic cell lines, and viral methods. “Synthetic mRNA” includes such mRNA that has been purified or otherwise obtained from an expression vehicle or system.

The words “complementary” or “complementarity” refer to the ability of a nucleic acid in a polynucleotide to form a base pair with another nucleic acid in a second polynucleotide. For example, the sequence A-G-T is complementary to the sequence T-C-A. Complementarity may be partial, in which only some of the nucleic acids match according to base pairing, or complete, where all the nucleic acids match according to base pairing.

Nucleic acid is “operably linked” when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the

5

polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding 15 sequence if it is positioned so as to facilitate translation. Generally, “operably linked” means that the DNA sequences being linked are near each other, and, in the case of a secretory leader, contiguous and in reading phase.

The terms “polypeptide,” “peptide” and “protein” are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

The term “amino acid” refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

A “conservative substitution” as used with respect to amino acids, refers to the substitution of an amino acid with a chemically similar amino acid. Amino acid substitutions which often preserve the structural and/or functional properties of the polypeptide in which the substitution is made are known in the art and are described, for example, by H. Neurath and R. L. Hill, 1979, in “The Proteins,” Academic Press, New York. The most commonly occurring exchanges are isoleucine/valine, tyrosine/phenylalanine, aspartic acid/glutamic acid, lysine/arginine, methionine/leucine, aspartic acid/asparagine, glutamic acid/glutamine, leucine/isoleucine, methionine/isoleucine, threonine/serine, tryptophan/phenylalanine, tyrosine/histidine, tyrosine/tryptophan, glutamine/arginine, histidine/asparagine, histidine/glutamine, lysine/asparagine, lysine/glutamine, lysine/glutamic acid, phenylalanine/leucine, phenylalanine/methionine, serine/alanine, serine/asparagine, valine/leucine, and valine/methionine. The following eight groups each contain amino acids that are conservative substitutions for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (see, e.g., Creighton, *Proteins* (1984)). In some embodiments, there may be at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, or at least 40 conservative substitutions. In some

6

embodiments, there may be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, or 40 conservative substitutions.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a “conservatively modified variant” where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

An amino acid or nucleotide base “position” is denoted by a number that sequentially identifies each amino acid (or nucleotide base) in the reference sequence based on its position relative to an N-terminus (or 5'-end). Due to deletions, insertions, truncations, fusions, and the like that must be taken into account when determining an optimal alignment, in general the amino acid residue number in a test sequence determined by simply counting from the N-terminus will not necessarily be the same as the number of its corresponding position in the reference sequence. For example, in a case where a variant has a deletion relative to an aligned reference sequence, there will be no amino acid in the variant that corresponds to a position in the reference sequence at the site of deletion. Where there is an insertion in an aligned reference sequence, that insertion will not correspond to a numbered amino acid position in the reference sequence. In the case of truncations or fusions there can be stretches of amino acids in either the reference or aligned sequence that do not correspond to any amino acid in the corresponding sequence.

The terms “numbered with reference to” or “corresponding to,” when used in the context of the numbering of a given amino acid or polynucleotide sequence, refers to the numbering of the residues of a specified reference sequence when the given amino acid or polynucleotide sequence is compared to a reference sequence. In embodiments the reference sequence is a *Candida intermedia* GXS1 protein having SEQ ID NO: 1. In embodiments, the comparison to the reference sequence is a sequence alignment between the given amino acid or polynucleotide sequence and the reference sequence.

“GXS1 protein” or “*Candida intermedia* GXS1 protein” is used according to its common, ordinary meaning and refers to proteins of the same or similar names and functional fragments and homologs thereof. The term includes recombinant or naturally occurring forms of GXS1 protein (e.g. Genbank ID: CAI44932.1; GI: 85057135; SEQ ID NO: 1), or variants thereof that maintain GXS1 protein activity (e.g. within at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 100% as compared to SEQ ID NO: 1). The term includes recombinant or naturally occurring forms of GXS1 protein or variants thereof that have sequence identity to SEQ ID NO: 1 (e.g. about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or 100% identity to SEQ ID NO: 1). GXS1 protein may refer to variants having mutated amino acid residues that modulate (e.g. increase or decrease when compared to GXS1 protein) GXS1 protein activity, expression, cellular targeting, or protein translocation. GXS1 protein may be modified as described herein (e.g. modified with a transporter motif sequence and/or glucose mitigation mutation).

SEQ ID NO: 1:
 MGLEDNRMVK RFVNVGEKKA GSTAMAIIVG LFAAS**GGVLF**
GYDTGTISGV MTMDYVLARY PSNKHSFTAD ESSLIVSILS
 VGTFFGALCA PFLNDTLGRR WCLILSALIV FNIGAILQVI
 STAIPLLCAG RVIAGFGVGL ISATIPLYQS ETAPKWIRGA
 IVSCYQWAIT IGLFLASCVN KGTEHMTNSG SYRIPLAIQC
 LWGLILGIGM IFLPETPRFW ISKGN**QEKAA** ESLARLRKLP
 IDHPDSLEEL RDITAAYEFE TVYKGSSWSQ VFSHKNHQLK
 RLFTGVAIQA FQQLTGVNFI FYYGTTFFKR AGVNGFTISL
 ATNIV**NV**GVST IPGILLMEVL GRRNMLMGGG TGM**SL**SQILV
 AIVGVATSEN NKSSQSVLVA FSCIFIAFFA ATWGPCA~~WVV~~
 VGELF**P**LRT AKSVSLCTAS NWLWNWGIAY ATPYMVDEDK
 GNLGS**NV**FFI WGGFNLACVF FAWYFIYETK GLSLEQVDEL
 YEHVSKAWKS KGFVPSKHSF REQVDQQMS KTEAIMSEEA
 SV

Residues corresponding to positions 36-41 are underlined and bolded for reference. Residues corresponding to positions 155, 225, 326, 354, 361, 407 and 446 are underlined for reference.

The term “recombinant” when used with reference to, for example, a cell, nucleic acid, or protein, indicates that the cell, nucleic acid, or protein, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express genes otherwise modified from those found in the native form of a cell (e.g. genes encoding a mutation in a native or non-native transporter protein, such as a transporter motif sequence as described herein). For example, a recombinant protein may be a protein that is expressed by a cell or organism that has been modified by the introduction of a heterologous nucleic acid (e.g. encoding the recombinant protein).

The word “expression” or “expressed” as used herein in reference to a DNA nucleic acid sequence (e.g. a gene) means the transcriptional and/or translational product of that sequence. The level of expression of a DNA molecule in a cell may be determined on the basis of either the amount of corresponding mRNA that is present within the cell or the amount of protein encoded by that DNA produced by the cell (Sambrook et al., 1989 *Molecular Cloning: A Laboratory Manual*, 18.1-18.88).

The term “gene” means the segment of DNA involved in producing a protein; it includes regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons). The leader, the trailer as well as the introns include regulatory elements that are necessary during the transcription and the translation of a gene. Further, a “protein gene product” is a protein expressed from a particular gene.

The term “isolated” refers to a nucleic acid, polynucleotide, polypeptide, protein, or other component that is partially or completely separated from components with which it is normally associated (other proteins, nucleic acids, cells,

etc.). In embodiments, an isolated polypeptide or protein is a recombinant polypeptide or protein.

A “yeast cell” as used herein, refers to a eukaryotic unicellular microorganism carrying out metabolic or other function sufficient to preserve or replicate its genomic DNA. Yeast cells may carry out fermentation of sugars described herein. In embodiments, fermentation may convert the sugar to a biofuel or biochemical as set forth herein. Yeast cells referenced herein include, for example, the following species: *Candida intermedia*, *Cryptococcus neoformans*, *Debaryomyces hansenii*, *Saccharomyces cerevisiae*, *Scheffersomyces stipitis*, or *Yarrowia lipolytica*. A “recombinant yeast cell” is a yeast cell which expresses a recombinant transporter protein.

The term “biofuel” as used herein refers to a convenient energy containing substance produced from living organisms (e.g. biomass conversion to a fuel). Thus, biofuels may be produced through, for example, fermentation of carbohydrates (e.g. sugars) found in biomass (e.g. lignocellulosic biomass). Biofuels may be solid, liquid, or gas forms. Biofuels include, for example, ethanol, biodiesel, vegetable oil, ether (oxygenated fuels), or gas (e.g. methane).

The term “biochemical” as used herein refers to production of chemicals by living organisms. Biochemicals herein include production of alcohols (e.g. methanol, butanol, ethanol, isobutanol, 2,3-butanediol, propanol); sugars (e.g. erythritol, mannitol, riboflavin); carotenoids (e.g. β -carotene, lycopene, astaxanthin); fatty acids (e.g. ricinoleic acid, linolenic acid, tetracetyl phytosphingosine); amino acids (e.g. valine, lysine, threonine); aromatics (e.g. indigo, vanillin, styrene, p-hydroxystyrene); flavonoids (e.g. naringenin, genistein, kaempferol, quercetin, chrysin, apigenin, luteolin); stillbenoids (e.g. resveratrol); terpenoids (e.g. β -amyryn, taxadiene, miltiradiene, paclitaxel, artemisinin, bisabolane); polyketides (e.g. aureothin, spectinabilin, lovastatin, geodin); acetone; or organic acids (e.g. citric acid, succinic acid, malic acid, lactic acid, polylactic acid, adipic acid, glucaric acid). See e.g. Curran K. A., Alper H. S., *Metabolic Engineering* 14:289-297 (2012).

A “transporter motif sequence” as used herein refers to an amino acid sequence that, when present in a protein (e.g. a sugar transporter protein such as a MFS transporter protein), increases the ability of the protein to transport a sugar or sugar-containing compound into a yeast cell. In embodiments, the transporter motif sequence imparts a hexose sugar transport preference or pentose sugar transport preference to the protein. Correspondingly, a transporter motif may refer to the specific sugar it transports into a yeast cell. For example, the transporter motif sequence may impart preference to hexose sugars to a transporter protein, thereby allowing the recombinant transporter protein to preferentially transport hexoses into a yeast cell. Such transporter motif sequences may be referred to herein as a “hexose transporter motif sequence.” In embodiments the transporter motif sequence imparts preference to a single hexose. The hexose may be galactose or mannose. Such transporter motif sequences may be referred to herein as a “galactose transporter motif sequence” and a “mannose transporter motif sequence” respectively. In embodiments, the transporter motif sequence imparts preference to more than one hexose sugar.

The transporter motif sequence may impart preference to pentose sugars to a transporter protein, thereby allowing the recombinant transporter protein to preferentially transport pentose into a yeast cell. In embodiments the transporter motif sequence imparts preference to a single pentose (e.g. xylose). The pentose may be xylose or arabinose. Such

transporter motif sequences may be referred to herein as a “xylose transporter motif sequence” and an “arabinose transporter motif sequence” respectively. In embodiments the transporter motif sequence imparts preference to more than one pentose sugar (e.g. xylose and arabinose). Such transporter motif sequences may be referred to as a “xylose/arabinose transporter motif sequence.”

In embodiments, the transporter motif sequence imparts preference to a hexose and a pentose. That is, in embodiments, the recombinant transporter protein having such a transporter motif sequence preferentially transports one hexose and/or one pentose. Such transporter motif sequences may be referred to by the sugars which are transported (e.g. galactose and arabinose). Accordingly, in embodiments, the transporter motif sequence imparts preference to galactose and arabinose. Such a transporter motif is herein referred to as a “galactose-arabinose transporter motif sequence” (i.e. a recombinant transporter protein that transports both galactose and arabinose, or transports galactose or arabinose).

The transporter motif sequence as described herein corresponds to residues corresponding to positions 36-41 of the *Candida intermedia* GXSI protein (“GXSI motif sequence”). One skilled in the art will immediately recognize the identity and location of residues corresponding to positions 36-41 of the *Candida intermedia* GXSI protein in other transporter proteins with different numbering systems. For example, by performing a simple sequence alignment with *Candida intermedia* GXSI protein the identity and location of residues corresponding to positions 36-41 of the *Candida intermedia* GXSI protein are identified in other yeast transport proteins as illustrated in Table 1. Insertion (e.g. substitution) of a transporter motif sequence into a yeast transport protein may thereby be performed resulting in a functional yeast transporter protein with an altered sugar transport preference (e.g. changing a preference for hexoses to a preference for pentoses). For example, amino acid residue positions 75-81 of *S. cerevisiae* HXT7 protein correspond to amino acid residue positions 36-41 of the *Candida intermedia* GXSI protein (see Table 1).

TABLE 1

Sequence alignment of 54 sequences from major facilitator superfamily sugar transporter proteins (SEQ ID NOs: 51-104, respectively in order of appearance in Table 1). Putative transporter motif sequences are illustrated in the box and corresponds as described herein to residue positions 36-41 of *C. intermedia* GXSI protein.

Dh2C02530p	KFRNFLDKTPNIYNVPIASISCSISGLM	PGIDISSMSLFIGDDKIKYFKH-----	63
Dh2E01166p	KLRLFLDKLPNIYNIYVIATISCSISGLM	PGIDISSMSAFLSNDAYLKYFGT-----	63
Dh2E01298p	KFRNFLDKFPNIHNVYIVVGISCSISGM	PGIDISSMSLFIGDDKLYDFNS-----	63
SsHGT2	KPRTFLDRLPNIYNVYIIASISCSISGM	PGFIDISSMSAFIGEDDYKNFFMN-----	63
Dh2A14300p	SLNKELDKPHTTYNIYVIAMITISGM	PGFVSSISAFISEPSYRRFFNY-----	61
Y10B06391p	QVGALQHRFPKLNHPYLTAAVATMGGL	LFGFDISSVSFAFVDTKPKYKEYFGY-----	59
Y10B01342p	-----MYKVHNPYLTAAVATMGML	PGFDISSVSFAFVGEDNYMNYFGH-----	43
BmHGT2	-----MGRITNPVYVLTALACTGGL	LFGFDISSMSAISSPNYLTYFPGKDLTVECPD	52
At5g59250	LASDAPESFWSWVILPFIFF-ALGGL	IFGYDIGATSGATLSLQSP---ALSGTITWVNF	139
At5g17010	---HVPENYSVVAAILPFLFP-ALGGL	LYGYEIGATSCATISLQEPMTLLSYAVPFSAV	89
SsAUT1	LNAEATNKWHIPRLIGVIALGSM	AAVQGMDESIVINGANLFYPKAFGVD---TMHNSD	161
Y10D00132	LNREITNKWDHPMKVYVVLVCCSL	AAVQGMDETIVINGANIFPAQFGIKEDSGVVSRSK	180
BmSTL1	----FLGMRGIKLNWAIGFAASAGFL	IFGYDQGVLSGLYTLPSWNAQFPEINTAAVGDS	73
SsXUT6	AKTNSYLGLRGHKLNFVAVSCFAGV	GFLIFGYDQGVMSLTLPSFENTFFPAMK-----	75
Dh2E01386p	--KNTMGLRGKPLRVAITICCTIGF	SLFGYDQGLMSGIITGKQFNEEPPTHGT-----	59
Dh2B05060p	--RTNIMGLRGKRLRVMFTVAVTLG	FSLFGYDQGLMSGITGQFNAAEPPTAGK-----	60
SsSTL1	--RRNRMGLRGKRLRVMFTVAVTLG	FSLFGYDQGLMSGITGQFNAAEPPTAGK-----	60
ScSTL1	--RTSHWGLTGKRLRYITIASMTG	FSLFGYDQGLMASLITGKQFNYPFPATKENG---D	70
BmHXT10	----IDVGLRGNWLLTVITASCAP	GFLIFGYDNGVMGGVVLGEGFNKTFMNPD-----	56
SsXUT2	-----GKQVSYAVTFTCELAFIL	IFGIEQGIIGNLINNQDFLNTFNGPTG-----	53
CnBC3990p	--HKTQRRLVGHNLLYSVSVFLSIG	VWVIFGYDQGVMSGIITGPFYKAYFPNQPTS-----	62
Y10F06776p	----MFSLTGKPLLYFTSVFVSLG	VVFLFGYDQGVMSGIITGPFYKAYPEHPTR-----	49
BmXUT3	VGATGAKGLIKNARTFAIAVFASMG	GGLIYGYNQGMFGQILSMHSFQEAASGVKGIT-----	78
SsXUT1	AGKSGVAGLVANSRSPFIIVFASL	GGLVYGYNQGMFGQISGMYSFKAIVEIKIQD---	77
SsXUT3	AHGNVVTIMMKDPVVFVLVILFASL	GGLIFGYDQGVISGIVTMESF--GAKFPRIFM----	63
SsXUT3-A	AHGNVVTIMMKDPVVFVLVILFASL	GGLIFGYDQGVISGIVTMESF--GAKFPRIFM----	63
SsXUT3-B	AHGNVVTIMMKDPVVFVLVILFASL	GGLIFGYDQGVISGIVTMESF--GAKFPRIFM----	63
DhXy1HP	SKGNIITVMSKDLPLVFCIIAFASIG	GGLIFGYDQGVISGIVTMESF--AAKFPRIFS----	64
ScGAL2	PIEIPKPMSEYVTVSLLCLCVAFGG	FMPGWDGTGISGFVVTDFLRRFG-MKHKDGTT--	113
ScHXT8	EVVVPEKPASAYATVSIIMCLMAF	GGFMSGWDTGTISGFVNQTDFLRRFGNYSHSKNT--	109
ScHXT1	AVAPPNTGKGVYVTVSICVMVAFGG	FIPGWDGTGTISGFVAQTDFLRRFG-MKHHDGSS--	107
ScHXT3	VLTNPNTGKGAAYTVSICVMVAFGG	FVFGWDTGTISGFVAQTDFLRRFG-MKHHDGSS--	104
ScHXT7	VVEIPKRPASAYVTVSICMIMIAF	GGFVFGWDTGTISGFINQTDFLRRFG-MKHKDGTT--	107
ScHXT9	PIDLPOKPLSAYTTVAIICLMIAF	GGFIFGWDGTGTISGFVNLSDFLRRFG-QKNDKGT--	103
ScHXT2	NAELPAKPIAAYTVIICLMIAFGG	FVFGWDTGTISGFVNQTDFKRRFG-QMKS DGT--	98
ScHXT10	SLDIPYKPIIAYVTVMGLCLMIAF	GGFIFGWDGTGTISGFINQTDFKRRFG-ELQRDGS--	91
CiGXF1	QVDAPQKGFVDYIVISIFCFMVA	GGFVFGFDTGTISGFVNMSDFKDRFG-QHHADGT--	86
ScHXT13	NVEPPKRGLIGYLVIIYLLCYPI	SFGGFLPGWDSGITAGFINMDFNKMNFSGSYKHS TGE--	100
BmGXP1	-MVFQVRGTPIGALTLFIAMLASM	GGFLFGWDTGQISGLTQMDPFRQRFATVNDPAIG-	58

TABLE 1-continued

Sequence alignment of 54 sequences from major facilitator superfamily sugar transporter proteins (SEQ ID NOs: 51-104, respectively in order of appearance in Table 1). Putative transporter motif sequences are illustrated in the box and corresponds as described herein to residue positions 36-41 of *C. intermedia* GXs1 protein.

ScHXT14	GQAAKISHNASLHIPVLLCLVISTGGFLFGWDIGTIGGMTNMVSPQEKFGTTNI IHDDET	105
BmGXs1	GPVARPASVKQSLPAILVAAASAFGGVLFYDGTGISGLIVMPNQETFGKVPVPGSTTGA	74
BmRGT2	GPVARPASVKQSLPAILVAAASAFGGVLFYDGTGISGLIVMPNQETFGKVPVPGSTTGA	74
CiGXs1	FVNVGEKKAGSTAMAIIVGLFAASGGVLFYDGTISGVMTMDYVLARY-----PSNK-	64
CiGXs1-A	FVNVGEKKAGSTAMAIIVGLFAASGGVLFYDGTISGVMTMDYVLARY-----PSNK-	64
CiGXs1-B	FVNVGEKKAGSTAMAIIVGLFAAFGGVLSGYDTGTISGVMTMDYVLARY-----PSNK-	64
Dh2D01474	YVNVGEKRAGSASMGIFVGAFAAFGGVLFYDGTISGIMAMNYVKGEF-----PANK-	64
Dh0D02167p	YVNVGEKRAGSASMGIFVGAFAAFGGVLFYDGTISGIMAMNYVKGEF-----PANK-	64
SsRGT2	YINFGKKAGSTTMGICVGLFAAFGGILFGYDGTISGIMAMDYVTARF-----PSNH-	64
Y10C06424p	IINRGEKPEGSAFMAAFVAVFVAFGGILFGYDGTISGVMAMPFVKKTF-----TDDG-	58
Y10C08943p	-----MAIIVAVFVAFGGLLFYDGTIAGIMTMGYVKEHF-----TDFGK	41
Dh2B14278p	YYKMQQKS-SSSAITVGLVAAVGGFLFYDGTGLINDIMEMTYVKDNF-----PANG-	69
EcXy1E	-----MNTQYNSSYIFSI TLVATLGGLLFGYDTAVISGTVESLHTVFPVAPQNLSESAAN-	54
SsXUT5	RSIGPLIPRNKHLFYGSVLLMSIVHPTIMGYDSMMVGSILNLDAVNNYFH-----	53
ScMAL11	KSMTLKQALLKYPKAALWSILVSTLVMREGYDTALLSALYALPVPQRKFGTLNGBGS---	148

A “glucose mitigation mutation” as used herein refers to an amino acid mutation that, when present in a recombinant transporter protein, reduces, minimizes, diminishes, or in certain embodiments, eliminates the inhibitory effect of glucose on the recombinant transporter when transporting a sugar other than glucose (e.g. xylose) into a yeast cell. A glucose mitigation mutation may, in embodiments, increase the ability of a recombinant transporter protein to transport a preferred sugar or sugar-containing compound into a yeast cell. Thus, in embodiments, a glucose mitigation mutation may increase the ability of a recombinant transporter protein to transport xylose into a yeast cell. A glucose mitigation mutation may include a single amino acid residue mutation (e.g. a “point mutation”) in a recombinant transporter protein. A glucose mitigation mutation may include two or more mutations (e.g. a “substitution set”) in a recombinant transporter protein. The glucose mitigation mutation may be in a transmembrane domain, an extracellular loop, or cytoplasmic loop of a recombinant transporter protein. In embodiments, the glucose mitigation mutation may be localized (i.e. glucose mitigation mutations located within a specified domain or region of a recombinant transporter protein) or distributed (i.e. glucose mitigation mutations located throughout the sequence of the recombinant transporter protein).

A “recombinant transporter protein” as used herein refers to a recombinantly expressed transmembrane protein which transports a sugar or sugar-containing compound (e.g. hexoses and pentoses) into a yeast cell. In embodiments, the recombinant transporter protein is a yeast recombinant transporter protein. In embodiments, the recombinant transporter protein is a transporter protein belonging to the major facilitator superfamily (“MFS”) transporter proteins. In embodiments, a recombinant transporter protein may transport a hexose (e.g. galactose) into a yeast cell. In embodiments, a recombinant transporter protein may transport a pentose (e.g. xylose or arabinose) into a yeast cell. A recombinant transporter protein may be engineered, using the transporter motif sequences described herein, to alter its sugar preference (e.g. a transporter protein having a preference to transport a hexose compound may be converted to a transporter protein having a preference to transport a pentose compound). A recombinant transporter protein may be characterized by the sugar it transports. Thus, a recombinant

transporter protein transporting xylose is herein referred to a “recombinant xylose transporter protein.” Likewise, recombinant transporter proteins transporting arabinose or galactose are herein referred to as a “recombinant arabinose transporter protein” and a “recombinant galactose transporter protein” respectively.

A recombinant transporter protein may be characterized as a transporter protein derived from a particular organism. Where a recombinant transporter protein is derived from a particular organism, the endogenous sequence of the recombinant transporter protein may be maintained and residues corresponding to positions 36-41 of the *Candida intermedia* GXs1 protein may be replaced with a transporter motif sequence. As an example, a *C. intermedia* gxsl transporter protein is a gxsl transporter protein, a homolog thereof, or a functional fragment thereof, found in *C. intermedia*. Amino acids 75-81 of *S. cerevisiae* hxt7 transporter protein may be replaced with a transporter motif sequence thereby forming a recombinant transporter protein with desired sugar transport characteristics as described herein. In embodiments, the recombinant transporter protein is a protein, functional fragment, or homolog thereof, identified by the following NCBI gene ID or NCBI accession numbers: 836043, 831564, AJ937350.1, AJ875406.1, 2901237, 2913528, 8998057, 8999011, 50419288, 948529, 4839826, 4852047, 4851844, 4840896, 4840252, 4841106, 4851701, 2907283, 2906708, 2908504, 2909312, 2909701, 4935064, 851943, 856640, 851946, 856494, 8998297, 2902950, 2902912, 853207, 852149, 855023, 853216, 853236, 850536, 855398, 4836720, 4836632, 4840859, 2913215, 2902914, 2910370, and 4838168 (SEQ ID NOs:7-50, respectively in order of appearance). Such recombinant transporter proteins may further be characterized by the sugar preference conferred (e.g. a *Candida intermedia* GXs1 recombinant xylose transporter protein).

A “pentose compound” or “pentose” is a monosaccharide-containing compound having 5 carbon atoms. Pentose compounds include aldopentoses (e.g. pentose compounds having an aldehyde moiety at carbon 1) and ketopentoses (e.g. pentose compounds having a ketone moiety at carbon 2 or carbon 3). Pentose compounds include, for example, D/L-arabinose, D/L-lyxose, D/L-ribose, D/L-xylose, D/L-ribulose, and D/L-xylulose. The term “monosaccharide-containing” refers to a compound that includes at least one monosaccharide.

13

A "hexose compound" or "hexose" is a monosaccharide-containing compound having 6 carbon atoms. Hexose compounds include aldohexoses (e.g. hexose compounds having an aldehyde moiety at carbon 1) and ketohexoses (e.g. hexose compounds having a ketone moiety at carbon 2). Hexose compounds include, for example, D/L-allose, D/L-altrose, D/L-glucose, D/L-mannose, D/L-galactose, and D/L-talose.

A "xylose compound" is xylose or a xylose-containing compound including at least one xylose moiety. Thus as used herein, the term xylose compound represents a single xylose, a chain including one or more xylose moieties, or a xylose moiety covalently or non-covalently bound to another chemical moiety (e.g. another sugar forming a xylose containing polysaccharide or xylose bound to lignin). An "arabinose compound" is arabinose or an arabinose-containing compound including at least one arabinose moiety. Thus as used herein, the term arabinose compound represents a single arabinose, a chain including one or more arabinose moieties, or an arabinose moiety covalently or non-covalently bound to another chemical moiety (e.g. another sugar forming an arabinose containing polysaccharide or arabinose bound to lignin). A "galactose compound" is galactose or a galactose-containing compound including at least one galactose moiety. Thus as used herein, the term galactose compound represents a single galactose, a chain including one or more galactose moieties, or a galactose moiety covalently or non-covalently bound to another chemical moiety (e.g. another sugar forming a galactose containing polysaccharide or bound to lignin).

A "sugar" as set forth herein, refers to monosaccharide and polysaccharide compounds metabolized by a yeast cell. In embodiments, a sugar may be a hexose sugar as described herein or a pentose sugar as described herein.

Polysaccharides herein include hexose-only polysaccharides, pentose-only polysaccharides, and hexose-pentose mixture polysaccharides. In embodiments, the xylose compound, the arabinose compound, or the galactose compound may be derived from or form part of a lignocellulosic biomass (e.g. plant dry matter that may be used in as a source for pentose compounds or hexose compounds and for production of biofuels or biochemicals), hemicellulose, marine biomass (e.g. seaweeds or algae that may be used in as a source for pentose compounds or hexose compounds and for production of biofuels or biochemicals) or other natural or synthetic sources for xylose, arabinose, or galactose, including but not limited to xylan or pectin. "Derived from" refers to extraction, removal, purification, or otherwise freeing a xylose compound, arabinose compound, or galactose compound from a source (e.g. lignocellulosic biomass) by either chemical processes (e.g. acid hydrolysis, ammonium explosion, or ionic liquids extraction) or through natural biological processes by organisms capable of using such sources for energy.

A "xylose growth media" refers to a yeast cell media containing a xylose compound in amounts sufficient to serve as a nutrient for growing or culturing recombinant yeast cells. The term refers to a media substantially free of glucose, and, in embodiments, is "glucose free" (i.e. the media contains no glucose). In embodiments, a xylose growth media includes trace amounts of glucose which are undetectable using known methods and which are insufficient to support significant growth of yeast cells. In embodiments, a xylose growth media includes trace amounts of glucose which are insufficient to cause inhibition of activity of a recombinant transporter protein (e.g. a recombinant xylose transporter protein) as described herein.

14

A "xylose-glucose growth media" refers to a yeast cell media containing a xylose compound in an amount sufficient to serve as a nutrient for growing or culturing recombinant yeast cells and a glucose compound. The term refers to a media that includes glucose in an amount sufficient to serve as a nutrient for growth or culturing recombinant yeast cells or in an amount sufficient to cause inhibition of activity of a recombinant transporter protein as described herein. The glucose may be present in the xylose-glucose growth media at a pre-determined concentration as described herein.

Xylose growth media and xylose-glucose growth media may be supplemented with other hexoses or pentoses described herein (e.g. mannose, galactose, or arabinose). Growth of a recombinant yeast cell in a xylose growth media may be compared to growth of a recombinant yeast cell in a xylose-glucose growth media. Thus, in embodiments, a recombinant xylose transporter protein may be selected for its xylose selectivity and/or its rate of transfer of a xylose compound into a yeast cell by comparing its growth in xylose growth media to its growth in xylose-glucose growth media. In embodiments, recombinant yeast cells having impaired growth in xylose-glucose growth media may indicate that the recombinant xylose transporter protein in the recombinant yeast is inhibited, at least in part, by glucose.

An "arabinose growth media" refers to a yeast cell media containing an arabinose compound in amounts sufficient to serve as a nutrient for growing or culturing recombinant yeast cells. The term refers to a media substantially free of glucose, and, in embodiments, is "glucose free" (i.e. the media contains no glucose). In embodiments, an arabinose growth media includes trace amounts of glucose which are undetectable using known methods and which are insufficient to support significant growth of yeast cells. In embodiments, an arabinose growth media includes trace amounts of glucose which are insufficient to cause inhibition of activity of a recombinant transporter protein (e.g. a recombinant arabinose transporter protein) as described herein.

An "arabinose-glucose growth media" refers to a yeast cell media containing an arabinose compound in an amount sufficient to serve as a nutrient for growing or culturing recombinant yeast cells and a glucose compound. The term refers to a media that includes glucose in an amount sufficient to serve as a nutrient for growth or culturing recombinant yeast cells or in an amount sufficient to cause inhibition of activity of a recombinant transporter protein as described herein. The glucose may be present in the arabinose-glucose growth media at a pre-determined concentration as described herein.

Arabinose growth media and arabinose-glucose growth media may be supplemented with other hexoses or pentoses described herein (e.g. mannose, galactose, or xylose). Growth of a recombinant yeast cell in an arabinose growth media may be compared to growth of a recombinant yeast cell in an arabinose-glucose growth media. Thus, in embodiments, a recombinant arabinose transporter may be selected for its arabinose selectivity and/or its rate of transfer of an arabinose compound into a yeast cell by comparing its growth in arabinose growth media to its growth in arabinose-glucose growth media. In embodiments, recombinant yeast cells having impaired growth in arabinose-glucose growth media may indicate that the recombinant arabinose transporter protein in the recombinant yeast is inhibited, at least in part, by glucose.

A "galactose growth media" refers to a yeast cell media containing a galactose compound in amounts sufficient to serve as a nutrient for growing or culturing recombinant yeast cells. The term refers to a media substantially free of

glucose, and, in embodiments, is “glucose free” (i.e. the media contains no glucose). In embodiments, a galactose growth media includes trace amounts of glucose which are undetectable using known methods and which are insufficient to support significant growth of yeast cells. In embodiments, a galactose growth media includes trace amounts of glucose which are insufficient to cause inhibition of activity of a recombinant transporter protein (e.g. a recombinant galactose transporter protein) as described herein.

A “galactose-glucose growth media” refers to a yeast cell media containing a galactose compound in an amount sufficient to serve as a nutrient for growing or culturing recombinant yeast cells and a glucose compound. The term refers to a media that includes glucose in an amount sufficient to serve as a nutrient for growth or culturing recombinant yeast cells or in an amount sufficient to cause inhibition of activity of a recombinant transporter protein as described herein. The glucose may be present in the galactose-glucose growth media at a pre-determined concentration as described herein.

Galactose growth media and galactose-glucose growth media may be supplemented with other hexoses or pentoses described herein (e.g. mannose, arabinose, or xylose). Growth of a recombinant yeast cell in a galactose growth media may be compared to growth of a recombinant yeast cell in a galactose-glucose growth media. Thus, in embodiments, a recombinant galactose transporter may be selected for its galactose selectivity and/or its rate of transfer of a galactose compound into a yeast cell by comparing its growth in galactose growth media to its growth in galactose-glucose growth media. In embodiments, recombinant yeast cells having impaired growth in galactose-glucose growth media may indicate that the recombinant galactose transporter protein in the recombinant yeast is inhibited, at least in part, by glucose.

As defined herein, the term “inhibition”, “inhibit”, “inhibiting” and the like refers to negatively affecting (e.g. decreasing) the activity or function of a recombinant transporter protein (e.g. recombinant xylose transporter protein) relative to the activity or function of the protein in the absence of the inhibitor (e.g. glucose). In embodiments, inhibition refers to a reduction in the growth rate of a recombinant yeast cell.

“Contacting” is used in accordance with its plain ordinary meaning and refers to the process of allowing at least two distinct species (e.g. chemical compounds including sugars, biomolecules or cells) to become sufficiently proximal to react, interact or physically touch. The term “contacting” includes allowing two species to react, interact, or physically touch, where the two species may be a sugar as described herein and a recombinant transporter protein as described herein. In embodiments contacting includes allowing a sugar described herein to interact with a recombinant transporter protein that is involved in transporting hexose or pentose compounds into a yeast cell.

I. Compositions

Provided herein are recombinant transporter proteins that include a transporter motif sequence and a glucose mitigation mutation. In one aspect, the recombinant transporter protein is a recombinant xylose transporter protein that includes a xylose transporter motif sequence and a glucose mitigation mutation.

1. Recombinant Xylose Transporter Protein

The xylose transporter motif sequence may correspond to amino acid residue positions 36, 37, 38, 39, 40, and 41 of

Candida intermedia GXS1 protein (SEQ ID NO: 1). In embodiments, the xylose transporter motif sequence corresponds to amino acid residue positions 36, 37, 38, 39, 40, and 41 of SEQ ID NO:1. The transporter motif sequence may have the sequence -G-G/F-X¹-X²-X³-G-, X¹ is D, C, G, H, I, L, or F. X² is A, D, C, E, G, H, or I. X³ is N, C, Q, F, G, L, M, S, T, or P. In embodiments, the transporter motif sequence is not -G-G-L-I-F-G- (SEQ ID NO:105) or -G-G-F-I-F-G- (SEQ ID NO:106).

X¹ may be D, C, G, I, L, or F. X¹ may be D, C, G, H, or F. X¹ may be D. X¹ may be C. X¹ may be G. X¹ may be I. X¹ may be L. X¹ may be H. X¹ may be F. X² may be D, C, E, G, H, or I. X² may be E, G, H, or I. X² may be H or I. X² may be H. X² may be I. X³ may be N, Q, F, M, S, T, or P. X³ may be F, M, S, or T. X³ may be S, T, or M. X³ may be T. X³ may be S. X³ may be M. In embodiments, when X¹ is F, X² may be I and X³ may be M or S.

The xylose transporter motif sequence may be -G-G-F-I-M-G- (SEQ ID NO:107), -G-F-F-I-M-G- (SEQ ID NO:108), -G-G-F-I-S-G- (SEQ ID NO:109), -G-F-F-I-S-G- (SEQ ID NO:110), -G-G-F-I-T-G- (SEQ ID NO:111), -G-F-F-I-T-G- (SEQ ID NO:112), -G-G-F-L-M-G- (SEQ ID NO:113), -G-F-F-L-M-G- (SEQ ID NO:114), -G-G-F-L-S-G- (SEQ ID NO:115), -G-F-F-L-S-G- (SEQ ID NO:116), -G-G-F-L-T-G- (SEQ ID NO:117), -G-F-F-L-T-G- (SEQ ID NO:118), -G-G-F-H-M-G- (SEQ ID NO:119), -G-F-F-H-M-G- (SEQ ID NO:120), -G-G-F-H-S-G- (SEQ ID NO:121), -G-F-F-H-S-G- (SEQ ID NO:122), -G-G-F-H-T-G- (SEQ ID NO:123) or -G-F-F-H-T-G- (SEQ ID NO:124). In embodiments, the xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ ID NO:107), -G-F-F-I-M-G- (SEQ ID NO:108), -G-G-F-I-S-G- (SEQ ID NO:109), -G-F-F-I-S-G- (SEQ ID NO:110), -G-G-F-I-T-G- (SEQ ID NO:111), or -G-F-F-I-T-G- (SEQ ID NO:112). In embodiments, the xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ ID NO:107), -G-F-F-I-M-G- (SEQ ID NO:108), -G-G-F-I-S-G- (SEQ ID NO:109), or -G-F-F-I-S-G- (SEQ ID NO:110). In embodiments, the xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ ID NO:107), or -G-F-F-I-M-G- (SEQ ID NO:108). The xylose transporter motif sequence may be -G-G-F-I-M-G- (SEQ ID NO:107). The xylose transporter motif sequence may be -G-F-F-I-M-G- (SEQ ID NO:108). The xylose transporter motif sequence may be -G-G-F-I-S-G- (SEQ ID NO:109). The xylose transporter motif sequence may be -G-F-F-I-S-G- (SEQ ID NO:110). The xylose transporter motif sequence may be -G-G-F-I-T-G- (SEQ ID NO:111). The xylose transporter motif sequence may be -G-F-F-I-T-G- (SEQ ID NO:112). The xylose transporter motif sequence may be -G-G-F-L-M-G- (SEQ ID NO:113). The xylose transporter motif sequence may be -G-F-F-L-M-G- (SEQ ID NO:114). The xylose transporter motif sequence may be -G-G-F-L-S-G- (SEQ ID NO:115). The xylose transporter motif sequence may be -G-F-F-L-S-G- (SEQ ID NO:116). The xylose transporter motif sequence may be -G-G-F-L-T-G- (SEQ ID NO:117). The xylose transporter motif sequence may be -G-F-F-L-T-G- (SEQ ID NO:118). The xylose transporter motif sequence may be -G-G-F-H-M-G- (SEQ ID NO:119). The xylose transporter motif sequence may be -G-F-F-H-M-G- (SEQ ID NO:120). The xylose transporter motif sequence may be -G-G-F-H-S-G- (SEQ ID NO:121). The xylose transporter motif sequence may be -G-F-F-H-S-G- (SEQ ID NO:122). The xylose transporter motif sequence may be -G-G-F-H-T-G- (SEQ ID NO:123). The xylose transporter motif sequence may be -G-F-F-H-T-G- (SEQ ID NO:124).

The glucose mitigation mutation may be within a protein domain corresponding to a transmembrane of a recombinant

transporter protein (e.g. one or more of transmembrane domains 1-12). The glucose mitigation mutation may be within two or more protein domains corresponding to transmembranes of a recombinant transporter protein. The glucose mitigation mutation may be within a protein domain corresponding to a transmembrane of *Candida intermedia* GXS1 protein. The transmembrane may be a protein domain corresponding to transmembrane 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 of *Candida intermedia* GXS1 protein. The transmembrane may be a protein domain corresponding to a transmembrane a protein of SEQ ID NO:51-104. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 1 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 2 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 3 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 4 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 5 of *Candida intermedia* GXS1 protein.

In embodiments, the glucose mitigation mutation is within a protein domain corresponding to transmembrane 5 of SEQ ID NO:1. In embodiments, the glucose mitigation mutation is within a protein domain corresponding to residue 160-179 of SEQ ID NO:1. In embodiments, the glucose mitigation mutation is at a position corresponding to T170 or I171 of SEQ ID NO:1. In embodiments, the glucose mitigation mutation is a T170N mutation. In embodiments, the glucose mitigation mutation is a I171F mutation.

The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 6 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 7 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 8 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 9 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 8 or 9 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 8 and 9 of *Candida intermedia* GXS1 protein (e.g. amino acid residues about 347 to about 366 of *Candida intermedia* GXS1 amino acid sequence (SEQ ID NO: 1)). The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 10 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 11 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 12 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to transmembrane 9 of *Candida intermedia* GXS1.

In embodiments, the glucose mitigation mutation is within a protein domain corresponding to an extracellular domain of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be in a protein domain corresponding to the 11-12 extracellular domain of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be within a protein domain corresponding to a cytoplasmic domain (i.e. intracellular) of *Candida intermedia* GXS1 protein. In embodiments, the glucose mitigation mutation is

in a protein domain corresponding to the 4-5 cytoplasmic domain (i.e. an intracellular domain between protein domains corresponding to transmembranes 4 and 5 of *Candida intermedia* GXS1 protein); the central cytoplasmic domain (i.e. an intracellular domain between protein domains corresponding to transmembranes 6 and 7 of *Candida intermedia* GXS1 protein); or the 10-11 cytoplasmic domain (i.e. an intracellular domain between protein domains corresponding to transmembranes 10 and 11 of *Candida intermedia* GXS1 protein) of *Candida intermedia* GXS1 protein.

The glucose mitigation mutation may be at a position corresponding to K155, T170, I171, N225, S354, A361, L407, or N446 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be at a position corresponding to T170 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be at a position corresponding to I171 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be at a position corresponding to K155 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be a conservative mutation at position 155 as described herein. In embodiments, the glucose mitigation mutation is a K155E mutation. The glucose mitigation mutation may be at a position corresponding to N225 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be a conservative mutation at position 225 as described herein. In embodiments, the glucose mitigation mutation is a N225D mutation. The glucose mitigation mutation may be at a position corresponding to S354 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be a conservative mutation at position 354 as described herein. In embodiments, the glucose mitigation mutation is a S354T mutation. The glucose mitigation mutation may be at a position corresponding to A361 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be a conservative mutation at position 361 as described herein. In embodiments, the glucose mitigation mutation is a A361T mutation. The glucose mitigation mutation may be at a position corresponding to L407 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be a conservative mutation at position 407 as described herein. In embodiments, the glucose mitigation mutation is a L407M mutation. The glucose mitigation mutation may be at a position corresponding to N446 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be a conservative mutation at position 446 as described herein. In embodiments, the glucose mitigation mutation is a N446S mutation. The glucose mitigation mutation may be at a position corresponding to N326 of *Candida intermedia* GXS1 protein. The glucose mitigation mutation may be a conservative mutation at position 326 as described herein. In embodiments, the glucose mitigation mutation is a N326S mutation. In embodiments, the glucose mitigation mutation is a N326H mutation.

The glucose mitigation mutation may include two or more positions corresponding to K155, T170, I171, N225, N326, S354, A361, L407, or N446 of *Candida intermedia* GXS1 protein as described hereinabove. The glucose mitigation mutation may include three or more positions corresponding to K155, T170, I171, N225, N326, S354, A361, L407, or N446 of *Candida intermedia* GXS1 protein as described hereinabove. The glucose mitigation mutation may include four or more positions corresponding to K155, T170, I171, N225, N326, S354, A361, L407, or N446 of *Candida intermedia* GXS1 protein as described hereinabove. The glucose mitigation mutation may include five or more posi-

tions corresponding to K155, T170, I171, N225, N326, S354, A361, L407, or N446 of *Candida intermedia* GXS1 protein as described hereinabove. The glucose mitigation mutation may include six or more positions corresponding to K155, T170, I171, N225, N326, S354, A361, L407, or N446 of *Candida intermedia* GXS1 protein as described hereinabove. The glucose mitigation mutation may include mutation at positions corresponding to K155, T170, I171, N225, N326, S354, A361, L407, and N446 of *Candida intermedia* GXS1 protein as described hereinabove. The glucose mitigation mutation may include mutation at positions corresponding to K155, T170, I171, N225, S354, A361, L407, and N446 of *Candida intermedia* GXS1 protein as described hereinabove.

The recombinant xylose transporter protein provided herein including embodiments thereof may further include an amino acid deletion. An amino acid deletion as provided herein is a deletion of at least one amino acid residue of a *Candida intermedia* GXS1 protein as described hereinabove. Thus, the sequence of a *Candida intermedia* GXS1 protein including an amino acid deletion includes at least one amino acid residue less relative to the sequence of a *Candida intermedia* GXS1 protein lacking said deletion. In embodiments, the deletion is at least 5 amino acids in length. In embodiments, the deletion is at least 10 amino acids in length. In embodiments, the deletion is at least 15 amino acids in length. In embodiments, the deletion is at least 20 amino acids in length. In embodiments, the deletion is at least 25 amino acids in length. In embodiments, the deletion is at least 30 amino acids in length. In embodiments, the deletion is at least 35 amino acids in length. In embodiments, the deletion is at least 40 amino acids in length. In embodiments, the deletion is at least 45 amino acids in length. In embodiments, the deletion is at least 50 amino acids in length. In embodiments, the deletion is at least 55 amino acids in length. In embodiments, the deletion is at least 60 amino acids in length. In embodiments, the deletion is at least 65 amino acids in length. In embodiments, the deletion is at least 70 amino acids in length. In embodiments, the deletion is at least 75 amino acids in length. In embodiments, the deletion is at least 80 amino acids in length. In embodiments, the deletion is at least 85 amino acids in length. In embodiments, the deletion is at least 90 amino acids in length. In embodiments, the deletion is at least 95 amino acids in length. In embodiments, the deletion is at least 100 amino acids in length.

In embodiments, the deletion is less than 50 amino acids in length. In embodiments, the deletion is less than 45 amino acids in length. In embodiments, the deletion is less than 40 amino acids in length. In embodiments, the deletion is less than 35 amino acids in length. In embodiments, the deletion is less than 30 amino acids in length. In embodiments, the deletion is less than 25 amino acids in length. In embodiments, the deletion is less than 20 amino acids in length. In embodiments, the deletion is less than 15 amino acids in length. In embodiments, the deletion is less than 10 amino acids in length. In embodiments, the deletion is within a protein domain corresponding to residue 497-522 of SEQ ID NO:1. In embodiments, the deletion is within a protein domain corresponding to residue 497-522 of a *Candida intermedia* GXS1 protein as described hereinabove.

2. Recombinant Arabinose Transporter Protein

Also provided herein is a recombinant arabinose transporter protein that includes an arabinose transporter motif sequence and a glucose mitigation mutation.

The arabinose transporter motif sequence may correspond to residue positions 36, 37, 38, 39, 40, and 41 of *Candida*

intermedia GXS1 protein. The arabinose transporter motif sequence may have the sequence -G-G-F-X⁴-X⁵-X⁶-G-. X⁴ is D, C, F, G, H, L, R, T, or P. X⁵ is A, C, E, F, H, K, S, P, or V. X⁶ is R, D, E, F, H, I, M, T, or Y. In embodiments, the arabinose transporter is not -G-G-L-V-Y-G- (SEQ ID NO:125), or -G-G-F-V-F-G- (SEQ ID NO:126).

X⁴ may be D, F, G, L, R, or T. X⁴ may be R, T, H, or F. X⁴ may be R. X⁴ may be T. X⁴ may be H. X⁴ may be F. X⁵ may be A, E, F, P, H, or V. X⁵ may be P, H, or V. X⁵ may be P. X⁵ may be H. X⁵ may be V. X⁶ may be T, H, F, M, or Y. X⁶ may be F or Y. X⁶ may be T or M. X⁶ may be T. X⁶ may be H. X⁶ may be F. X⁶ may be M. X⁶ may be Y. In embodiments, X⁴ is F or T, X⁵ is P or I, and X⁶ is M or T.

The arabinose transporter motif sequence may be -G-G-F-H-M-G- (SEQ ID NO:119), -G-F-F-H-M-G- (SEQ ID NO:120), -G-G-R-P-T-G- (SEQ ID NO:127), -G-F-R-P-T-G- (SEQ ID NO:128), -G-G-T-P-T-G- (SEQ ID NO:129), or -G-F-T-P-T-G- (SEQ ID NO:130). The arabinose transporter motif sequence may be -G-G-F-H-M-G- (SEQ ID NO:119), or -G-F-F-H-M-G- (SEQ ID NO:120). The arabinose transporter motif sequence may be -G-G-R-P-T-G- (SEQ ID NO:127), -G-F-R-P-T-G- (SEQ ID NO:128). The arabinose transporter motif sequence may be -G-G-T-P-T-G- (SEQ ID NO:129) or -G-F-T-P-T-G- (SEQ ID NO:130). The arabinose transporter motif sequence may be -G-G-F-H-M-G- (SEQ ID NO:119). The arabinose transporter motif sequence may be -G-F-F-H-M-G- (SEQ ID NO:120). The arabinose transporter motif sequence may be -G-G-R-P-T-G- (SEQ ID NO:127). The arabinose transporter motif sequence may be -G-F-R-P-T-G- (SEQ ID NO:128). The arabinose transporter motif sequence may be -G-G-T-P-T-G- (SEQ ID NO:129). The arabinose transporter motif sequence may be -G-F-T-P-T-G- (SEQ ID NO:130).

The glucose mitigation mutation of the recombinant arabinose transporter protein is as described hereinabove for the “recombinant xylose transporter protein” and includes embodiments thereof.

3. Recombinant Galactose Transporter Protein

Provided herein is a recombinant galactose transporter protein that includes an galactose transporter motif sequence and a glucose mitigation mutation.

The galactose transporter motif sequence is as described hereinabove for the “arabinose transporter motif sequence” and includes embodiments thereof. The glucose mitigation mutation of the recombinant arabinose transporter protein is as described hereinabove for the “recombinant xylose transporter protein” and includes embodiments thereof.

Also provided herein is a recombinant galactose-arabinose transporter protein that includes a galactose-arabinose transporter motif sequence and a glucose mitigation mutation. The galactose-arabinose transporter motif sequence may be as described hereinabove for the “arabinose transporter motif sequence” and includes embodiments thereof. The glucose mitigation mutation of the recombinant galactose-arabinose transporter protein is as described hereinabove for the “recombinant xylose transporter protein” and includes embodiments thereof.

II. Nucleic Acids

In another aspect is a nucleic acid encoding a recombinant xylose transporter protein described herein, including embodiments thereof. In yet another aspect is a nucleic acid encoding a recombinant arabinose transporter protein described herein, including embodiments thereof. In still another aspect is a nucleic acid encoding a recombinant galactose transporter protein described herein, including

embodiments thereof. In another aspect is a nucleic acid encoding a recombinant galactose-arabinose transporter protein described herein, including embodiments thereof. The nucleic acids may be RNA or DNA. The nucleic acids may be cDNA. The nucleic acids may be single- or double-stranded RNA or single- or double-stranded DNA. The nucleic acids may be located on a plasmid or other vector. The nucleic acids may be introduced and expressed by a yeast cell using conventional techniques known to those in the art.

III. Recombinant Yeast Cells

Provided herein are recombinant yeast cells that include a recombinant transporter protein as described herein, including embodiments thereof. Also provided herein are recombinant yeast cells that include a nucleic acid encoding a recombinant xylose transporter protein described herein, including embodiments thereof.

1. Recombinant Yeast Cell Including a Recombinant Xylose Transporter Protein

In one aspect is a recombinant yeast cell that includes a recombinant xylose transporter protein as described herein, including embodiments thereof. In embodiments, the growth rate of the recombinant yeast cell including a recombinant transporter protein as described herein can be measured. The growth rate may be determined in xylose growth media (i.e. in the absence of glucose). The growth rate may be determined in xylose-glucose growth media. In embodiments, the growth rates (i.e. growth rate in the absence and presence of glucose) are compared to determine the differential growth rate of the recombinant yeast cells. If the growth rate of recombinant yeast cells grown in xylose-glucose growth media is less than the growth rate of recombinant yeast cells grown in xylose growth media, the differential growth rate may indicate the presence of glucose inhibition of the recombinant transporter protein. As described herein, inclusion of a glucose mitigating mutation decreases, minimizes, or may eliminate glucose inhibition of a recombinant xylose transporter protein.

In embodiments, the growth rate of the recombinant yeast cell in a xylose-glucose growth media is about 5% to about 150% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 140% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 130% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 120% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 110% of the growth rate of the recombinant yeast cell in xylose growth media.

The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 100% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 90% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 80% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media

may be about 5% to about 70% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 60% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 50% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 40% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 30% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 20% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 5% to about 10% of the growth rate of the recombinant yeast cell in xylose growth media.

The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 150% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 140% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 130% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 120% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 110% of the growth rate of the recombinant yeast cell in xylose growth media.

The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 100% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 90% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 80% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 70% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 60% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 50% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 40% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 30% of the growth rate of the recombinant yeast cell in xylose growth media. The growth rate of the recombinant yeast cell in a xylose-glucose growth media may be about 10% to about 20% of the growth rate of the recombinant yeast cell in xylose growth media.

In embodiments, the xylose growth media includes xylose at a concentration of about 10 g/L to about 275 g/L. The xylose growth media may include xylose at a concentration of about 10 g/L to about 250 g/L. The xylose growth media may include xylose at a concentration of about 10 g/L to about 225 g/L. The xylose growth media may include xylose at a concentration of about 10 g/L to about 200 g/L. The xylose growth media may include xylose at a concentration of about 10 g/L to about 175 g/L. The xylose growth media may include xylose at a concentration of about 10 g/L to about 150 g/L. The xylose growth media may include xylose at a concentration of about 10 g/L to about 125 g/L. The xylose growth media may include xylose at a concentration of about 10 g/L to about 100 g/L. The xylose growth media may include xylose at a concentration of about 10 g/L to about 75 g/L. The xylose growth media may include xylose at a concentration of about 10 g/L to about 50 g/L. The xylose growth media may include xylose at a concentration of about 10 g/L to about 25 g/L.

The xylose growth media may include xylose at a concentration of about 125 g/L. The xylose growth media may include xylose at a concentration of about 150 g/L. The xylose growth media may include xylose at a concentration of about 175 g/L. The xylose growth media may include xylose at a concentration of about 200 g/L. The xylose growth media may include xylose at a concentration of about 225 g/L. The xylose growth media may include xylose at a concentration of about 250 g/L. The xylose growth media may include xylose at a concentration of about 275 g/L. The xylose growth media may include xylose at a concentration of about 300 g/L.

In embodiments, the xylose-glucose growth media includes xylose at a concentration as described herein for xylose growth media. In embodiments, the xylose-glucose growth media includes glucose at a concentration of about 0.05 g/L to about 20 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 15 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 10 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 5 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 4 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 3 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 2 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 1 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 0.5 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 0.1 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.1 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.5 g/L. The xylose-glucose growth media may include glucose at a concentration of about 1 g/L. The xylose-glucose growth media may include glucose at a concentration of about 2 g/L. The xylose-glucose growth media may include glucose at a concentration of about 3 g/L. The xylose-glucose growth media may include glucose at a concentration of about 4 g/L. The xylose-glucose growth media may include glucose at a concentration of about 5 g/L. The xylose-glucose growth media may include glucose at a concentration of about 10 g/L. The xylose-glucose growth media may include glucose

at a concentration of about 15 g/L. The xylose-glucose growth media may include glucose at a concentration of about 20 g/L.

In embodiments, the xylose-glucose growth media includes glucose at a concentration of about 0.05 g/L to about 100 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 90 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 80 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 70 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 60 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 50 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 40 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 30 g/L. The xylose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 25 g/L.

The xylose-glucose growth media may include glucose at a concentration of about 25 g/L. The xylose-glucose growth media may include glucose at a concentration of about 30 g/L. The xylose-glucose growth media may include glucose at a concentration of about 40 g/L. The xylose-glucose growth media may include glucose at a concentration of about 50 g/L. The xylose-glucose growth media may include glucose at a concentration of about 60 g/L. The xylose-glucose growth media may include glucose at a concentration of about 70 g/L. The xylose-glucose growth media may include glucose at a concentration of about 80 g/L. The xylose-glucose growth media may include glucose at a concentration of about 90 g/L. The xylose-glucose growth media may include glucose at a concentration of about 100 g/L.

2. Recombinant Yeast Cell Including a Recombinant Arabinose Transporter Protein

In another aspect is a recombinant yeast cell that includes a recombinant arabinose transporter protein as described herein, including embodiments thereof. In embodiments, the growth rate of the recombinant yeast cell including a recombinant arabinose transporter protein as described herein can be measured. The growth rate may be determined in arabinose growth media (i.e. in the absence of glucose). The growth rate may be determined in arabinose-glucose growth media. In embodiments, the growth rates (i.e. growth rate in the absence and presence of glucose) are compared to determine the differential growth rate of the recombinant yeast cells. If the growth rate of recombinant yeast cells grown in arabinose-glucose growth media is less than the growth rate of recombinant yeast cells grown in arabinose growth media, the differential growth rate may indicate glucose inhibits the activity of the recombinant arabinose transporter protein. As described herein, inclusion of a glucose mitigating mutation decreases, minimizes, or may eliminate glucose inhibition of the recombinant arabinose transporter protein.

In embodiments, the growth rate of the recombinant yeast cell in an arabinose-glucose growth media is about 5% to about 150% of the growth rate of the recombinant yeast cell in arabinose growth media. The growth rate of the recombinant yeast cell in an arabinose-glucose growth media may be about 5% to about 140% of the growth rate of the recombinant yeast cell in arabinose growth media. The growth rate of the recombinant yeast cell in an arabinose-glucose growth media may be about 5% to about 130% of

The arabinose growth media may include arabinose at a concentration of about 125 g/L. The arabinose growth media may include arabinose at a concentration of about 150 g/L. The arabinose growth media may include arabinose at a concentration of about 175 g/L. The arabinose growth media may include arabinose at a concentration of about 200 g/L. The arabinose growth media may include arabinose at a concentration of about 225 g/L. The arabinose growth media may include arabinose at a concentration of about 250 g/L. The arabinose growth media may include arabinose at a concentration of about 275 g/L. The arabinose growth media may include arabinose at a concentration of about 300 g/L.

In embodiments, the arabinose-glucose growth media includes arabinose at a concentration described herein for an arabinose growth media. In embodiments, the arabinose-glucose growth media includes glucose at a concentration of about 0.05 g/L to about 100 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 90 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 80 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 70 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 60 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 50 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 40 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 30 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 25 g/L.

The arabinose-glucose growth media may include glucose at a concentration of about 25 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 30 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 40 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 50 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 60 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 70 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 80 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 90 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 100 g/L.

In embodiments, the arabinose-glucose growth media includes glucose at a concentration of about 0.05 g/L to about 20 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 15 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 10 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 5 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 4 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 3 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 2 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 1 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 0.5 g/L. The arabinose-glucose growth media may include glucose at a

concentration of about 0.05 g/L to about 0.1 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 0.5 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 1 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 2 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 3 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 4 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 5 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 10 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 15 g/L. The arabinose-glucose growth media may include glucose at a concentration of about 20 g/L.

3. Recombinant Yeast Cell Including a Recombinant Galactose Transporter Protein

In another aspect is a recombinant yeast cell that includes a recombinant galactose transporter protein as described herein, including embodiments thereof. In embodiments, the growth rate of the recombinant yeast cell including a recombinant transporter protein as described herein can be measured. The growth rate may be determined in galactose growth media (i.e. in the absence of glucose). The growth rate may be determined in galactose-glucose growth media. In embodiments, the growth rates (i.e. growth rate in the absence and presence of glucose) are compared to determine the differential growth rate of the recombinant yeast cells. If the growth rate of recombinant yeast cells grown in galactose-glucose growth media is less than the growth rate of recombinant yeast cells grown in galactose growth media, the differential growth rate may indicate glucose inhibits the activity of the recombinant galactose transporter protein. As described herein, inclusion of a glucose mitigating mutation decreases, minimizes, or may eliminate glucose inhibition experienced by the recombinant galactose transporter protein.

In embodiments, the growth rate of the recombinant yeast cell in a galactose-glucose growth media is about 5% to about 150% of the growth rate of the recombinant yeast cell in galactose growth media. The growth rate of the recombinant yeast cell in a galactose-glucose growth media may be about 5% to about 140% of the growth rate of the recombinant yeast cell in galactose growth media. The growth rate of the recombinant yeast cell in a galactose-glucose growth media may be about 5% to about 130% of the growth rate of the recombinant yeast cell in galactose growth media. The growth rate of the recombinant yeast cell in a galactose-glucose growth media may be about 5% to about 120% of the growth rate of the recombinant yeast cell in galactose growth media. The growth rate of the recombinant yeast cell in a galactose-glucose growth media may be about 5% to about 110% of the growth rate of the recombinant yeast cell in galactose growth media.

The growth rate of the recombinant yeast cell in a galactose-glucose growth media may be about 5% to about 100% of the growth rate of the recombinant yeast cell in galactose growth media. The growth rate of the recombinant yeast cell in a galactose-glucose growth media may be about 5% to about 90% of the growth rate of the recombinant yeast cell in galactose growth media. The growth rate of the recombinant yeast cell in a galactose-glucose growth media may be about 5% to about 80% of the growth rate of the recombinant yeast cell in galactose growth media. The growth rate of the recombinant yeast cell in a galactose-glucose growth media may be about 5% to about 70% of the

growth media may include glucose at a concentration of about 0.05 g/L to about 25 g/L.

The galactose-glucose growth media may include glucose at a concentration of about 25 g/L. The galactose-glucose growth media may include glucose at a concentration of about 30 g/L. The galactose-glucose growth media may include glucose at a concentration of about 40 g/L. The galactose-glucose growth media may include glucose at a concentration of about 50 g/L. The galactose-glucose growth media may include glucose at a concentration of about 60 g/L. The galactose-glucose growth media may include glucose at a concentration of about 70 g/L. The galactose-glucose growth media may include glucose at a concentration of about 80 g/L. The galactose-glucose growth media may include glucose at a concentration of about 90 g/L. The galactose-glucose growth media may include glucose at a concentration of about 100 g/L.

In embodiments, the galactose-glucose growth media includes glucose at a concentration of about 0.05 g/L to about 20 g/L. The galactose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 15 g/L. The galactose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 10 g/L. The galactose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 5 g/L. The galactose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 4 g/L. The galactose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 3 g/L. The galactose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 2 g/L. The galactose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 1 g/L. The galactose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 0.5 g/L. The galactose-glucose growth media may include glucose at a concentration of about 0.05 g/L to about 0.1 g/L. The galactose-glucose growth media may include glucose at a concentration of about 0.05 g/L. The galactose-glucose growth media may include glucose at a concentration of about 0.1 g/L. The galactose-glucose growth media may include glucose at a concentration of about 0.5 g/L. The galactose-glucose growth media may include glucose at a concentration of about 1 g/L. The galactose-glucose growth media may include glucose at a concentration of about 2 g/L. The galactose-glucose growth media may include glucose at a concentration of about 3 g/L. The galactose-glucose growth media may include glucose at a concentration of about 4 g/L. The galactose-glucose growth media may include glucose at a concentration of about 5 g/L. The galactose-glucose growth media may include glucose at a concentration of about 10 g/L. The galactose-glucose growth media may include glucose at a concentration of about 15 g/L. The galactose-glucose growth media may include glucose at a concentration of about 20 g/L.

IV. Methods

Also provided herein are methods of transporting xylose into a recombinant yeast cell. In one aspect, the method includes contacting a recombinant yeast cell with a xylose compound described herein, where the recombinant yeast cell includes a recombinant xylose transporter protein as described herein, including embodiments thereof. The recombinant xylose transporter protein transports the xylose compound into the recombinant yeast cell. In embodiments, the only sugar (i.e. carbon source) present is a xylose

compound. The recombinant xylose transporter protein is as described herein, including embodiments thereof. By extension, the xylose transporter motif sequence and the glucose mitigation mutation are as described herein, including embodiments thereof.

In another aspect, the method includes contacting a recombinant yeast cell with a xylose compound, where the xylose compound is the only sugar (i.e. carbon source) in the media, and where the recombinant yeast cell includes a recombinant xylose transporter protein as described herein, including embodiments thereof.

1. Transporting Xylose into a Recombinant Yeast Cell

The xylose compound may be derived from lignocellulosic biomass, hemicellulose, or xylan. Thus, in embodiments, the xylose compound is not the only sugar (i.e. carbon source) present. The xylose compound may be derived from lignocellulosic biomass. The xylose compound may be derived from hemicellulose. The xylose compound may be derived from xylan. In embodiments, the recombinant yeast cell metabolizes the xylose compound. The xylose compound may be present at a concentration as described hereinabove for the "xylose growth media". In embodiments, the recombinant yeast cell converts the xylose compound to a biofuel as described herein (e.g. ethanol) or to a biochemical as described herein. The recombinant yeast cell may convert the xylose compound to a biofuel as described herein (e.g. ethanol). The recombinant yeast cell may convert the xylose compound to a biochemical as described herein. In embodiments, the only sugar (i.e. carbon source) available is the xylose compound.

In embodiments, the recombinant xylose transporter transports a xylose compound into a recombinant yeast cell in a xylose-glucose growth media at a rate of at least 1 nmol min⁻¹ gDCW⁻¹ to 15 nmol min⁻¹ gDCW⁻¹. The term "gDCW" provided herein is well known in the art and refers to gram dry cell weight. The recombinant xylose transporter may transport a xylose compound into a recombinant yeast cell in a xylose-glucose growth media at a rate of at least 2 nmol min⁻¹ gDCW⁻¹ to 15 nmol min⁻¹ gDCW⁻¹. The recombinant xylose transporter may transport a xylose compound into a recombinant yeast cell in a xylose-glucose growth media at a rate of at least 3 nmol min⁻¹ gDCW⁻¹ to 15 nmol min⁻¹ gDCW⁻¹. The recombinant xylose transporter may transport a xylose compound into a recombinant yeast cell in a xylose-glucose growth media at a rate of at least 4 nmol min⁻¹ gDCW⁻¹ to 15 nmol min⁻¹ gDCW⁻¹. The recombinant xylose transporter may transport a xylose compound into a recombinant yeast cell in a xylose-glucose growth media at a rate of at least 5 nmol min⁻¹ gDCW⁻¹ to 15 nmol min⁻¹ gDCW⁻¹. The recombinant xylose transporter may transport a xylose compound into a recombinant yeast cell in a xylose-glucose growth media at a rate of at least 6 nmol min⁻¹ gDCW⁻¹ to 15 nmol min⁻¹ gDCW⁻¹. The recombinant xylose transporter may transport a xylose compound into a recombinant yeast cell in a xylose-glucose growth media at a rate of at least 7 nmol min⁻¹ gDCW⁻¹ to 15 nmol min⁻¹ gDCW⁻¹. The recombinant xylose transporter may transport a xylose compound into a recombinant yeast cell in a xylose-glucose growth media at a rate of at least 8 nmol min⁻¹ gDCW⁻¹ to 15 nmol min⁻¹ gDCW⁻¹. The recombinant xylose transporter may transport a xylose compound into a recombinant yeast cell in a xylose-glucose growth media at a rate of at least 9 nmol min⁻¹ gDCW⁻¹ to 15 nmol min⁻¹ gDCW⁻¹. The recombinant xylose transporter may transport a xylose compound into a recombinant yeast cell in a xylose-glucose growth media at a rate of at least 10 nmol min⁻¹ gDCW⁻¹ to 15 nmol min⁻¹ gDCW⁻¹.

port a galactose compound into a recombinant yeast cell in a galactose-glucose growth media at a rate of at least $130 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$ to $150 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$. The recombinant galactose transporter may transport a galactose compound into a recombinant yeast cell in a galactose-glucose growth media at a rate of at least $140 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$ to $150 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$.

The recombinant galactose transporter may transport a galactose compound into a recombinant yeast cell in a galactose-glucose growth media at a rate of at least $60 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$. The recombinant galactose transporter may transport a galactose compound into a recombinant yeast cell in a galactose-glucose growth media at a rate of at least $70 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$. The recombinant galactose transporter may transport a galactose compound into a recombinant yeast cell in a galactose-glucose growth media at a rate of at least $80 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$. The recombinant galactose transporter may transport a galactose compound into a recombinant yeast cell in a galactose-glucose growth media at a rate of at least $90 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$. The recombinant galactose transporter may transport a galactose compound into a recombinant yeast cell in a galactose-glucose growth media at a rate of at least $100 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$. The recombinant galactose transporter may transport a galactose compound into a recombinant yeast cell in a galactose-glucose growth media at a rate of at least $110 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$. The recombinant galactose transporter may transport a galactose compound into a recombinant yeast cell in a galactose-glucose growth media at a rate of at least $120 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$. The recombinant galactose transporter may transport a galactose compound into a recombinant yeast cell in a galactose-glucose growth media at a rate of at least $130 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$. The recombinant galactose transporter may transport a galactose compound into a recombinant yeast cell in a galactose-glucose growth media at a rate of at least $140 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$. The recombinant galactose transporter may transport a galactose compound into a recombinant yeast cell in a galactose-glucose growth media at a rate of at least $150 \text{ nmol min}^{-1} \text{ gDCW}^{-1}$.

V. Examples

In previous research, was developed a xylose specific transporter hereafter termed "CiGXS1-FIM" ("FIM"), based on a hexose transporter from *C. intermedia*, GXS1. The FIM mutation imparted specificity in transporting xylose over glucose. (11) The presence of glucose, however, inhibited the performance of FIM in transporting xylose. Herein directed evolution was conducted to reduce the observed glucose inhibition.

A library of randomly mutated FIM was generated by error-prone PCR with a library size of over 1×10^5 mutants (as measured by independent *E. coli* colonies post-transformation). The mutant FIM was then transformed into *S. cerevisiae* ETKXG strain, a triple hexokinase knockout strain which is not able to grow on glucose, and screened on the xylose dependent growth based advantage on the drop-out plates with 20 g/L of xylose and 2.5 g/L of glucose. The 140 selected mutants from the plates were then tested for the growth on the medium with 20 g/L of xylose and 2.5 g/L glucose using Bioscreen C and the top 6 mutants were selected for further confirmation. The growth rates of the selected mutants on xylose in the presence of various concentration of glucose were then confirmed using Bioscreen C (FIG. 1).

The mutant 105, which has 6 mutations: K155E, N225D, S354T, A361T, L407M, N446S (FIG. 2), showed reduced glucose inhibition. Indeed, the mutant 105 showed significantly higher growth rates in the all tested conditions compared to FIM and wild-type transporters. Mutant 105 shows nearly a 30-fold increase in the growth rate on a xylose medium with the presence of glucose. The growth rate of the mutant 105 in the xylose 20 g/L+glucose 2.5 g/L medium was slightly higher than the growth rate of the wild-type transporter in the xylose only medium. Though the mutant 105 showed the highest reduction in glucose inhibition, the growth rate on xylose only was slightly reduced compared to FIM and wild-type. See FIG. 1

The mutant 78, which has a single mutation, N326S, showed reduced glucose inhibition without decrease in xylose transport capability. In contrast to mutant 105, mutant 78 showed no reduction in the xylose transport performance. This suggests the mutant 78 represents a promising candidate for further round of directed evolution to develop xylose transporters with reduced glucose inhibition.

REFERENCES

1. Reijenga K A, et al. (2001) Control of glycolytic dynamics by hexose transport in *Saccharomyces cerevisiae*. *Biophysical Journal* 80(2):626-634.
2. Gardonyi M, Jeppsson M, Liden G, Gorwa-Grausland M F, & Hahn-Hagerdal B (2003) Control of xylose consumption by xylose transport in recombinant *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering* 82(7):818-824.
3. Elbing K, et al. (2004) Role of hexose transport in control of glycolytic flux in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* 70(9):5323-5330.
4. Wahlbom C F, Otero R R C, van Zyl W H, Hahn-Hagerdal B, & Jonsson L J (2003) Molecular analysis of a *Saccharomyces cerevisiae* mutant with improved ability to utilize xylose shows enhanced expression of proteins involved in transport, initial xylose metabolism, and the pentose phosphate pathway. *Applied and Environmental Microbiology* 69(2):740-746.
5. Bengtsson O, et al. (2008) Identification of common traits in improved xylose-growing *Saccharomyces cerevisiae* for inverse metabolic engineering. *Yeast* 25(11):835-847.
6. Jeffries T W & Jin Y S (2004) Metabolic engineering for improved fermentation of pentoses by yeasts. *Applied Microbiology and Biotechnology* 63(5):495-509.
7. Hahn-Hagerdal B, Karhumaa K, Fonseca C, Spencer-Martins I, & Gorwa-Grauslund M F (2007) Towards industrial pentose-fermenting yeast strains. *Applied Microbiology and Biotechnology* 74(5):937-953.
8. Martin C H, Nielsen D R, Solomon K V, & Prather K L J (2009) Synthetic metabolism: engineering biology at the protein and pathway scales. *Chemistry & Biology* 16(3):277-286.
9. Tyo K E J, Kocharin K, & Nielsen J (2010) Toward design-based engineering of industrial microbes. *Current Opinion in Microbiology* 13(3):255-262.
10. Curran K A & Alper H S (2012) Expanding the chemical palate of cells by combining systems biology and metabolic engineering. *Metabolic Engineering* 14(4):289-297.
11. Eric Young, Alice Tong, Hang Bui, Caitlin Spofford, and Hal Alper, 2014. *Rewiring yeast sugar transporter preference through modifying a conserved protein motif*. *PNAS* 111(1), 131-136

89

VI. P Embodiments

Embodiment P 1

A recombinant xylose transporter protein comprising a xylose transporter motif sequence and a glucose mitigation mutation.

Embodiment P 2

The recombinant xylose transporter protein of embodiment P 1, wherein said xylose transporter motif sequence corresponds to amino acid residue positions 36, 37, 38, 39, 40, and 41 of *Candida intermedia* GXS1 protein, and wherein said xylose transporter motif sequence is -G-G/F-X¹-X²-X³-G-; wherein, X¹ is D, C, G, H, I, L, or F; X² is A, D, C, E, G, H, or I; and X³ is N, C, Q, F, G, L, M, S, T, or P.

Embodiment P 3

The recombinant xylose transporter protein of embodiment P 1 or embodiment P 2, wherein said xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ ID NO:107), -G-F-F-I-M-G- (SEQ ID NO:108), -G-G-F-I-S-G- (SEQ ID NO:109), -G-F-F-I-S-G- (SEQ ID NO:110), -G-G-F-I-T-G- (SEQ ID NO:111), -G-F-F-I-T-G- (SEQ ID NO:112), -G-G-F-L-M-G- (SEQ ID NO:113), -G-F-F-L-M-G- (SEQ ID NO:114), -G-G-F-L-S-G- (SEQ ID NO:115), -G-F-F-L-S-G- (SEQ ID NO:116), -G-G-F-L-T-G- (SEQ ID NO:117), -G-F-F-L-T-G- (SEQ ID NO:118), -G-G-F-H-M-G- (SEQ ID NO:119), -G-F-F-H-M-G- (SEQ ID NO:120), -G-G-F-H-S-G- (SEQ ID NO:121), -G-F-F-H-S-G- (SEQ ID NO:122), -G-G-F-H-T-G- (SEQ ID NO:123) or -G-F-F-H-T-G- (SEQ ID NO:124).

Embodiment P 4

The recombinant xylose transporter protein of any one of embodiments P 1 to 3, wherein said xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ ID NO:107), -G-F-F-I-M-G- (SEQ ID NO:108), -G-G-F-I-S-G- (SEQ ID NO:109), or -G-F-F-I-S-G- (SEQ ID NO:110).

Embodiment P 5

The recombinant xylose transporter protein of any one of embodiments P 1 to 4, wherein said xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ ID NO:107).

Embodiment P 6

The recombinant xylose transporter protein of any one of embodiments P 1 to 5, wherein said glucose mitigation mutation is within a protein domain corresponding to transmembrane 9 of *Candida intermedia* GXS1 protein.

Embodiment P 7

The recombinant xylose transporter protein of any one of embodiments P 1 to 6, wherein said glucose mitigation mutation is at a position corresponding to K155, N225, S354, A361, L407, or N446 of *Candida intermedia* GXS1 protein.

Embodiment P 8

The recombinant xylose transporter protein of any one of embodiments P 1 to 5, wherein said glucose mitigation

90

mutation is within a protein domain corresponding to transmembrane 8 of *Candida intermedia* GXS1 protein.

Embodiment P 9

The recombinant xylose transporter protein of any one of embodiments P 1 to 5, or embodiment P 8, wherein said glucose mitigation mutation is at a position corresponding N326 of *Candida intermedia* GXS1 protein.

Embodiment P 10

The recombinant xylose transporter protein of embodiment P 9, wherein said glucose mitigation mutation is a N326S mutation.

Embodiment P 11

A recombinant yeast cell comprising a recombinant xylose transporter protein of any one of embodiments P 1 to 10.

Embodiment P 12

The recombinant yeast cell of embodiment P 11, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 10% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment P 13

The recombinant yeast cell of embodiment P 11, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 20% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment P 14

The recombinant yeast cell of embodiment P 11, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 30% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment P 15

The recombinant yeast cell of embodiment P 11, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 40% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment P 16

The recombinant yeast cell of embodiment P 11, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 50% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment P 17

The recombinant yeast cell of embodiment P 11, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 60% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment P 18

The recombinant yeast cell of embodiment P 11, wherein the growth rate of said recombinant yeast cell in a xylose-

91

glucose growth media is at least about 70% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment P 19

The recombinant yeast cell of embodiment P 11, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 80% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment P 20

The recombinant yeast cell of embodiment P 11, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 90% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment P 21

The recombinant yeast cell of embodiment P 11, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 100% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment P 22

The recombinant yeast cell of embodiment P 11, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 110% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment P 23

The recombinant yeast cell of embodiment P 11, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 120% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment P 24

The recombinant yeast cell of any one of embodiments P 11 to 23, wherein said xylose-glucose growth media comprises about 0.05 g/L to about 20 g/L glucose.

Embodiment P 25

The recombinant yeast cell of any one of embodiments P 11 to 24, wherein said xylose-glucose growth media comprises about 2.5 g/L glucose.

Embodiment P 26

The recombinant yeast cell of any one of embodiments P 11 to 25, wherein said xylose-glucose growth media comprises about 5 g/L glucose.

Embodiment P 27

The recombinant yeast cell of any one of embodiments P 11 to 26, wherein said xylose-glucose growth media comprises about 10 g/L glucose.

Embodiment P 28

The recombinant yeast cell of any one of embodiments P 11 to 27, wherein said xylose-glucose growth media comprises about 20 g/L glucose.

92

Embodiment P 29

The recombinant yeast cell of any one of embodiments P 11 to 28, wherein said xylose-glucose growth media comprises about 0.05 g/L to about 300 g/L xylose.

Embodiment P 30

The recombinant yeast cell of any one of embodiments P 11 to 29, wherein said xylose growth media comprises about 0.05 g/L to about 300 g/L xylose.

Embodiment P 31

The recombinant yeast cell of any one of embodiments P 11 to 30, wherein said xylose growth media comprises about 20 g/L xylose.

Embodiment P 32

A method of transporting xylose into a recombinant yeast cell, said method comprising: i) contacting a recombinant yeast cell with a xylose compound, wherein said recombinant yeast cell comprises a recombinant xylose transporter protein, said recombinant xylose transporter protein comprising a xylose transporter motif sequence and a glucose mitigation mutation; and ii) allowing said recombinant xylose transporter protein to transport said xylose compound into said recombinant yeast cell.

Embodiment P 33

The method of embodiment P 32, wherein said xylose transporter motif sequence corresponds to amino acid residue positions 36, 37, 38, 39, 40, and 41 of *Candida intermedia* GXS1 protein, and wherein said xylose transporter motif sequence is -G-G/F-X¹-X²-X³-G-; wherein, X¹ is D, C, G, H, I, L, or F; X² is A, D, C, E, G, H, or I; and X³ is N, C, Q, F, G, L, M, S, T, or P.

Embodiment P 34

The method of embodiment P 32 or embodiment P 33, wherein said xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ ID NO:107), -G-F-F-I-M-G- (SEQ ID NO:108), -G-G-F-I-S-G- (SEQ ID NO:109), -G-F-F-I-S-G- (SEQ ID NO:110), -G-G-F-I-T-G- (SEQ ID NO:111), -G-F-F-I-T-G- (SEQ ID NO:112), -G-G-F-L-M-G- (SEQ ID NO:113), -G-F-F-L-M-G- (SEQ ID NO:114), -G-G-F-L-S-G- (SEQ ID NO:115), -G-F-F-L-S-G- (SEQ ID NO:116), -G-G-F-L-T-G- (SEQ ID NO:117), -G-F-F-L-T-G- (SEQ ID NO:118), -G-G-F-H-M-G- (SEQ ID NO:119), -G-F-F-H-M-G- (SEQ ID NO:120), -G-G-F-H-S-G- (SEQ ID NO:121), -G-F-F-H-S-G- (SEQ ID NO:122), -G-G-F-H-T-G- (SEQ ID NO:123) or -G-F-F-H-T-G- (SEQ ID NO:124).

Embodiment P 35

The method of any one of embodiments P 32 to 34, wherein said xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ ID NO:107), -G-F-F-I-M-G- (SEQ ID NO:108), -G-G-F-I-S-G- (SEQ ID NO:109), or -G-F-F-I-S-G- (SEQ ID NO:110).

93

Embodiment P 36

The method of any one of embodiments P 32 to 35, wherein said xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ ID NO:107).

Embodiment P 37

The method of any one of embodiments P 32 to 36, wherein said glucose mitigation mutation is within a protein domain corresponding to transmembrane 9 of *Candida intermedia* GXS1 protein.

Embodiment P 38

The method of any one of embodiments P 32 to 37, wherein said glucose mitigation mutation is at a position corresponding to K155, N225, S354, A361, L407, or N446 of *Candida intermedia* GXS1 protein.

Embodiment P 39

The method of any one of embodiments P 32 to 36, wherein said glucose mitigation mutation is within a protein domain corresponding to transmembrane 8 of *Candida intermedia* GXS1 protein.

Embodiment P 40

The method of any one of embodiments P 32 to 36, or embodiment P 39, wherein said glucose mitigation mutation is at a position corresponding N326 of *Candida intermedia* GXS1 protein.

Embodiment P 41

The method of embodiment P 40, wherein said glucose mitigation mutation is a N326S mutation.

Embodiment P 42

The method of any one of embodiments P 32 to 41, wherein said recombinant yeast cell metabolizes said xylose compound.

Embodiment P 43

The method of any one of embodiments P 32 to 42, wherein said recombinant yeast cell converts said xylose compound to a biofuel.

Embodiment P 44

The method of any one of embodiments P 32 to 43, wherein said xylose compound forms part of lignocellulosic biomass, hemicellulose, or xylan.

Embodiment P 45

The method of any one of embodiments P 32 to 44, wherein said recombinant xylose transporter transports said xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 10% of the rate said recombinant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

Embodiment P 46

The method of any one of embodiments P 32 to 44, wherein said recombinant xylose transporter transports said

94

xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 20% of the rate said recombinant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

Embodiment P 47

The method of any one of embodiments P 32 to 44, wherein said recombinant xylose transporter transports said xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 30% of the rate said recombinant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

Embodiment P 48

The method of any one of embodiments P 32 to 44, wherein said recombinant xylose transporter transports said xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 40% of the rate said recombinant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

Embodiment P 49

The method of any one of embodiments P 32 to 44, wherein said recombinant xylose transporter transports said xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 50% of the rate said recombinant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

Embodiment P 50

The method of any one of embodiments P 32 to 44, wherein said recombinant xylose transporter transports said xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 60% of the rate said recombinant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

Embodiment P 51

The method of any one of embodiments P 32 to 44, wherein said recombinant xylose transporter transports said xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 70% of the rate said recombinant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

Embodiment P 52

The method of any one of embodiments P 32 to 44, wherein said recombinant xylose transporter transports said xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 80% of the rate said recombinant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

Embodiment P 53

The method of any one of embodiments P 32 to 44, wherein said recombinant xylose transporter transports said xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 90% of the rate said recom-

95

binant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

Embodiment P 54

The method of any one of embodiments P 32 to 44, wherein said recombinant xylose transporter transports said xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 100% of the rate said recombinant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

Embodiment P 55

The method of any one of embodiments P 32 to 44, wherein said recombinant xylose transporter transports said xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 110% of the rate said recombinant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

Embodiment P 56

The method of any one of embodiments P 32 to 44, wherein said recombinant xylose transporter transports said xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 120% of the rate said recombinant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

Embodiment P 57

The method of any one of embodiments P 45 to 56, wherein said xylose-glucose growth media comprises about 0.05 g/L to about 20 g/L glucose.

Embodiment P 58

The method of any one of embodiments P 45 to 57, wherein said xylose-glucose growth media comprises about 2.5 g/L glucose.

Embodiment P 59

The method of any one of embodiments P 45 to 58, wherein said xylose-glucose growth media comprises about 5 g/L glucose.

Embodiment P 60

The method of any one of embodiments P 45 to 59, wherein said xylose-glucose growth media comprises about 10 g/L glucose.

Embodiment P 61

The method of any one of embodiments P 45 to 60, wherein said xylose-glucose growth media comprises about 20 g/L glucose.

Embodiment P 62

The method of any one of embodiments P 45 to 61, wherein said xylose-glucose growth media comprises about 0.05 g/L to about 300 g/L xylose.

96

Embodiment P 63

The method of any one of embodiments P 45 to 62, wherein said xylose growth media comprises about 0.05 g/L to about 300 g/L xylose.

Embodiment P 64

The method of any one of embodiments P 45 to 63, wherein said xylose growth media comprises about 20 g/L xylose.

Embodiment P 65

The method of any one of embodiments P 32 to 64, wherein said recombinant xylose transporter protein transports said xylose compound into said recombinant yeast cell in a xylose-glucose growth media growth media at a rate of at least 5 nmol min⁻¹ gDCW⁻¹.

Embodiment P 66

A nucleic acid encoding the recombinant xylose transporter protein of one of embodiments P 1 to 10.

VII. Further Embodiments

Embodiment 1

A recombinant xylose transporter protein comprising a xylose transporter motif sequence and a glucose mitigation mutation.

Embodiment 2

The recombinant xylose transporter protein of embodiment 1, wherein said xylose transporter motif sequence corresponds to amino acid residue positions 36, 37, 38, 39, 40, and 41 of *Candida intermedia* GXS1 protein.

Embodiment 3

The recombinant xylose transporter protein of embodiment 1 or embodiment 2, wherein said xylose transporter motif sequence is -G-G/F-X¹-X²-X³-G-; wherein, X¹ is D, C, G, H, I, L, or F; X² is A, D, C, E, G, H, or I; and X³ is N, C, Q, F, G, L, M, S, T, or P.

Embodiment 4

The recombinant xylose transporter protein of one of embodiments 1-3, wherein said xylose transporter motif sequence is G-G-F-I-M-G- (SEQ ID NO:107), -G-F-F-I-M-G- (SEQ ID NO:108), -G-G-F-I-S-G- (SEQ ID NO:109), -G-F-F-I-S-G- (SEQ ID NO:110), -G-G-F-I-T-G- (SEQ ID NO:111), -G-F-F-I-T-G- (SEQ ID NO:112), -G-G-F-L-M-G- (SEQ ID NO:113), -G-F-F-L-M-G- (SEQ ID NO:114), -G-G-F-L-S-G- (SEQ ID NO:115), -G-F-F-L-S-G- (SEQ ID NO:116), -G-G-F-L-T-G- (SEQ ID NO:117), -G-F-F-L-T-G- (SEQ ID NO:118), -G-G-F-H-M-G- (SEQ ID NO:119), -G-F-F-H-M-G- (SEQ ID NO:120), -G-G-F-H-S-G- (SEQ ID NO:121), -G-F-F-H-S-G- (SEQ ID NO:122), -G-G-F-H-T-G- (SEQ ID NO:123) or -G-F-F-H-T-G- (SEQ ID NO:124).

Embodiment 5

The recombinant xylose transporter protein of any one of embodiments 2 to 4, wherein said xylose transporter motif

97

sequence is G-G-F-I-M-G- (SEQ ID NO:107), -G-F-F-I-M-G- (SEQ ID NO:108), -G-G-F-I-S-G- (SEQ ID NO:109), or -G-F-F-I-S-G- (SEQ ID NO:110).

Embodiment 6

The recombinant xylose transporter protein of any one of embodiments 1 to 5, wherein said xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ ID NO:107).

Embodiment 7

The recombinant xylose transporter protein of any one of embodiments 1 to 6, wherein said glucose mitigation mutation is within a protein domain corresponding to transmembrane 9 of *Candida intermedia* GXS1 protein.

Embodiment 8

The recombinant xylose transporter protein of any one of embodiments 1 to 7, wherein said glucose mitigation mutation is at a position corresponding to K155, N225, S354, A361, L407, or N446 of *Candida intermedia* GXS1 protein.

Embodiment 9

The recombinant xylose transporter protein of any one of embodiments 1 to 6, wherein said glucose mitigation mutation is within a protein domain corresponding to transmembrane 8 of *Candida intermedia* GXS1 protein.

Embodiment 10

The recombinant xylose transporter protein of any one of embodiments 1 to 6, or embodiment 9, wherein said glucose mitigation mutation is at a position corresponding to N326 of *Candida intermedia* GXS1 protein.

Embodiment 11

The recombinant xylose transporter protein of embodiment 10, wherein said glucose mitigation mutation is a N326H mutation.

Embodiment 12

The recombinant xylose transporter protein of embodiment 10, wherein said glucose mitigation mutation is a N326S mutation.

Embodiment 13

The recombinant xylose transporter protein of any one of embodiments 1 to 6, wherein said glucose mitigation mutation is within a protein domain corresponding to transmembrane 5 of *Candida intermedia* GXS1 protein.

Embodiment 14

The recombinant xylose transporter protein of any one of embodiments 1 to 6, or embodiment 13, wherein said glucose mitigation mutation is within a protein domain corresponding to residue 160-179 of *Candida intermedia* GXS1 protein.

Embodiment 15

The recombinant xylose transporter protein of any one of embodiments 1 to 6, 13 or 14, wherein said glucose miti-

98

gation mutation is at a position corresponding to T170 or I171 of *Candida intermedia* GXS1 protein.

Embodiment 16

The recombinant xylose transporter protein of embodiment 15, wherein said glucose mitigation mutation is a T170N mutation.

Embodiment 17

The recombinant xylose transporter protein of embodiment 15, wherein said glucose mitigation mutation is a I171F mutation.

Embodiment 18

The recombinant xylose transporter protein of one of embodiments 1-17 further comprising an amino acid deletion.

Embodiment 19

The recombinant xylose transporter protein of embodiment 18, wherein said deletion is within a protein domain corresponding to residue 497-522 of *Candida intermedia* GXS1 protein.

Embodiment 20

The recombinant xylose transporter protein of embodiment 18 or embodiment 19, wherein said deletion is at least 10 amino acids in length.

Embodiment 21

A recombinant yeast cell comprising a recombinant xylose transporter protein of any one of embodiments 1 to 20

Embodiment 22

The recombinant yeast cell of embodiment 21, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 10% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment 23

The recombinant yeast cell of embodiment 21, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 20% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment 24

The recombinant yeast cell of embodiment 21, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 30% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment 25

The recombinant yeast cell of embodiment 21, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 40% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment 26

The recombinant yeast cell of embodiment 21, wherein the growth rate of said recombinant yeast cell in a xylose-

99

glucose growth media is at least about 50% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment 27

The recombinant yeast cell of embodiment 21, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 60% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment 28

The recombinant yeast cell of embodiment 21, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 70% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment 29

The recombinant yeast cell of embodiment 21, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 80% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment 30

The recombinant yeast cell of embodiment 21, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 90% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment 31

The recombinant yeast cell of embodiment 21, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 100% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment 32

The recombinant yeast cell of embodiment 21, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 110% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment 33

The recombinant yeast cell of embodiment 21, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 120% of the growth rate of said recombinant yeast cell in a xylose growth media.

Embodiment 34

The recombinant yeast cell of any one of embodiments 21 to 33, wherein said xylose-glucose growth media comprises about 0.05 g/L to about 20 g/L glucose.

Embodiment 35

The recombinant yeast cell of any one of embodiments 21 to 34, wherein said xylose-glucose growth media comprises about 2.5 g/L glucose.

100

Embodiment 36

The recombinant yeast cell of any one of embodiments 21 to 35, wherein said xylose-glucose growth media comprises about 5 g/L glucose.

Embodiment 37

The recombinant yeast cell of any one of embodiments 21 to 36, wherein said xylose-glucose growth media comprises about 10 g/L glucose.

Embodiment 38

The recombinant yeast cell of any one of embodiments 21 to 37, wherein said xylose-glucose growth media comprises about 20 g/L glucose.

Embodiment 39

The recombinant yeast cell of any one of embodiments 21 to 38, wherein said xylose-glucose growth media comprises about 0.05 g/L to about 300 g/L xylose.

Embodiment 40

The recombinant yeast cell of any one of embodiments 21 to 39, wherein said xylose growth media comprises about 0.05 g/L to about 300 g/L xylose.

Embodiment 41

The recombinant yeast cell of any one of embodiments 21 to 40, wherein said xylose growth media comprises about 20 g/L xylose.

Embodiment 42

A method of transporting xylose into a recombinant yeast cell, said method comprising: i) contacting a recombinant yeast cell with a xylose compound, wherein said recombinant yeast cell comprises a recombinant xylose transporter protein, said recombinant xylose transporter protein comprising a xylose transporter motif sequence and a glucose mitigation mutation; and ii) allowing said recombinant xylose transporter protein to transport said xylose compound into said recombinant yeast cell.

Embodiment 43

The method of embodiment 42, wherein said xylose transporter motif sequence corresponds to amino acid residue positions 36, 37, 38, 39, 40, and 41 of *Candida intermedia* GXS1 protein.

Embodiment 44

The method of embodiment 42 or embodiment 43, wherein said xylose transporter motif sequence is -G-G/F-X¹-X²-X³-G-; wherein, X¹ is D, C, G, H, I, L, or F; X² is A, D, C, E, G, H, or I; and X³ is N, C, Q, F, G, L, M, S, T, or P.

Embodiment 45

The method of one of embodiments 42-44, wherein said xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ

101

ID NO:107), -G-F-F-I-M-G- (SEQ ID NO:108), -G-G-F-I-S-G- (SEQ ID NO:109), -G-F-F-I-S-G- (SEQ ID NO:110), -G-G-F-I-T-G- (SEQ ID NO:111), -G-F-F-I-T-G- (SEQ ID NO:112), -G-G-F-L-M-G- (SEQ ID NO:113), -G-F-F-L-M-G- (SEQ ID NO:114), -G-G-F-L-S-G- (SEQ ID NO:115), -G-F-F-L-S-G- (SEQ ID NO:116), -G-G-F-L-T-G- (SEQ ID NO:117), -G-F-F-L-T-G- (SEQ ID NO:118), -G-G-F-H-M-G- (SEQ ID NO:119), -G-F-F-H-M-G- (SEQ ID NO:120), -G-G-F-H-S-G- (SEQ ID NO:121), -G-F-F-H-S-G- (SEQ ID NO:122), -G-G-F-H-T-G- (SEQ ID NO:123) or -G-F-F-H-T-G- (SEQ ID NO:124).

Embodiment 46

The method of any one of embodiments 42 to 45, wherein said xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ ID NO:107), -G-F-F-I-M-G- (SEQ ID NO:108), -G-G-F-I-S-G- (SEQ ID NO:109), or -G-F-F-I-S-G- (SEQ ID NO:110).

Embodiment 47

The method of any one of embodiments 42 to 46, wherein said xylose transporter motif sequence is -G-G-F-I-M-G- (SEQ ID NO:107).

Embodiment 48

The method of any one of embodiments 42 to 47, wherein said glucose mitigation mutation is within a protein domain corresponding to transmembrane 9 of *Candida intermedia* GXS1 protein.

Embodiment 49

The method of any one of embodiments 42 to 48, wherein said glucose mitigation mutation is at a position corresponding to K155, N225, S354, A361, L407, or N446 of *Candida intermedia* GXS1 protein.

Embodiment 50

The method of any one of embodiments 42 to 47, wherein said glucose mitigation mutation is within a protein domain corresponding to transmembrane 8 of *Candida intermedia* GXS1 protein.

Embodiment 51

The method of any one of embodiments 42 to 47, or embodiment 50, wherein said glucose mitigation mutation is at a position corresponding N326 of *Candida intermedia* GXS1 protein.

Embodiment 52

The method of embodiment 51, wherein said glucose mitigation mutation is a N326H mutation.

Embodiment 53

The method of embodiment 51, wherein said glucose mitigation mutation is a N326S mutation.

102

Embodiment 54

The method of any one of embodiments 42 to 47, wherein said glucose mitigation mutation is within a protein domain corresponding to transmembrane 5 of *Candida intermedia* GXS1 protein.

Embodiment 55

The method of one of embodiments 42 to 47, or embodiment 54, wherein said glucose mitigation mutation is within a protein domain corresponding to residue 160-179 of *Candida intermedia* GXS1 protein.

Embodiment 56

The method of embodiment 54 or 55, wherein said glucose mitigation mutation is at a position corresponding to T170 or I171 of *Candida intermedia* GXS1 protein.

Embodiment 57

The method of embodiment 56, wherein said glucose mitigation mutation is a T170N mutation.

Embodiment 58

The method of embodiment 56, wherein said glucose mitigation mutation is a I171F mutation.

Embodiment 59

The method of any one of embodiments 42 to 58, further comprising an amino acid deletion.

Embodiment 60

The method of embodiment 59, wherein said deletion is within a protein domain corresponding to residue 497-522 of *Candida intermedia* GXS1 protein.

Embodiment 61

The method of any one of embodiments 42 to 53, wherein said recombinant yeast cell metabolizes said xylose compound.

Embodiment 62

The method of any one of embodiments 42 to 61, wherein said recombinant yeast cell converts said xylose compound to a biofuel.

Embodiment 63

The method of any one of embodiments 42 to 62, wherein said xylose compound forms part of lignocellulosic biomass, hemicellulose, or xylan.

Embodiment 64

The method of any one of embodiments 42 to 63, wherein said recombinant xylose transporter transports said xylose compound into said yeast in a xylose-glucose growth media at a rate at least about 10% of the rate said recombinant xylose transporter transports said xylose compound into said yeast in a xylose growth media.

105

Embodiment 83

The method of any one of embodiments 64 to 82, wherein said xylose growth media comprises about 20 g/L xylose.

Embodiment 84

The method of any one of embodiments 42 to 83, wherein said recombinant xylose transporter protein transports said

106

xylose compound into said recombinant yeast cell in a xylose-glucose growth media growth media at a rate of at least 5 nmol min⁻¹ gDCW⁻¹

Embodiment 85

A nucleic acid encoding the recombinant xylose transporter protein of one of embodiments 1 to 21.

5

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 130

<210> SEQ ID NO 1

<211> LENGTH: 522

<212> TYPE: PRT

<213> ORGANISM: Candida intermedia

<400> SEQUENCE: 1

```

Met Gly Leu Glu Asp Asn Arg Met Val Lys Arg Phe Val Asn Val Gly
1          5          10          15
Glu Lys Lys Ala Gly Ser Thr Ala Met Ala Ile Ile Val Gly Leu Phe
20          25          30
Ala Ala Ser Gly Gly Val Leu Phe Gly Tyr Asp Thr Gly Thr Ile Ser
35          40          45
Gly Val Met Thr Met Asp Tyr Val Leu Ala Arg Tyr Pro Ser Asn Lys
50          55          60
His Ser Phe Thr Ala Asp Glu Ser Ser Leu Ile Val Ser Ile Leu Ser
65          70          75          80
Val Gly Thr Phe Phe Gly Ala Leu Cys Ala Pro Phe Leu Asn Asp Thr
85          90          95
Leu Gly Arg Arg Trp Cys Leu Ile Leu Ser Ala Leu Ile Val Phe Asn
100         105         110
Ile Gly Ala Ile Leu Gln Val Ile Ser Thr Ala Ile Pro Leu Leu Cys
115         120         125
Ala Gly Arg Val Ile Ala Gly Phe Gly Val Gly Leu Ile Ser Ala Thr
130         135         140
Ile Pro Leu Tyr Gln Ser Glu Thr Ala Pro Lys Trp Ile Arg Gly Ala
145         150         155         160
Ile Val Ser Cys Tyr Gln Trp Ala Ile Thr Ile Gly Leu Phe Leu Ala
165         170         175
Ser Cys Val Asn Lys Gly Thr Glu His Met Thr Asn Ser Gly Ser Tyr
180         185         190
Arg Ile Pro Leu Ala Ile Gln Cys Leu Trp Gly Leu Ile Leu Gly Ile
195         200         205
Gly Met Ile Phe Leu Pro Glu Thr Pro Arg Phe Trp Ile Ser Lys Gly
210         215         220
Asn Gln Glu Lys Ala Ala Glu Ser Leu Ala Arg Leu Arg Lys Leu Pro
225         230         235         240
Ile Asp His Pro Asp Ser Leu Glu Glu Leu Arg Asp Ile Thr Ala Ala
245         250         255
Tyr Glu Phe Glu Thr Val Tyr Gly Lys Ser Ser Trp Ser Gln Val Phe
260         265         270
Ser His Lys Asn His Gln Leu Lys Arg Leu Phe Thr Gly Val Ala Ile
275         280         285
Gln Ala Phe Gln Gln Leu Thr Gly Val Asn Phe Ile Phe Tyr Tyr Gly
290         295         300

```

-continued

Thr Thr Phe Phe Lys Arg Ala Gly Val Asn Gly Phe Thr Ile Ser Leu
 305 310 315 320
 Ala Thr Asn Ile Val Asn Val Gly Ser Thr Ile Pro Gly Ile Leu Leu
 325 330 335
 Met Glu Val Leu Gly Arg Arg Asn Met Leu Met Gly Gly Ala Thr Gly
 340 345 350
 Met Ser Leu Ser Gln Leu Ile Val Ala Ile Val Gly Val Ala Thr Ser
 355 360 365
 Glu Asn Asn Lys Ser Ser Gln Ser Val Leu Val Ala Phe Ser Cys Ile
 370 375 380
 Phe Ile Ala Phe Phe Ala Ala Thr Trp Gly Pro Cys Ala Trp Val Val
 385 390 395 400
 Val Gly Glu Leu Phe Pro Leu Arg Thr Arg Ala Lys Ser Val Ser Leu
 405 410 415
 Cys Thr Ala Ser Asn Trp Leu Trp Asn Trp Gly Ile Ala Tyr Ala Thr
 420 425 430
 Pro Tyr Met Val Asp Glu Asp Lys Gly Asn Leu Gly Ser Asn Val Phe
 435 440 445
 Phe Ile Trp Gly Gly Phe Asn Leu Ala Cys Val Phe Phe Ala Trp Tyr
 450 455 460
 Phe Ile Tyr Glu Thr Lys Gly Leu Ser Leu Glu Gln Val Asp Glu Leu
 465 470 475 480
 Tyr Glu His Val Ser Lys Ala Trp Lys Ser Lys Gly Phe Val Pro Ser
 485 490 495
 Lys His Ser Phe Arg Glu Gln Val Asp Gln Gln Met Asp Ser Lys Thr
 500 505 510
 Glu Ala Ile Met Ser Glu Glu Ala Ser Val
 515 520

<210> SEQ ID NO 2
 <211> LENGTH: 522
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 2

Met Gly Leu Glu Asp Asn Arg Met Val Lys Arg Phe Val Asn Val Gly
 1 5 10 15
 Glu Lys Lys Ala Gly Ser Thr Ala Met Ala Ile Ile Val Gly Leu Phe
 20 25 30
 Ala Ala Ser Gly Gly Phe Ile Met Gly Tyr Asp Thr Gly Thr Ile Ser
 35 40 45
 Gly Val Met Thr Met Asp Tyr Val Leu Ala Arg Tyr Pro Ser Asn Lys
 50 55 60
 His Ser Phe Thr Ala Asp Glu Ser Ser Leu Ile Val Ser Ile Leu Ser
 65 70 75 80
 Val Gly Thr Phe Phe Gly Ala Leu Cys Ala Pro Phe Leu Asn Asp Thr
 85 90 95
 Leu Gly Arg Arg Trp Cys Leu Ile Leu Ser Ala Leu Ile Val Phe Asn
 100 105 110
 Ile Gly Ala Ile Leu Gln Val Ile Ser Thr Ala Ile Pro Leu Leu Cys
 115 120 125
 Ala Gly Arg Val Ile Ala Gly Phe Gly Val Gly Leu Ile Ser Ala Thr
 130 135 140

-continued

```

Ile Pro Leu Tyr Gln Ser Glu Thr Ala Pro Lys Trp Ile Arg Gly Ala
145                               150                               155                               160

Ile Val Ser Cys Tyr Gln Trp Ala Ile Thr Ile Gly Leu Phe Leu Ala
                               165                               170                               175

Ser Cys Val Asn Lys Gly Thr Glu His Met Thr Asn Ser Gly Ser Tyr
                               180                               185                               190

Arg Ile Pro Leu Ala Ile Gln Cys Leu Trp Gly Leu Ile Leu Gly Ile
                               195                               200                               205

Gly Met Ile Phe Leu Pro Glu Thr Pro Arg Phe Trp Ile Ser Lys Gly
210                               215                               220

Asn Gln Glu Lys Ala Ala Glu Ser Leu Ala Arg Leu Arg Lys Leu Pro
225                               230                               235                               240

Ile Asp His Pro Asp Ser Leu Glu Glu Leu Arg Asp Ile Thr Ala Ala
                               245                               250                               255

Tyr Glu Phe Glu Thr Val Tyr Gly Lys Ser Ser Trp Ser Gln Val Phe
                               260                               265                               270

Ser His Lys Asn His Gln Leu Lys Arg Leu Phe Thr Gly Val Ala Ile
                               275                               280                               285

Gln Ala Phe Gln Gln Leu Thr Gly Val Asn Phe Ile Phe Tyr Tyr Gly
290                               295                               300

Thr Thr Phe Phe Lys Arg Ala Gly Val Asn Gly Phe Thr Ile Ser Leu
305                               310                               315                               320

Ala Thr Asn Ile Val Asn Val Gly Ser Thr Ile Pro Gly Ile Leu Leu
                               325                               330                               335

Met Glu Val Leu Gly Arg Arg Asn Met Leu Met Gly Gly Ala Thr Gly
                               340                               345                               350

Met Ser Leu Ser Gln Leu Ile Val Ala Ile Val Gly Val Ala Thr Ser
                               355                               360                               365

Glu Asn Asn Lys Ser Ser Gln Ser Val Leu Val Ala Phe Ser Cys Ile
370                               375                               380

Phe Ile Ala Phe Phe Ala Ala Thr Trp Gly Pro Cys Ala Trp Val Val
385                               390                               395                               400

Val Gly Glu Leu Phe Pro Leu Arg Thr Arg Ala Lys Ser Val Ser Leu
                               405                               410                               415

Cys Thr Ala Ser Asn Trp Leu Trp Asn Trp Gly Ile Ala Tyr Ala Thr
                               420                               425                               430

Pro Tyr Met Val Asp Glu Asp Lys Gly Asn Leu Gly Ser Asn Val Phe
                               435                               440                               445

Phe Ile Trp Gly Gly Phe Asn Leu Ala Cys Val Phe Phe Ala Trp Tyr
450                               455                               460

Phe Ile Tyr Glu Thr Lys Gly Leu Ser Leu Glu Gln Val Asp Glu Leu
465                               470                               475                               480

Tyr Glu His Val Ser Lys Ala Trp Lys Ser Lys Gly Phe Val Pro Ser
                               485                               490                               495

Lys His Ser Phe Arg Glu Gln Val Asp Gln Gln Met Asp Ser Lys Thr
                               500                               505                               510

Glu Ala Ile Met Ser Glu Glu Ala Ser Val
515                               520

```

<210> SEQ ID NO 3

<211> LENGTH: 522

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

-continued

<400> SEQUENCE: 3

Met Gly Leu Glu Asp Asn Arg Met Val Lys Arg Phe Val Asn Val Gly
 1 5 10 15
 Glu Lys Lys Ala Gly Ser Thr Ala Met Ala Ile Ile Val Gly Leu Phe
 20 25 30
 Ala Ala Ser Gly Gly Phe Ile Met Gly Tyr Asp Thr Gly Thr Ile Ser
 35 40 45
 Gly Val Met Thr Met Asp Tyr Val Leu Ala Arg Tyr Pro Ser Asn Lys
 50 55 60
 His Ser Phe Thr Ala Asp Glu Ser Ser Leu Ile Val Ser Ile Leu Ser
 65 70 75 80
 Val Gly Thr Phe Phe Gly Ala Leu Cys Ala Pro Phe Leu Asn Asp Thr
 85 90 95
 Leu Gly Arg Arg Trp Cys Leu Ile Leu Ser Ala Leu Ile Val Phe Asn
 100 105 110
 Ile Gly Ala Ile Leu Gln Val Ile Ser Thr Ala Ile Pro Leu Leu Cys
 115 120 125
 Ala Gly Arg Val Ile Ala Gly Phe Gly Val Gly Leu Ile Ser Ala Thr
 130 135 140
 Ile Pro Leu Tyr Gln Ser Glu Thr Ala Pro Lys Trp Ile Arg Gly Ala
 145 150 155 160
 Ile Val Ser Cys Tyr Gln Trp Ala Ile Thr Ile Gly Leu Phe Leu Ala
 165 170 175
 Ser Cys Val Asn Lys Gly Thr Glu His Met Thr Asn Ser Gly Ser Tyr
 180 185 190
 Arg Ile Pro Leu Ala Ile Gln Cys Leu Trp Gly Leu Ile Leu Gly Ile
 195 200 205
 Gly Met Ile Phe Leu Pro Glu Thr Pro Arg Phe Trp Ile Ser Lys Gly
 210 215 220
 Asn Gln Glu Lys Ala Ala Glu Ser Leu Ala Arg Leu Arg Lys Leu Pro
 225 230 235 240
 Ile Asp His Pro Asp Ser Leu Glu Glu Leu Arg Asp Ile Thr Ala Ala
 245 250 255
 Tyr Glu Phe Glu Thr Val Tyr Gly Lys Ser Ser Trp Ser Gln Val Phe
 260 265 270
 Ser His Lys Asn His Gln Leu Lys Arg Leu Phe Thr Gly Val Ala Ile
 275 280 285
 Gln Ala Phe Gln Gln Leu Thr Gly Val Asn Phe Ile Phe Tyr Tyr Gly
 290 295 300
 Thr Thr Phe Phe Lys Arg Ala Gly Val Asn Gly Phe Thr Ile Ser Leu
 305 310 315 320
 Ala Thr Asn Ile Val His Val Gly Ser Thr Ile Pro Gly Ile Leu Leu
 325 330 335
 Met Glu Val Leu Gly Arg Arg Asn Met Leu Met Gly Gly Ala Thr Gly
 340 345 350
 Met Ser Leu Ser Gln Leu Ile Val Ala Ile Val Gly Val Ala Thr Ser
 355 360 365
 Glu Asn Asn Lys Ser Ser Gln Ser Val Leu Val Ala Phe Ser Cys Ile
 370 375 380
 Phe Ile Ala Phe Phe Ala Ala Thr Trp Gly Pro Cys Ala Trp Val Val
 385 390 395 400
 Val Gly Glu Leu Phe Pro Leu Arg Thr Arg Ala Lys Ser Val Ser Leu

-continued

	405		410		415
Cys Thr Ala Ser Asn Trp Leu Trp Asn Trp Gly Ile Ala Tyr Ala Thr	420		425		430
Pro Tyr Met Val Asp Glu Asp Lys Gly Asn Leu Gly Ser Asn Val Phe	435		440		445
Phe Ile Trp Gly Gly Phe Asn Leu Ala Cys Val Phe Phe Ala Trp Tyr	450		455		460
Phe Ile Tyr Glu Thr Lys Gly Leu Ser Leu Glu Gln Val Asp Glu Leu	465		470		475
Tyr Glu His Val Ser Lys Ala Trp Lys Ser Lys Gly Phe Val Pro Ser	485		490		495
Lys His Ser Phe Arg Glu Gln Val Asp Gln Gln Met Asp Ser Lys Thr	500		505		510
Glu Ala Ile Met Ser Glu Glu Ala Ser Val	515		520		

<210> SEQ ID NO 4

<211> LENGTH: 522

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 4

Met Gly Leu Glu Asp Asn Arg Met Val Lys Arg Phe Val Asn Val Gly	1	5	10	15
Glu Lys Lys Ala Gly Ser Thr Ala Met Ala Ile Ile Val Gly Leu Phe	20	25	30	
Ala Ala Ser Gly Gly Phe Ile Met Gly Tyr Asp Thr Gly Thr Ile Ser	35	40	45	
Gly Val Met Thr Met Asp Tyr Val Leu Ala Arg Tyr Pro Ser Asn Lys	50	55	60	
His Ser Phe Thr Ala Asp Glu Ser Ser Leu Ile Val Ser Ile Leu Ser	65	70	75	80
Val Gly Thr Phe Phe Gly Ala Leu Cys Ala Pro Phe Leu Asn Asp Thr	85	90	95	
Leu Gly Arg Arg Trp Cys Leu Ile Leu Ser Ala Leu Ile Val Phe Asn	100	105	110	
Ile Gly Ala Ile Leu Gln Val Ile Ser Thr Ala Ile Pro Leu Leu Cys	115	120	125	
Ala Gly Arg Val Ile Ala Gly Phe Gly Val Gly Leu Ile Ser Ala Thr	130	135	140	
Ile Pro Leu Tyr Gln Ser Glu Thr Ala Pro Lys Trp Ile Arg Gly Ala	145	150	155	160
Ile Val Ser Cys Tyr Gln Trp Ala Ile Asn Ile Gly Leu Phe Leu Ala	165	170	175	
Ser Cys Val Asn Lys Gly Thr Glu His Met Thr Asn Ser Gly Ser Tyr	180	185	190	
Arg Ile Pro Leu Ala Ile Gln Cys Leu Trp Gly Leu Ile Leu Gly Ile	195	200	205	
Gly Met Ile Phe Leu Pro Glu Thr Pro Arg Phe Trp Ile Ser Lys Gly	210	215	220	
Asn Gln Glu Lys Ala Ala Glu Ser Leu Ala Arg Leu Arg Lys Leu Pro	225	230	235	240
Ile Asp His Pro Asp Ser Leu Glu Glu Leu Arg Asp Ile Thr Ala Ala				

-continued

245					250					255					
Tyr	Glu	Phe	Glu	Thr	Val	Tyr	Gly	Lys	Ser	Ser	Trp	Ser	Gln	Val	Phe
			260					265					270		
Ser	His	Lys	Asn	His	Gln	Leu	Lys	Arg	Leu	Phe	Thr	Gly	Val	Ala	Ile
		275						280					285		
Gln	Ala	Phe	Gln	Gln	Leu	Thr	Gly	Val	Asn	Phe	Ile	Phe	Tyr	Tyr	Gly
		290						295					300		
Thr	Thr	Phe	Phe	Lys	Arg	Ala	Gly	Val	Asn	Gly	Phe	Thr	Ile	Ser	Leu
		305						310					315		320
Ala	Thr	Asn	Ile	Val	Asn	Val	Gly	Ser	Thr	Ile	Pro	Gly	Ile	Leu	Leu
				325						330					335
Met	Glu	Val	Leu	Gly	Arg	Arg	Asn	Met	Leu	Met	Gly	Gly	Ala	Thr	Gly
			340						345					350	
Met	Ser	Leu	Ser	Gln	Leu	Ile	Val	Ala	Ile	Val	Gly	Val	Ala	Thr	Ser
		355							360				365		
Glu	Asn	Asn	Lys	Ser	Ser	Gln	Ser	Val	Leu	Val	Ala	Phe	Ser	Cys	Ile
		370						375					380		
Phe	Ile	Ala	Phe	Phe	Ala	Ala	Thr	Trp	Gly	Pro	Cys	Ala	Trp	Val	Val
		385						390					395		400
Val	Gly	Glu	Leu	Phe	Pro	Leu	Arg	Thr	Arg	Ala	Lys	Ser	Val	Ser	Leu
				405						410					415
Cys	Thr	Ala	Ser	Asn	Trp	Leu	Trp	Asn	Trp	Gly	Ile	Ala	Tyr	Ala	Thr
				420						425				430	
Pro	Tyr	Met	Val	Asp	Glu	Asp	Lys	Gly	Asn	Leu	Gly	Ser	Asn	Val	Phe
		435							440					445	
Phe	Ile	Trp	Gly	Gly	Phe	Asn	Leu	Ala	Cys	Val	Phe	Phe	Ala	Trp	Tyr
		450							455					460	
Phe	Ile	Tyr	Glu	Thr	Lys	Gly	Leu	Ser	Leu	Glu	Gln	Val	Asp	Glu	Leu
		465							470					475	480
Tyr	Glu	His	Val	Ser	Lys	Ala	Trp	Lys	Ser	Lys	Gly	Phe	Val	Pro	Ser
				485						490					495
Lys	His	Ser	Phe	Arg	Glu	Gln	Val	Asp	Gln	Gln	Met	Asp	Ser	Lys	Thr
			500							505					510
Glu	Ala	Ile	Met	Ser	Glu	Glu	Ala	Ser	Val						
		515							520						

<210> SEQ ID NO 5

<211> LENGTH: 496

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 5

Met	Gly	Leu	Glu	Asp	Asn	Arg	Met	Val	Lys	Arg	Phe	Val	Asn	Val	Gly
1				5						10				15	
Glu	Lys	Lys	Ala	Gly	Ser	Thr	Ala	Met	Ala	Ile	Ile	Val	Gly	Leu	Phe
			20						25					30	
Ala	Ala	Ser	Gly	Gly	Phe	Ile	Met	Gly	Tyr	Asp	Thr	Gly	Thr	Ile	Ser
			35						40					45	
Gly	Val	Met	Thr	Met	Asp	Tyr	Val	Leu	Ala	Arg	Tyr	Pro	Ser	Asn	Lys
		50							55					60	
His	Ser	Phe	Thr	Ala	Asp	Glu	Ser	Ser	Leu	Ile	Val	Ser	Ile	Leu	Ser
				70						75				80	
Val	Gly	Thr	Phe	Phe	Gly	Ala	Leu	Cys	Ala	Pro	Phe	Leu	Asn	Asp	Thr

-continued

85					90					95					
Leu	Gly	Arg	Arg	Trp	Cys	Leu	Ile	Leu	Ser	Ala	Leu	Ile	Val	Phe	Asn
		100						105					110		
Ile	Gly	Ala	Ile	Leu	Gln	Val	Ile	Ser	Thr	Ala	Ile	Pro	Leu	Leu	Cys
		115					120					125			
Ala	Gly	Arg	Val	Ile	Ala	Gly	Phe	Gly	Val	Gly	Leu	Ile	Ser	Ala	Thr
	130					135					140				
Ile	Pro	Leu	Tyr	Gln	Ser	Glu	Thr	Ala	Pro	Lys	Trp	Ile	Arg	Gly	Ala
145					150					155					160
Ile	Val	Ser	Cys	Tyr	Gln	Trp	Ala	Ile	Thr	Ile	Gly	Leu	Phe	Leu	Ala
				165					170						175
Ser	Cys	Val	Asn	Lys	Gly	Thr	Glu	His	Met	Thr	Asn	Ser	Gly	Ser	Tyr
			180						185					190	
Arg	Ile	Pro	Leu	Ala	Ile	Gln	Cys	Leu	Trp	Gly	Leu	Ile	Leu	Gly	Ile
		195					200						205		
Gly	Met	Ile	Phe	Leu	Pro	Glu	Thr	Pro	Arg	Phe	Trp	Ile	Ser	Lys	Gly
	210					215					220				
Asn	Gln	Glu	Lys	Ala	Ala	Glu	Ser	Leu	Ala	Arg	Leu	Arg	Lys	Leu	Pro
225					230					235					240
Ile	Asp	His	Pro	Asp	Ser	Leu	Glu	Glu	Leu	Arg	Asp	Ile	Thr	Ala	Ala
				245					250						255
Tyr	Glu	Phe	Glu	Thr	Val	Tyr	Gly	Lys	Ser	Ser	Trp	Ser	Gln	Val	Phe
		260						265						270	
Ser	His	Lys	Asn	His	Gln	Leu	Lys	Arg	Leu	Phe	Thr	Gly	Val	Ala	Ile
		275					280						285		
Gln	Ala	Phe	Gln	Gln	Leu	Thr	Gly	Val	Asn	Phe	Ile	Phe	Tyr	Tyr	Gly
	290					295					300				
Thr	Thr	Phe	Phe	Lys	Arg	Ala	Gly	Val	Asn	Gly	Phe	Thr	Ile	Ser	Leu
305					310					315					320
Ala	Thr	Asn	Ile	Val	His	Val	Gly	Ser	Thr	Ile	Pro	Gly	Ile	Leu	Leu
				325					330						335
Met	Glu	Val	Leu	Gly	Arg	Arg	Asn	Met	Leu	Met	Gly	Gly	Ala	Thr	Gly
		340						345					350		
Met	Ser	Leu	Ser	Gln	Leu	Ile	Val	Ala	Ile	Val	Gly	Val	Ala	Thr	Ser
		355					360						365		
Glu	Asn	Asn	Lys	Ser	Ser	Gln	Ser	Val	Leu	Val	Ala	Phe	Ser	Cys	Ile
	370					375					380				
Phe	Ile	Ala	Phe	Phe	Ala	Ala	Thr	Trp	Gly	Pro	Cys	Ala	Trp	Val	Val
385					390					395					400
Val	Gly	Glu	Leu	Phe	Pro	Leu	Arg	Thr	Arg	Ala	Lys	Ser	Val	Ser	Leu
				405					410						415
Cys	Thr	Ala	Ser	Asn	Trp	Leu	Trp	Asn	Trp	Gly	Ile	Ala	Tyr	Ala	Thr
			420					425						430	
Pro	Tyr	Met	Val	Asp	Glu	Asp	Lys	Gly	Asn	Leu	Gly	Ser	Asn	Val	Phe
		435					440						445		
Phe	Ile	Trp	Gly	Gly	Phe	Asn	Leu	Ala	Cys	Val	Phe	Phe	Ala	Trp	Tyr
450						455							460		
Phe	Ile	Tyr	Glu	Thr	Lys	Gly	Leu	Ser	Leu	Glu	Gln	Val	Asp	Glu	Leu
465					470					475					480
Tyr	Glu	His	Val	Ser	Lys	Ala	Trp	Lys	Ser	Lys	Gly	Phe	Val	Pro	Ser
				485					490						495

-continued

```

<211> LENGTH: 496
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 6

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

Glu Lys Lys Ala Gly Ser Thr Ala Met Ala Ile Ile Val Gly Leu Phe
20          25          30

Ala Ala Ser Gly Gly Phe Ile Met Gly Tyr Asp Thr Gly Thr Ile Ser
35          40          45

Gly Val Met Thr Met Asp Tyr Val Leu Ala Arg Tyr Pro Ser Asn Lys
50          55          60

His Ser Phe Thr Ala Asp Glu Ser Ser Leu Ile Val Ser Ile Leu Ser
65          70          75          80

Val Gly Thr Phe Phe Gly Ala Leu Cys Ala Pro Phe Leu Asn Asp Thr
85          90          95

Leu Gly Arg Arg Trp Cys Leu Ile Leu Ser Ala Leu Ile Val Phe Asn
100         105         110

Ile Gly Ala Ile Leu Gln Val Ile Ser Thr Ala Ile Pro Leu Leu Cys
115         120         125

Ala Gly Arg Val Ile Ala Gly Phe Gly Val Gly Leu Ile Ser Ala Thr
130         135         140

Ile Pro Leu Tyr Gln Ser Glu Thr Ala Pro Lys Trp Ile Arg Gly Ala
145         150         155         160

Ile Val Ser Cys Tyr Gln Trp Ala Ile Thr Phe Gly Leu Phe Leu Ala
165         170         175

Ser Cys Val Asn Lys Gly Thr Glu His Met Thr Asn Ser Gly Ser Tyr
180         185         190

Arg Ile Pro Leu Ala Ile Gln Cys Leu Trp Gly Leu Ile Leu Gly Ile
195         200         205

Gly Met Ile Phe Leu Pro Glu Thr Pro Arg Phe Trp Ile Ser Lys Gly
210         215         220

Asn Gln Glu Lys Ala Ala Glu Ser Leu Ala Arg Leu Arg Lys Leu Pro
225         230         235         240

Ile Asp His Pro Asp Ser Leu Glu Glu Leu Arg Asp Ile Thr Ala Ala
245         250         255

Tyr Glu Phe Glu Thr Val Tyr Gly Lys Ser Ser Trp Ser Gln Val Phe
260         265         270

Ser His Lys Asn His Gln Leu Lys Arg Leu Phe Thr Gly Val Ala Ile
275         280         285

Gln Ala Phe Gln Gln Leu Thr Gly Val Asn Phe Ile Phe Tyr Tyr Gly
290         295         300

Thr Thr Phe Phe Lys Arg Ala Gly Val Asn Gly Phe Thr Ile Ser Leu
305         310         315         320

Ala Thr Asn Ile Val His Val Gly Ser Thr Ile Pro Gly Ile Leu Leu
325         330         335

Met Glu Val Leu Gly Arg Arg Asn Met Leu Met Gly Gly Ala Thr Gly
340         345         350

Met Ser Leu Ser Gln Leu Ile Val Ala Ile Val Gly Val Ala Thr Ser
355         360         365

Glu Asn Asn Lys Ser Ser Gln Ser Val Leu Val Ala Phe Ser Cys Ile
370         375         380

```

-continued

Phe Ile Ala Phe Phe Ala Ala Thr Trp Gly Pro Cys Ala Trp Val Val
 385 390 395 400

Val Gly Glu Leu Phe Pro Leu Arg Thr Arg Ala Lys Ser Val Ser Leu
 405 410 415

Cys Thr Ala Ser Asn Trp Leu Trp Asn Trp Gly Ile Ala Tyr Ala Thr
 420 425 430

Pro Tyr Met Val Asp Glu Asp Lys Gly Asn Leu Gly Ser Asn Val Phe
 435 440 445

Phe Ile Trp Gly Gly Phe Asn Leu Ala Cys Val Phe Phe Ala Trp Tyr
 450 455 460

Phe Ile Tyr Glu Thr Lys Gly Leu Ser Leu Glu Gln Val Asp Glu Leu
 465 470 475 480

Tyr Glu His Val Ser Lys Ala Trp Lys Ser Lys Gly Phe Val Pro Ser
 485 490 495

<210> SEQ ID NO 7

<211> LENGTH: 3086

<212> TYPE: DNA

<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 7

cgaagatgag agataagcga ccaaaattag cgggagaatc ctcaactgtc tcaactccgat 60

ccgatcatgg ctttcgctgt ctccggttcag tcacatttcg caatcagagc gttaaaacga 120

gaccacttca agaacccttc tctctgtact ttctgctcgt gttttaaatc gaggcctgac 180

tcgtcttacc ttagtttaaa ggaacgtact tgcttcgctt ccaaaccggg ttagtcaact 240

actagataca gacatatatt ccagggtgaaa agcaattttc tctgattttg atttcttcta 300

aaaaaaaaatc aaaagttttg atttttccgg ttttggttca tggttaggtc ggagctgaga 360

cgggaggaga gttcccgac agtggagaag tagctgattc gcttgcttct gatgcaccag 420

agtcattttc ttggtcttct gtgatactcc cgtgcgtatt gatccttctc gttaatttga 480

atttaactgt tctcgagtaa tggatttga ttgcctgtgt aatgatcttg cctaagtact 540

taatctctta gcaatgcctt atgtatcctc aattatagaa gatagaaaat tagttatttg 600

cctagattgc actgatttaa tagtaattta cttgaatcgt aggtttatct tcccggcttt 660

gggaggatta ttgtttgggt atgacattgg ggctacctcc ggtgctacgc tctcaactca 720

ggttatttta aagtatctat tttcatctag ttacatttta gctttcagaa tttaacatt 780

atgtactgtc tcatatgggc ttgaacagtc acctgcgctt agcggaacta catggtttaa 840

cttctcaact gttcagctag gacttgtggt atgttatttg gagatcgata tttctgttag 900

ttaagccata gagttagcag aaaatgatag tttttactgc atttgtgtgt gtaggttagc 960

ggatccttgt atggagccct tcttggtcca atttctgtct atggcgttgc tgatttcctt 1020

ggtaagtctt gtttttttgg gttgacttct cgttcttctt aactgaatgc aagtatctca 1080

ttctggtttt cttcacatct ttatgaagga agaaggcggg aacttattat agctgctgtt 1140

ctctatctcc tcgggtctct gatcaactggc tgtgccctgt atcttaatat tctcttagtt 1200

ggaaggcttc tctatggcct tggatttggg ttggtgagct ccggaacct gaatcgttat 1260

ggtaatttct ttgtactttt tgttgttgat tggtagttaa tgttttatgt aaattggttt 1320

tgcttgatag gcaatgcatg gggctcccct ctatattgct gagacatgcc catctcaaat 1380

ccgtggaact ttgatctctc tgaagaact cttcatctga ttgggaattt tggttaagtgt 1440

ctgatgtcaa tctcttccca gtatgatttc tgcgtaaata ttgatttctt tcttgtgcag 1500

-continued

ttgggttttt ctggtggaag cttccagatt gatgtagttg gagggtgccg ttacatgtat	1560
ggatttggtta cgctgtgtgc tttgctgatg ggactaggca tgtggagtct cctgcatct	1620
cctcgctggt tgctgcttag agctgtocaa ggtaaaggac aattacaaga atacaaagag	1680
aaggccatgc ttgccctcag caaattacgt ggcagacctc caggtgataa aatctcagag	1740
aagttggtag atgatgccta tttatctgtg aaaacggcct atgaagatga gaaatctggg	1800
ggaaacttcc tggaaagtatt ccaagggcct aatttgaaag ctttgacaat tgggtggaggt	1860
ttagtctct tccaacaggt gattcttctt cgctgtttcc atttggatga atgtgtgagc	1920
atTTTTGAAA taatttacac tctgcttctg tgtgacagat aactggacag cctagtgttc	1980
tttattatgc gggttcgatt cttcaggat gctcgcctta acattgaaat gaatgagatt	2040
acctactaat ttttactgcc tttagtcgga tgtttaatga gactttatgc tcaattatct	2100
attcaaagac tgctggattc tctgctgctg ctgatgcaac tcgagtctct gttattattg	2160
gtgttttcaa ggtggccctt tttatTTTTT tgtttggatg tgtaaatatc ttatTTTTCA	2220
acaagcttcc agttattcaa tactaacctc ttcaattgat aacgctctcg tagttactga	2280
tgacatgggt agctgttgcg aaagttgatg atctcggcag acgaccttta ctgattggag	2340
gtgtcagcgg cattgtatga attcatttta tgtctatatac ttctgtttct tattttccaa	2400
agaaaagata tcattttctta tattttcttc aaattccagg cgttgtcctt gtttctactg	2460
tcagcatact acaagtttct cggaggcttt ccccttctcg ctgttgggtc actgcttctc	2520
tatgttggtt gttaccaggt ttgactctta gctcactgtt agctgtggct acatagtttt	2580
atctagcata ttactaaagt ctcagcgcga acagatctca tttggacca tcagctggct	2640
aatggtgtca gagatTTTCC cgctccgcac aagagggaga gggatcagtc ttgcagttct	2700
tacaaacttt ggctccaatg ctattgtgac atttgcattt tcacctttaa aggtatattt	2760
ttctgttctt gcttgttcaa ccttgaagt tattagtaac tcttatcaaa atatgcattc	2820
tctgtaggaa tttcttgag ccgagaatct tttccttctc tttgggggca tagcactggg	2880
atcactgctg tttgtaatac tagtagttcc agagaccaag ggtctcagct tggagaagaat	2940
tgaatcaaaa atcttgaagt gaaacgttga agaacatatt tgttatagtt gattctgggtg	3000
aaatacatgt atgaggatgt gactattctc tttgtaacat aatatgtttc gtaatcaatg	3060
ggaaaaccaa acctTTTTCA tgatta	3086

<210> SEQ ID NO 8

<211> LENGTH: 5119

<212> TYPE: DNA

<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 8

gtgaagatga ttcaatgtcg ttgagtggta gtagttaaat cttcaaaacc cgggagaatt	60
atataatggt gtttcgttca gcaaacctaa ctaaactcag aaccagaatt catatTTTCG	120
tcgctcaaat cacagctatg gcgcttgatc ctgagcagca gcaaccaatc tctctgttt	180
ccagagaggt atacgtctcc tttcattgtc tacttctctg ttttgtttcc tgttcgatta	240
aacggtgtga gagagtTTTCC atttgaagta tacgtctcct ttcactctgaa gcttaattct	300
ctagtaaatt cgtacttctt gacttgccac attaggaatg ttccttttga gatcaacttt	360
ggctTTTAGC tgacaaatTT ctggaaaaat tctgaaatTT gctatgttgg aatctTAGAG	420
attgtattta agagtcatat attcaacttt tggggactcg cttcaattct gagaagaat	480

-continued

ttagaggctt ttacggatth ttagecatca gcctcacagt tatctttaag catgaaatth	540
ttagacccta aatgctgtgt acaattggth tgttggtthc atgatttgth tgttccttac	600
tgttttttht tttttggthg gcaactthct gaaactgtht atgctgttht gtttggtahg	660
tcatctgthg agatcagccc agaaaggthg cctcttatta aagagahatca tgtcccagah	720
aactactctg tctgtgcctg cattctcccg tgagatctct catctctctc ccaagththg	780
tagttcttht aatgattcaa atgaaacact cttgaaacat tccgathgth tctthththg	840
gthtatthth gtatcathh gcataathg thctctgact aggaaactag aacgctgthg	900
atggtacctt gcatagctth caaatagah thctththgth atthcatacc atctgcagth	960
atthggacct atgthcttht gctgathgth gththththg ththgathg thctththg	1020
atthgathat gcatgthah agththctgth thctththtca gththththt thctgctctc	1080
ggaggactgc thththgtht tgaaththgth gcaactthct gthgcaacct thctctthg	1140
gthaaagtht aagththctg thactatctg thctctacahh atthththt atthcaacth	1200
thththcaagth gctgathctg ththththcah gctthaaath ththaaagctc thgaaaththg	1260
actgathat cagthctgah acathgath thctactthg cagacctth ththctctthg	1320
atthctctctg thathththg thctactgath gthctgthth thathththg thctacttht	1380
taggagctca thactththg thctthctat gctgthctct ththctgctg cthctcttht	1440
aagthggaath thctgththt actththctc agthgathth gthctgathth thagththth	1500
agthththth gathctthct thththctct gthctththg thctththt thctctgctc	1560
thththgath gthththgca thththgath agathaaath gctthcaahh thgththth	1620
thththctcag acctththgth cactththctg thctththt gthctcaathth thgctththt	1680
thctgctgath gthththgth agthththct thctactthct thctththth gthctgath	1740
ththgcaath ththctththg thaaahgath agthththgth ththththgath thggaagth	1800
gaaahgath gctgaththg gctgaththt ththctctgth thgagctct thgactgth	1860
thgctthctc ctathctgth ctgathath gctgththt ththththgth thctgthgth	1920
thgthgath gathctgct ththththgth thctacttht agctththct thgctgth	1980
atggtctcah cththththt ththgathth gctgththct aagathctgath ctctgthth	2040
gagthgthgath aathththgth thththgath thththgctc thctacttht atthththgth	2100
gthgactct thctagthct ththththt thactththct cctctththh gthgthgth	2160
thctththct gththththt aactthctct aactththct atgathctth cagththgth	2220
aathththt ctathctthg thgagathgth gththththct agthgathct ththgaaah	2280
thaththgth ctathctthg thctaththt cagathctc atathththg actgathct	2340
thgctctctg thththththt gthththctc ththththgct agthgathgath aactththt	2400
actththgath aathththt atctththgth thththctgth gaththctat agctactc	2460
thgctctgth aactgathth ththctctca thththgath gathgctgth ctgathth	2520
thgctgath thctcaagth ctgathctgth gathctgth atctththh gaththctth	2580
thgctctthg gathgththt agthctctgath acctgath thththgctgth gathththt	2640
thgthgath thgaththgath atctththct accththth thctctctat thctctcah	2700
aaththctc agthgththt thgathctgth agthctctc thctgathth thctgththg	2760
ctgathctgth atgcaahh thctctctthg gathththt thggaathth gathgththg	2820
ctctctgath ctctgthth gctthththg cctgctctc agthgaaagth gaththgth	2880

-continued

aaccaacgag	aggctgcaat	caaatctctc	tgctgcctta	gaggccctgc	ttttgttgat	2940
tcagctgctg	aacaagtaaa	tgagattttg	gctgaactaa	cttttgtggg	cgaggataaa	3000
gaagtccaat	ttggtgaatt	atttcaaggg	aagtgcctga	aagctctcat	cataggagga	3060
ggtctagtct	tgtttcagca	ggttggtata	gtcaaatcta	atcattgtgt	tctcttgaaa	3120
gaataattac	taatgaattc	ttgtctaattg	gcagataact	ggacaaccaa	gtgtgcttta	3180
ttatgcacca	tcaatactgc	aggttttctt	cttctctaaa	taaactttct	gtatagtata	3240
atattggctt	tgttgcaact	cgaaaactcac	gactgggatt	ttaacttctc	cccccatccc	3300
acatttctcg	ttagaccgct	ggcttttcag	ccgcaggtga	tgcaacaagg	gtctcaattc	3360
tacttggtct	attgaagggt	ttcactgttt	cccactggc	taaaatttaa	gttacttttc	3420
caatgtaata	aagaaatgcc	gttttagtta	ttttgccttc	aaagaaataa	caccgcctc	3480
tgttgttgca	gttgattatg	acaggagtag	ctggtgtagt	tatcgacaga	cttgaagga	3540
gacctttact	tcttggtgga	ggttggtgca	tggtatgtcc	tctgtcttta	tcccttcctc	3600
tttctcatcc	atgttccggt	tatacacctt	aagcaaaagt	acagtaaat	taggtcaagc	3660
tcactaaatg	ttcagctcaa	tgatagagag	aaccaatac	tcctaataac	attttatttt	3720
caatgtcaaa	caggttggtt	cattattcct	gctggggcca	tactacctct	ttttcagcgc	3780
ctcaccagtt	gttgctgtag	ttgcaactgct	gctttatgtg	ggctggtacc	aggtaaacaa	3840
atagttgatt	agaatthtgc	taatcaaaact	tcactaccac	aatgtcaatg	cttgccttaa	3900
aagtatcata	gactttttga	tataactcgt	ggcatctaaa	gttggagtcc	ttataacaag	3960
tagttccatt	ggagtttatg	tacaaaagtg	attcatggat	ttcattattt	tcgccttaaa	4020
acttttgata	tgcttctagt	tcttttttcc	aacatgtttg	gctttctact	tttcaataag	4080
agcttgaacg	tattaccatc	aatcctgcag	ctatcttttg	gtccaattgg	ttggctgatg	4140
atthcagaga	tatttccctc	taaaactaaga	ggtcagagtc	tcagctctagc	agtgcttggtg	4200
aatthtgggtg	caaacgcact	tgtgacattt	gcgttttcac	ccctaaaggt	aacctaccat	4260
tcccttgagg	acaaaaatc	tcgtttgttt	aagcttggtta	tggccattac	tgtagaaatg	4320
gtagcaaccc	ttttgcaaga	cttttttact	ttagaagaaa	ttgtcaatca	tatggtttagt	4380
cgtggtgaaa	ttctagtoga	atgatatcga	caccacagct	tagagtagtt	tttctctga	4440
tatctaactg	gaacttttccg	tttggtctca	atggctctgt	tttgcaggaa	ctgctgggag	4500
ctgggatact	gttttgtggg	tttggtgtga	tatgtgtatt	gtctctggta	ttcatattct	4560
tcattgtgcc	agagacaaag	ggtctcactc	ttgaggaaat	tgaagccaaa	tgtctctaaa	4620
aaagaggtct	gttctttgct	tagaaaccac	aaagtcgtgt	gcttctctcac	atgattttga	4680
ttcatattgt	taatcagtga	ggaaaataat	agtgcaggtt	cagaaacaaa	taaacatag	4740
aatctgccgc	acaagacggg	aaatgaatct	tcagagacca	ctaagttatt	tgaagcaatg	4800
ttacttcaaa	ggctcggttg	tttctagcaa	aatacatgta	cgagaattca	taaatacaga	4860
aatctttgta	atgattatta	gcgctctgat	gaagttagaa	aataaaaaaa	gaaaacatca	4920
tagaagaatt	taaatttgta	gaatatgtcc	taaccagtga	tgtttcgaaa	tccgaagggt	4980
tctcaaagtt	tgtatthttt	ttaaacgatt	ccacgattct	gcaatgctgc	attatgatat	5040
agaacattat	gctgaataga	agatathttt	cgggatttgt	aagacttgat	gtgatatagt	5100
ataatggaac	attgtggctc					5119

-continued

<211> LENGTH: 1644

<212> TYPE: DNA

<213> ORGANISM: Candida intermedia

<400> SEQUENCE: 9

```

atgtcacaag attcgcattc ttctggtgcc gctacaccag tcaatggttc catccttgaa    60
aaggaaaaag aagactctcc agttcttcaa gttgatgccc cacaaaaggg tttcaaggac    120
tacattgtca tttctatcct ctgttttatg gttgccttcg gtggtttcgt cttcggtttc    180
gacactggta ccatttcogg tttcgtgaac atgtctgact ttaaagacag attcgggtcaa    240
caccacgctg atggctactc ttacttctcc gacgttagag ttggtttgat gattttctatt    300
ttcaacgttg gttgcgctgt cgggtgtatt ttcctctgca aggtcctgga tgtctggggg    360
agaagaattg gtcttatggt ctccatggct gtctacgttg ttggtattat tattcagatc    420
tcttcatcca ccaagtggta ccagttcttc attggtctgc ttattgtctg tttgggtggt    480
ggtaaccgtt ctgtcgtttc cccacttttc atctctgagg tttctccaaa gcaaattaga    540
ggtaactttag tgtcgtcgtt ccagttgtgt atcaccttgg gtatcttctt gggttactgt    600
actacttacg gtactaagac ctacactgac tctagacagt ggagaattcc tttgggtttg    660
tgtttcgctt gggctatcct gttggttgct ggtatgttga acatgccaga gtctccaaga    720
tacttggttg agaagcacag aattgatgag gccaaagat ccattgccag atccaacaag    780
atccctgagg aggaccatt cgtctacact gaggttcagc ttattcaggc cggtattgag    840
agagaagctt tggtctgca ggcactcttg aaggagtga tcaactggta gccaaagatc    900
ttcagaagag ttatcatggg tattatgctt cagtccttgc aacagttgac cggtgacaac    960
tacttctctc actacggtac taccattttc caggctgtcg gtttgaagga ttctttccag   1020
acttctatca ttttgggtat tgtcaactt gcttccacct tcgttggtat ctatgtcatt   1080
gagagattgg gtagaagatt gtgtcttttg accggttccg ctgctatggt catctgtttc   1140
atcatctact ctttgattgg tactcagcac ttgtacaagc aaggttactc caacgagacc   1200
tccaacactt acaaggcttc tggtaacgct atgatcttca tcaactgtct ttacattttc   1260
ttctttgctt ctacctgggc tgggtggtgt tactgtatca tttccgagtc ctaccattg   1320
agaattagat ccaaggccat gtctattgct accgctgcta actggttgtg gggtttcttg   1380
atctccttct tcactccatt catcaccagt gccatccact tctactacgg tttcgttttc   1440
actggttgtt tggctttctc tttcttttac gtctacttct tcgtctacga aaccaagggt   1500
ctttctttgg aggaggttga tgagatgtac gcttccggtg ttcttccact caagtctgcc   1560
agctgggttc caccaaaatc tgagcacatg gctcactctg ccggttacgc tgggtctgac   1620
aaggccaccg acgaacaggt ttaa                                     1644

```

<210> SEQ ID NO 10

<211> LENGTH: 1569

<212> TYPE: DNA

<213> ORGANISM: Candida intermedia

<400> SEQUENCE: 10

```

atgggtttgg aggacaatag aatggttaag cgtttcgtca acggtggcga gaagaaggct    60
ggctctactg ccattggccat catcgtcggg ctttttctct cttctggtgg tgcctcttcc   120
ggatacgata ctggtactat ttctggtgtg atgaccatgg actacgttct tgcctgttac   180
ccttccaaca agcactcttt tactgtgat gaatcttctt tgattgtttc tatcttgtct   240
gttggtactt tctttggtgc actttgtgct ccattcctta acgacacctc cggtagacgt   300

```

-continued

tgggtgtctta ttctttctgc tcttattgtc ttcaacattg gtgctatctt gcaggtcate	360
tctactgcca ttccattgct ttgtgctggg agagtatttg caggttttgg tgtcggtttg	420
atctctgcta ctattccatt gtaccaatct gagactgctc caaagtggat cagaggtgcc	480
attgtctctt gttaccagtg ggtattacc attggtcttt tcttggcctc ttgtgtcaac	540
aagggtactg agcacatgac taactctgga tcttacagaa ttccactgac tattcaatgt	600
ctttggggtc ttatctggg tatcggtatg atcttcttgc cagagactcc aagattctgg	660
atctccaagg gtaaccagga gaaggctgct gagtctttgg ccagattgag aaagcttcca	720
attgaccacc cagactctct cgaggaatta agagacatca ctgctgttta cgagttcgag	780
actgtgtacg gtaagtctc ttggagccag gtgttctctc acaagaacca ccagttgaag	840
agattgttca ctgggtggc tatccaggct ttccagcaat tgaccgggtg taacttcatt	900
ttctactaog gtactacott cttcaagaga gctgggttta acggtttcac tatctccttg	960
gccactaaca ttgtcaatgt cggttctact attccagga ttcttttgat ggaagtcctc	1020
ggtagaagaa acatgttgat ggggtggtct actggtatgt ctctttctca attgatcgtt	1080
gccattgttg gtgtgtctac ctcgaaaaac aacaagtctt cccagtcctg cctgtgtgct	1140
ttctcctgta ttttcattgc cttctctgct gccacctggg gtccatgtgc ttgggttgtt	1200
gttggtgagt tgttccatt gagaaccaga getaagtctg tctccttggt tactgcttcc	1260
aactggttgt ggaactgggg tattgcttac gctactccat acatggtgga tgaagacaag	1320
ggtaacttgg gttccaatgt gttctctcgc tggggtggtt tcaacttggc ttgtgttttc	1380
ttcgcttggg acttcatcta cgagaccaag ggtctttctt tggagcaggt cgacgagttg	1440
tacgagcatg tcagcaaggc ttggaagtct aagggtctcg ttccatctaa gcaactcttc	1500
agagagcagg tggaccagca aatggactcc aaaactgaag ctattatgtc tgaagaagct	1560
tctgtttaa	1569

<210> SEQ ID NO 11

<211> LENGTH: 1584

<212> TYPE: DNA

<213> ORGANISM: Debaryomyces hansenii

<400> SEQUENCE: 11

atgggttttag aagataatgc gcttattaga aagtatgtca atgtcgggga aaaaagggct	60
gggtcagcat ctatggggat tttcgtaggg gccttcgag ctttcggggg tgttttggtc	120
gggtatgata ctggaacctt ttcaggtatc atggccatga actatgtcaa aggagaattt	180
cctgccaata aggagagttt tacgtcaaaa gaaagtctgt tgattgttcc catcttatca	240
gcaggtaact tcttcggtgc gttgttagca cggtttatgt ctgatacttt aggtagaaga	300
tggtcattaa ttatttcaac attcattgtt ttcaacttgg gagtgatctt acaaactgtt	360
tcaactggta ttccattact atgtgctgga agagctattg ctggttttgg tgttggtctt	420
atatctgctg tcattccatt atatcaatca gaagccactc caaaatggat tagaggggct	480
gttgtatctt gttatcaatg ggcattact attggcttat tggtagcggc ttgtgttaac	540
caaggtaact ataataagaa tgactcgggt tcatacagaa ttccaattgc tgttcaactt	600
ttatggctgt taattttggg tactggatg atcttcttgc ctgacacgcc acgtttctgg	660
atccacaag gtaatgaatc tgaagctaag aaatcgttga agattttaag aaaattacca	720
cttgaccacc cagacttaat tgaagaatac gaagacatca aagctgttta cgatttcgag	780

-continued

tgttctttcg gaaaatcgtc ttggatggat ctttttacca caaggaatag gcaattgaag	840
agattattta cgggtgttgc tcttcaagca tttcaacaat taactgggtg taactttatc	900
ttctattttg gtacctcttt cttcaagagt gctgggtattg aaaacgaatt tcttatttct	960
ttagccacca gtattgtaa cgtaggtatg actgttccag gtatcttttt aattgaatta	1020
gtgggtcgtc gttctatggt gttgtggggg gctgttggtg tgtegggttc tcaatttatc	1080
gttgtatttg ttggatttgc cactgatagt gccgacgcta acaaagtctt gatcgcatc	1140
acttgtttct tcattgcatt ctttgcctct acctggggtc caattgcttg ggttgttgtt	1200
ggtgagattt tcccattaag aactagagct aagtcogttg cattatctgc cgcttccaac	1260
tggttttggg actgggctat cgcatacgtc actccatact tggtagaaga tggaagggt	1320
aatgctaact taggaacaaa cgttttcttc atctggggtg gatgtaactt tctttgtatt	1380
ttgttcaact acgtcttcat ttacgaaaac aagggttact ctttgaaca aattgatgaa	1440
ttatatgaaa aagttccaca tgettggaaa tctcgtggct tcattccttc tgcccatgct	1500
ttcagagaag atgctccaga gtcaatcagc tctatgggaa aggatagga aaaagttact	1560
gaaattgaaa ccacttcagt ctaa	1584

<210> SEQ ID NO 12

<211> LENGTH: 2196

<212> TYPE: DNA

<213> ORGANISM: *Debaryomyces hansenii*

<400> SEQUENCE: 12

atgtggccat ttctagatag gatcatgat gatcactcga cggaggagga atattacaaa	60
aagatgcaac agaagtgtc ttcgtcgagt gccattacag tggggctagt ggcggcggtg	120
gggggatatt tgatgggta cgatactggg ctcatatacg atattatgga gatgacgtac	180
gtgaaggata acttcccggc gaatggacat agcttcagcg tacacgagcg agcgttgata	240
acggccatct tgctactcgg aacatttttt gggcggttga tagcgcegtt gatateggat	300
acgtggggta gaaagttttc gatcattgta tcgtcagctc tcatatttaa tgtgggaaac	360
atcttgacgg tgcacccac agaggtggtg cttttatgtg ttgggagagc tgtttcgggg	420
ttgtcagtggt gtatactctc agccattgtg ccgttgatc aagctgaagc ttgcceaaaa	480
tgggtaagag gttccattgt gtatacttac caatgggcga taacatgggg gttgttgatt	540
gctagtgcta tatgtcaggg ggcaaaaaac attatgaatt cgggctcgtg tagaataccg	600
gtgggaatac agttcttctg ggcaattatt ttgtctgtcg ggatgctttt tttaccgaa	660
tctctctgat tccatgtaca gaaggataat atccaagaag cattgaaatg cttagctaga	720
ctaagaaaag tgccgacaga tgaccctgat ttgatagaag aactagttga aatcaaggcc	780
aactatgact acgagttatc gtttggaaag gcatcctata tagattgctt taagagtgg	840
ggaggaagaa ataagcaact cagcaggtat ctcacaggtg ttgggggtcca ggcatttcaa	900
caagttcag gaatcaattt tattttctat tacgggggta atttctctgc aagttcagga	960
attaagaatt attacttaat gtcatttcta acttatgcag ttaacacact tttacaata	1020
ccagggataa ttttgattga agtcatagga agaaggaaac tattattatt tggcggcatt	1080
ggaatggctg tttcaaatc cataatagca attgttgggtg ttagtatgtc agataggtca	1140
atcagttcta taatagtgt gtcattttct tgtgtcttta ttgcattttt tgcacatca	1200
tggggcgggt ccgtatgggc gctctcttca gatatatttg gtattgggtat tagacaaaa	1260
gcaatctcgc taacggcagc gacaaaattg ttagttaatt ttacctttgc atttataacg	1320

-continued

```

ccatatttga ttgatactgg taaacatacc gcagcattgg gtaacaaaat tttttttatc 1380
tggggtgggtt gtaatgctct aggcgctgta tttgtgtact ttatggttta cgaaacgaaa 1440
ggattaaaaat tggaagaat tgatttcacg tacaaaaact gtgtgaacgc tagagcttcg 1500
actaagttca aatctcaaaa gattgtatac gccaatcaaa tttcaacacc gatctcagag 1560
cttcttaatc ccaaccgatc gcatattgcg attgagaaat caagtggaag taacaataat 1620
ggcgatgatg atgagaatag cgaagaaaat caccatgatt ttggttttaga gaacggaaac 1680
gtcttacaca acaatttggg taataggaat ataactctta tcccttataa gaatattatt 1740
tcacctctgc gatcattcag ttcagattct tcaagcgatt ccgactcacc ttcaccactc 1800
aatgattatg aaagataact acatagcttg caaaaagaag gcagccatca tgacacctct 1860
caaaactcaa cgtcttaaat tactgacaat aagctaagtc atcttaattc catgcataat 1920
gcccactcac ggggtaattt tactgaggag gacttgaagt acctcaatga tgaatatgat 1980
ttagcgatgt cgcagaaaata cccaggcaca actacggcgc ctctaaaacc aagcacgaat 2040
atgactgtta tagcggcgcc attttttgac gctccaccaa gcgattcaga cacggacgac 2100
gaaagcgaag acgatgaagt cggtgacgaa gatagcgaag cagaaccaga aactactaag 2160
ggctcaccgg ttgataccaa ggactctagt tcgtag 2196

```

<210> SEQ ID NO 13

<211> LENGTH: 1593

<212> TYPE: DNA

<213> ORGANISM: *Debaryomyces hansenii*

<400> SEQUENCE: 13

```

atgttcaata aatcagatt tggtttctgc agcttaaaca aggagcttga taagtttcat 60
acgacatata atatatatgt tattgctatg attacaacaa tctcaggat gatgtttggg 120
tttgatgtct cttcaatctc tgcatttata tctgaacct cctacaggcg gttcttcaat 180
tatccaaact ctacgacaca gggtgcaatt actgcacca tgtctgcagg gagcttcttg 240
ggtgcaatat tatcatcttt tgtatcagaa agaattggga gaagaacctc attgcttttt 300
tgtgccatgt tctgggtttt aggttccata atccaatcat catgcagaaa ccttgggtcaa 360
ttaatcgcag gtcgcataat atccggggtg ggtgttggtg ttggatccgc aataacccca 420
atatattggt cagaagtatc tctgcccct tctagaggtg ttattggagg tttatttcaa 480
ttagctatca catttgggat attgataatg ttttacatag gatattggtg tacatttatt 540
aacggtcagg cgtctttcag attagcttgg gcgctacaaa tgattcccgg tttggttttg 600
tttctggtg tattcatttt acctgagtca ccacgctggt tagcaacaaa tagtaagtgg 660
gagcaagcag aagaagtgat tcgaagaatt aacgaaaaag acaaaacagg aagatatttg 720
attgaattag aggagttgaa gaaaagtatc acaattcata aattatcaaa agatataggg 780
tacctagatc ttttcaggaa aaagaactat aaaagtagta ttgttggtat ttcggctcaa 840
atctggaatc agcttacogg aatgaacgtg atgatgtatt atatagtcta tatctttgaa 900
atggtgggat aactggaaa tacagtttta gttagttaa gtatccaata tgttatcaat 960
tttggcgtga ctttaatagc tttgcctta tctgattatg ttggaagaag aagactaatg 1020
ctaactcggg gtgtactaat gatggcatgg ctatttctgt ttgggtgact ctttgcagcc 1080
tattctgaaa aggtagaaaa cgtcaccagt gacgcaaccg tggttgttac tatcccogaa 1140
gaacatcgaa atattggaaa agccattggt gcttgcctct atttattcgt tgccaccttc 1200

```

-continued

```

gctagtacgt gggcagtatg ttcatgggtg tatttttctg aagtgtcccc aactaggaca 1260
agatcaaaag caggactgtt agcagttgca agtgactggg caataaattt tgccatagct 1320
ctttttactt cttcagcatt tagaaatatt acttggaaaa cttactttgt ttttgaaca 1380
ttctgcgggc ctatgacogt tcataccttc ctttcatacc ctgaaactcg tatgaaaacg 1440
ctagaagaaa ttgatatgac tttccaaaat tgtatcccac catggaggtc cgcgggagtt 1500
actattcaaa gtcttgacgg ttatttaagt gattctaagg aaaataacgt aacgcatgtc 1560
gaacaagtgg aatcaaatag cataagtaag tag 1593

```

<210> SEQ ID NO 14

<211> LENGTH: 1731

<212> TYPE: DNA

<213> ORGANISM: Debaryomyces hansenii

<400> SEQUENCE: 14

```

atgtcttcgt tattgaccaa caaatatttc aaagactatt acaataatcc aagtccaacg 60
gcgattggta ctatgatcgc catcctagag attggtgctt tgatatacgc gttcgttgcg 120
ggtaaagtgt gtgatcatat cgaagaaga cgcactatac gttatggatc gtttatcttc 180
ataatcgggg ggcttatcca aacttcgtca atcaatatta ttaacttagg ggcaggaagg 240
tttatcagtg gtgtggcaat tgggtttttg actactatca ttccatgcta tcagtctgaa 300
atcagtccac ctgatgatag aggattctat gcgtgcttag aatttactgg taacatcatt 360
ggttactcta caagtatttg ggtggattat ggggtttcgt tcatagaaaa tgactactca 420
tggagaactc cattggccat ccaatgtgtc atgggtggcc ttttgttat tggatcattc 480
gtaatcgtag aaactccaag atggctttta gatcatgac atgatattga aggaatgatt 540
gtgatttctg atttgatagg tgatgggat gttgaagacg aattatctaa aacggaatat 600
agaaacatta aggaaaacat tctcatcgca agagtggaag gaggggaacg ttcatatcgt 660
tatatgtaa ctaagtataa gaaacgtctt tcagttgctg gcttttcgca gatgtttgca 720
caattgaatg gtattaatat ggtttcctat tatgcgccta tgatttttga gctggccggt 780
tgggttggtg gacaggccat ttaatgact ggtattaatt ccatagttta cgtattgagt 840
acaattccgc catggtatth ggtggatgg tggggcagaa agccgttatt gttgtctggc 900
gctgttgta tgggtatacc attattagtg atttcgtatt cattattcct tgataacatt 960
tatacccaa acatagtggt tgtatcggtg atcatattca atgctgcatt tggtgcaagt 1020
tgggggcaaa tcccatggat gatgaatgaa gtattaccta atagtattag atctaagggt 1080
gctgcatgt ccaactgcaac taattggcta ttcaacttca tcgtagggga aatgacccca 1140
atattgttgg ataccattaa atggagaacc tatttgatct cagccgtctc gtgtgcatta 1200
tcatttttat gcgtccactt cttgtttccg gaaacaaagg ggttgagctt agaggacatg 1260
ggctccgtat ttgatgataa ctctgtctatt ttttcgttcc attcgggtgc atccagtgga 1320
aacaatgcat caactactac tatcaataac tacggggctg cagatagaga tagcggtatc 1380
gaggttcgca gaactagtat ctccgttgaa actccaaacc atgcaaacat agttaatgaa 1440
gcatttcgcc aatccccagc atcaattgct cgtaattcaa aggccaaacc agagttagac 1500
ggctttatta ctggaatcc accaccgta ccacctgacc tctcggctct ggactccaat 1560
gtacctccac aagaaatoga accaccttca ttcgacgtca ttttcaata taaagttcgc 1620
caattagaga aaccaaacat attccaaaag gctttccgtg caatctccgg tgatagtacg 1680
ttcaaaccac cacagatoga cgaagaacgt acttctctct cgaataattg a 1731

```

-continued

<210> SEQ ID NO 15
 <211> LENGTH: 1674
 <212> TYPE: DNA
 <213> ORGANISM: Debaryomyces hansenii

<400> SEQUENCE: 15

```

atgactactg ctgttgatt agaagataat tccaaagga atattattac tgttatgagt    60
aaagatccgt tagtgttttg cataattgcg tttgcatcta ttgggggttt actctttggg    120
tacgatcagg gagttattag tggtatagtt actatggagt catttgcggc aaaattcccc    180
agaatttttt cggatcctga ttataagggt tggtttgttt ctacattttt gttgtgtgct    240
tggtttggtt cgttgataaa ctctcctgtt gttgatcgat ttggaagaag ggatactata    300
agaatagctt gcgttgatt tgttatcgga tctgtctttc aatgtgcagg tacgtcggtt    360
agtatgcttt tcgctgggag agcagttgct ggtatcgggtg tcggacaatt gactatggta    420
gtcccaatth atagtgcgga attagctcca ccttcagtaa gaggtggttt agttgtaata    480
cagcaattht caataacaat aggtatthttg atatcattth ggataaatta tggcactcaa    540
tttattggag gaactaaatg tgctcctgac caagattaca aaggcgatac tttcgacca    600
tacattgatg ttcctcaagg tggttgttac ggtcaaaaag atgcttctcg gagaattcca    660
tttgggttac aaattgcacc agctttcatt ttaggtattg gtatgtcttt tttccaaga    720
tcgccaagat ggctcttctc aagaaaaaga gaagaggagg catgggaggc tttgaattat    780
ttgaggagaa gaaataatcc tgatatgatt gatgctgaat tcaatgaaat taaatcggac    840
gtattatttg aacaaaaata taacgagagg aaatttcaag gaaaaacagg aatgtctttc    900
tttataacat cgtattggga tctagtctct accaaatcaa attttaagag agtttttata    960
ggctctcgcg ttatgttctt ccaacaatth attggttgta atgcaattat ttattacgca    1020
ccaacaatat ttagtcaatt aggaatggat tccaacacca cagcattggt aggaacaggg    1080
gtctatggga ttgttaattg cctttcaact attcctgcta tctttgcgat tgacagattt    1140
ggtagaaaaa ctttattgat ggctggggca gctggaactt ttgtttcggt ggttatagta    1200
ggtgcaattg tcggcacata tggtgatact ttatctaagc ataaaactgc tggtagagcg    1260
gcaatagctt ttatttttat atatgatttc aacttttcat atagctgggc acctattgga    1320
tgggttttac catccgaat ctctctctatt ggaatcagat caaaagetat atccattact    1380
acttcatcaa catggatgaa taatttcata attggtttgg tcaccccccg tatggtggag    1440
acaatgaaat ggggaacata tatctttttt gcagcatttg ctataattgc atttgctttc    1500
acttggtttg ttatcccaga aactaaagga gtaccttag aagaaatgga tttagttttt    1560
ggtgatttag atcggttga agaaaagcaa aacttctctc gcatgaaatga attatcaaaa    1620
atggattcaa ttaaagctac aacagatatt tcggaagcac attattctga ttaa    1674

```

<210> SEQ ID NO 16
 <211> LENGTH: 1476
 <212> TYPE: DNA
 <213> ORGANISM: Escherichia coli

<400> SEQUENCE: 16

```

atgaataccc agtataattc cagttatata ttttcgatta ccttagtcgc tacattaggt    60
ggtttattat ttggctacga caccgcggtt atttcoggta ctgttgagtc actcaatacc    120
gtctttgttg ctccacaaaa cttaagtga tccgctgcca actccctggt agggttttgc    180

```

-continued

gtggccagcg ctctgattgg ttgcatcacc ggcggtgccc tgggtgggta ttgcagtaac	240
cgcttcgggc gtcgtgattc acttaagatt gctgctgtcc tgttttttat ttctgggtga	300
ggttctgcct ggccagaact tggttttacc tctataaacc cggacaacac tgtgcctggt	360
tatctggcag gttatgtccc ggaatttgtt atttatcgca ttattggcgg tattggcggt	420
ggtttagcct caatgctctc gccaatgtat attgcggaac tggctccagc tcatattcgc	480
gggaaactgg tctcttttaa ccagtttgcg attattttcg ggcaactttt agtttactgc	540
gtaaaactatt ttattgcocg ttccgggtgat gccagctggc tgaatactga cggtggcggt	600
tatatgtttg cctcggaatg tatccctgca ctgctgttct taatgctgct gtataccgtg	660
ccagaaagtc ctcgctggct gatgtcgcgc ggcaagcaag aacagggcga aggtatcctg	720
cgcaaaatta tgggcaacac gcttgcaact caggcagtac aggaaattaa aactccctg	780
gatcatggcc gcaaaacogg tggctcgtctg ctgatgtttg gcgtgggctg gattgtaac	840
ggcgtaatgc tctccatctt ccagcaattt gtcggcatca atgtgggtgct gtactacgcg	900
ccggaagtgt tcaaaacgct gggggccagc acggatatcg cgctggtgca gaccattatt	960
gtcggagtta tcaacctcac cttcacogtt ctggcaatta tgaagggtga taaatttgggt	1020
cgtaagccac tgcaaatat cggcgcactc ggaatggcaa tgggtatggt tagcctcgggt	1080
accgcgtttt aactcaggc accgggtatt gtggcgctac tgtcgtatgct gttctatggt	1140
gccccttttg ccatgtctcg gggcccggtg tgcctgggtc tgcctgctgga aatcttccc	1200
aatgctatct gtggtaaagc gctggcaatc gcggtggcgg cccagtggtg ggcgaactac	1260
ttcgtctcct ggaccttccc gatgatggac aaaaactcct ggctggtggc ccatttccac	1320
aacggtttct cctactggat ttacggttgt atgggcgctc tggcagcact gtttatgtgg	1380
aaatttgtcc cggaaaccaa aggtaaaacc cttgaggagc tggaaagcgt ctgggaaccg	1440
gaaacgaaga aacacaaca aactgctacg ctgtaa	1476

<210> SEQ ID NO 17

<211> LENGTH: 1701

<212> TYPE: DNA

<213> ORGANISM: Scheffersomyces stipitidis

<400> SEQUENCE: 17

atgcacggtg gtggtgacgg taacgatatc acagaaatta ttgcagccag acgtctccag	60
atcgtcggta agtctggtgt ggctggttta gtcgcaact caagatcttt ctccatcgca	120
gtctttgcat ctcttggtgg attggtctac ggttacaatc aaggtatggt cgggtcaaatt	180
tccggtatgt actcattctc caaagctatt ggtgttgaag agattcaaga caatcctact	240
ttgcaagggt tgttgacttc tattcttgaa cttggtgcct gggttggtgt cttgatgaac	300
ggttacattg ctgatagatt gggctgtaag aagtcagttg ttgctggtgt tttctcttc	360
ttcatcggtg tcattgtaca agctgttgcg cgtggtggta actacgacta catcttaggt	420
ggtagatttg tcgtcgggat tgggtgggtt attctttcta tggttgtgcc attgtacaat	480
gctgaagttt ctcccaccaga aattcgtggt tctttggtg ctttgcaaca attggctatt	540
actttcggta ttatgatttc ttactggatt acctacgcta ccaactacat tgggtggtact	600
ggctctggtc aaagtaaagc ttcttggttg gttcctatct gtatccaatt ggttccagct	660
ttgctcttgg gtgttggtat cttcttcatg cctgagctcc caagatgggt gatgaacgaa	720
gacagagaag acgaaatggtt gtcggtctct tccaacttgc gttccttgag taaggaagat	780
actcttgctc aaatggaatt ccttgaatg aaggcacaaa agttggtcga aagagaactt	840

-continued

```

tctgcaaagt acttcectca cctccaagac ggttctgcca agagcaactt cttgattggt    900
ttcaaccaat acaagtccat gattactcac taccacaact tcaagcgtgt tgcagttgcc    960
tgtttaatta tgacctoca acaatggact ggtgttaact tcatcttgta ctatgctcca    1020
ttcatcttca gttctttagg tttgtctgga aacaccattt ctctttttagc ttctgggtgt    1080
gtcggtatcg tcatgttctc tgctaccatt ccagctgttc tttgggtcga cagacttgggt    1140
agaaagccag ttttgatttc cggtgccatt atcatgggta tttgtcactt tgttgtggct    1200
gcaatcttag gtcagttcgg tggtaacttt gtcaaccact ccggtgctgg ttgggttget    1260
gttgtcttcc tttggatttt cgctatcggg ttccggttact cttgggggtcc atgtgcttgg    1320
gtccttgttg ccgaagtctt cccattgggt ttgctgtgta aggggtgttc tatcgggtgcc    1380
tcttctaact ggttgaacaa ctctcgtgtc gccatgteta ccccagattt tgttgetaag    1440
gctaagttcg gtgcttaacat tttcttaggt ttgatgtgta ttttcgggtc cgcatacgtt    1500
caattcttct gtcagaaac taagggtcgt accttgaag aaattgatga acttttcgggt    1560
gacacctctg gtacttccaa gatggaaaag gaaatccatg agcaaaagct taaggaagtt    1620
ggtttcttc aattgctcgg tgaagaaaat gcttctgaat ccgaaaacag caaggctgat    1680
gtctaccacg ttgaaaaata a                                         1701

```

<210> SEQ ID NO 18

<211> LENGTH: 1404

<212> TYPE: DNA

<213> ORGANISM: *Scheffersomyces stipitidis*

<400> SEQUENCE: 18

```

atgaagtatt ttcaaactctg gaaatcaggc aaacaagtaa gctacgctgt tacattcact    60
tgtgaattgg catttattct ttttggtatt gaacagggta ttattggtaa tcttattaac    120
aaccaggact tcctaaacac ttttgaaac cccaccggta gttatttagg tattatcgtt    180
tctatctata ccttaggggtg ttttttgggt tgtgttatga acttcttcat tggatgatcga    240
atgggcagaa gaagcaaaat tgcttctca atgacagtta tcacaattgg tgttgetctt    300
caatgtagtt ccttttcagt tgaacaattg atgattggaa gatttatcac tgggcttggg    360
actggttggg aaacttctac ttgtccaatg tatcaggcag aactttcacc tccaaaagtt    420
agaggacggt tgggtgtctc agaagcattg tttgttgag ttggtttaat ctatgcatac    480
tggtttgatt atgctcttct tttcaattct ggtcctattg catgggagact tctcttggc    540
tctcagattg tgttcgcctt tgttgttttc tgtttcactt tcacaatacc cgaatcccct    600
agatacatgt tttacaaagg agagaaagaa gaagccaaaa gaattttatc ttatgtcttt    660
ggaaagccag gagatcatcc tgacattctt aagggaatgga atgatattaa tgatgctggt    720
atthtggaaa cttcagaagg agctttctcg tgggcaaac ttttcaagcc cgataaggca    780
agaactggat acagagctctt cttggcatac atgagcatgt ttgcgcaaca gttgagtggg    840
gttaatgtag ttaactacta tattacattt gttttgatta acagtgttgg catogaagac    900
aacttggccc taattcttgg tgggttggcc gtcactctgtt tcaactgttgg ttcattagtt    960
cctactttct ttgctgatag gatgggaaga agattgcctt cagcagttgg agcttttggc    1020
tgtggtgttt gtatgatgct aatttcaatc ttattaagtt ttcaagacaa tccaaagttg    1080
aagaagagca gtggagctgg tgctgtggct ttctttttcg ttttcaact tgtcttggc    1140
tccactggta attgtattcc atggctgatg atttcaagac ttatcccctc tcatgcacgt    1200

```

-continued

gctaaaggat cttcattagc tacatcaagt aactggcttt ggaatttctt tgttggtgag	1260
atcactccaa ctatcattga aaagttgaag tggaaagcat atttgatctt tatgtgctgc	1320
aacttctcct tcgtaccaat gttttacttt ttctttcccg agacaaagaa ccttacttta	1380
gaagccattg acgatttggt ctca	1404

<210> SEQ ID NO 19

<211> LENGTH: 1656

<212> TYPE: DNA

<213> ORGANISM: Scheffersomyces stipitidis

<400> SEQUENCE: 19

atgagagaag ttggtattct tgatgttgc catggcaacg ttgtaactat aatgatgaaa	60
gatccagtag tatttttggg gattttattt gcacccttg gaggtttgct ttttggttat	120
gatcaagggg ttattagtgg cattgtcaca atggaatctt ttggtgcaaa attcccaga	180
atttttatgg atgccgatta caagggttgg tttgtgtcta cttttttgct atgcgcatgg	240
tttggtccta ttattaatac tccaattggt gataggtttg gaagacgtga ttctatcaca	300
atctcttgty ttatttttgt cattggttct gcgttccaat gtgctggcat taatacaagt	360
atggtatttg gtggcgctgc tgttctggtt cttgcagtcg gtcaattaac catggtagtt	420
ccaatgtaca tgcggaatt ggctcctcca tcggtgagag gtgggttggg tgaattcag	480
caactttcga ttacaattgg tatcatgatt tcttattggt tggattatgg cactcatttt	540
attggaggta ctagatgtgc tctagtcac ccataccaag gtgaaacttt taaccctaat	600
gtggatgttc ctccaggtgg ctgctatggt caaagtgatg ccagttggag aattcctttt	660
ggtgttcaga ttgctccagc agtggttggg ggtattggaa tgatattttt cccaagatct	720
cccagatggt tactctctaa aggtcgcgac gaagaagctt ggagctcttt gaaatatctc	780
agaagaaaga gtcagtagga tcaagtcgaa agagagtttg ctgaaattaa ggcagaggtc	840
gtttatgaag acaagtacaa ggaaaagaga ttccctggta agactggagt tgctttaaca	900
cttactggat actgggatat tcttactact aaatctcact tcaagagagt ttttattgga	960
tcagctgtca tgttcttcca acaattcatt ggctgcaatg caataattta ttacgcacct	1020
acaattttca cacaattggg aatgaactct acaactactt ccttgcttgg tactggtctt	1080
tatggtattg ttaattgtct ttcccacctt ccagcagtggt tcttgatcga tagatgtgga	1140
agaaagactt tgtaaatggc aggtgctatt ggaactttta tttccttggg tattgtcggc	1200
gcaatcgttg gcaagtatgg cgatcgttta tctgaattca agacagcagg gagaactgca	1260
attgctttca ttttcattta tgatgtgaat ttctcgtaca gttgggctcc aattggatgg	1320
gttttaccct cagagatttt cccaatcggc atcagatcca atgccatctc cataactacc	1380
tcacttactt ggatgaataa ttttattatt ggcttggcca ctccacatat gttagaacaa	1440
atgaagtggg gcaacttaac tttttttgca gcgtttgcta ttattgcggt ctttttcaact	1500
tggttatca tcccggaaac caaggaggtt ccattggaag aatggatgc cgtggttggc	1560
gatactgcag cattgcagga aaagaatttg gttaccatta cgtcagtttc tgaatctgac	1620
gccaaggatc gcaactcgat tgaaatgtca gaataa	1656

<210> SEQ ID NO 20

<211> LENGTH: 1629

<212> TYPE: DNA

<213> ORGANISM: Scheffersomyces stipitidis

<400> SEQUENCE: 20

-continued

atgtcttcgt tattgactaa cgaatacttc aaagactact accacaaacc gactcctggt	60
gaagtgggta ctatgattgc tatcttagag atcggcgcac tttttcctc ctccatagct	120
ggaagagtag gtgacatcgt tggcagaaga agaaccatta gatacgggtc tttcattttt	180
gtagtaggcg gtcttgtaga agctacttcg gtcaatattg tcaatctctc actaggaaga	240
ttgattgcog gtattgcoat tggctttttg acaaccatca tcccattgcta ccagtctgaa	300
atcagccccc cagacgatag aggtttctat gcctgtttgg agttcaccgg aaatatcatt	360
ggatatgcta gtagtatttg ggtagactac gggttttcat ttttagacaa tgatttcagc	420
tggaggagcc cattgtatgt tcaggttggt attggctcca tghtatttat tggttcattc	480
cttattgtag aaaccctag atggctcttg gatcacaacc atgatatcga aggcatgatt	540
gtcattttctg acttgtatgc agatgggatg gtggaagacg atgatgctat tgctgagtac	600
agaaacataa aggaaagtgt cttgatagcc agagtgaag gcggagagag atcgtaccag	660
tatttggtca ccaaatatac caagagactt tctgtggcat gcttttcgca aatgtttgcc	720
cagatgaatg gtataaacat ggtatcttac tatgctccta tgatcttoga atctgctggc	780
tgggttggtg gacaagctat cttgatgact ggtatcaact ccattatcta catctttagt	840
accattcctc catggactt agttgattct tggggcagaa aacctttgct tttatctgga	900
tctgtgctca tgggtgttcc gctcttaacc attgcttggc cgttattctt aaacaacaca	960
tacacaccog gggttgtggt tggcagtgta atcgtattca atgctgcttt tggatacagt	1020
tgggggtccaa ttccttggtc catgagcgaat gtgttcoccta actcagttag atcaaaagg	1080
gctgccatgt ctactgcaac caactggctc tttaacttta ttggtggaga gatgacacct	1140
atcttgttgg atacaattac ctggagaact tacttgatcc cggcaacttc gtgtgtatta	1200
tcgttttttg ctggttgatt tttatttcca gagaccaagg gtttagcatt ggaggatatg	1260
ggctccgcat tcgatgataa ttctgcaata ttttcataac actcaactcc ttcactggg	1320
tatggtgcga ccgagtctaa cagtaatgcc aggagagcaa gtgtcatctc ttcagaaaac	1380
taccaggata gtttgcatca gacagcggct tcattggcta gaaatcctc aagcatgagg	1440
cctgattaag atggcataat cacaggagct gctacccttt cggcagtacc accattaaaa	1500
ccaataaagt ctgatgcgtc agtccattca gtcgatgcca taattccaag catttccagc	1560
aatattccgc aggaaattga accaccaacc tttgatgaaa tctttaagta caagttgaat	1620
gagatggaa	1629

<210> SEQ ID NO 21

<211> LENGTH: 1488

<212> TYPE: DNA

<213> ORGANISM: Scheffersomyces stipitidis

<400> SEQUENCE: 21

atgacggaaa gaagcattgg acctttaatc cccagaaata agcacttatt ctatggatcc	60
gtattattga tgagtattgt tcaccaactc atcatgggat acgattccat gatggttgg	120
agtattctta atctagatgc atatgtaaat tatttccact taacggctgc taccactgga	180
ctcaataactg ctgacgatg gcttgggcaa gtaattgcca cattgacagt tattctgtat	240
ttcaatgaca aatttggttag aagaagtcca gtttgataa gtattgcaat cagtttgggt	300
ggggttgcac tgcaatcagc agcccaaac attgagatgt ttattatcgg aagaatagtt	360
attggttttg gaatatctat tgggtttgtc tcatctacca ttttgtaag tgaactagcc	420

-continued

cctccagaca aaagaggatt tattcttga ttgagtttta caagctttct agtaggaagt	480
ttaattgcag cagggtgcac atatggaaca agaaatgctc ctggagactg gtgttgaga	540
atcccatcaa ttattcaagg ggctccagat attgttgcta ttattaacat actctttatt	600
tcagaatcac caaggtgggt gattgcaaag gaaagattca gcgaagctcg tgaattatt	660
tctatcatta gtgatgttcc tattgaagat gcacatgaag aatgtgaaa gatacatgcc	720
catattcaaa ctgagaagac tgctttccct ggcaataagt ggaaacaaat ggtgagctcc	780
aagagcaata caagaagagt tattatcttg ttcacacagg ccatagtac tgaatggcc	840
ggttcttcag ttggatcgta ctatctttca attatattaa ctcaagctgg ggtcaaagat	900
tcgaatgata gactaagagt aaatattgtg atgagttcgt ggtcattggt aattgctctt	960
tccggatgtc taatgtttga cagaattgga agaaaaatgc aatcgctcat ttcgttatca	1020
ggtatgatca tatgctttat agtttttaggt gttttggta aagaatatgg cgatggctcat	1080
agcaagagcg gaagttagc agctgtgcc atgatgtttt tattcacagg attttactca	1140
ttcactttca ctccattgaa ctctttgtac cctccagaat tgttcccta cgtggtgaga	1200
agtacaggag ttacactctt taatattttc aacggctgct ggggactttt cgcaagtttc	1260
attttacca ttgcaatgaa tggaaattggc tggaaatttt acatcattaa tgettgtat	1320
gacgtcatat tccttccaat aataatgttc tgttgattg agacaaaggg aattaattg	1380
gatacaatta gtgaagtatt gcacggaaga ggacctgaag atgaagaaag cattgaagaa	1440
agtcacagcc taatcagaca aggttttgtt gttaatacaa agaagtaa	1488

<210> SEQ ID NO 22

<211> LENGTH: 1635

<212> TYPE: DNA

<213> ORGANISM: Scheffersomyces stipitidis

<400> SEQUENCE: 22

atgtccagtg ttgaaaaag tgctgaaact gcttctata cgtcgcaggc cagcgcaagc	60
ggctctgcaa agaccaacag ctaccttggc ctccagaggcc acaaacttaa ttttgetgct	120
tcttgttttg ctgggtgttg tttcttactt ttcggttacg atcaaggtgt catgggttca	180
ttggtgacct tgccatcctt cgaaaacact ttcccggcca tgaaggctag caacaacgct	240
accttacaag gcgccgttat tgcactttat gaaatcggtt gtatgtcttc ttctttagca	300
accatttacc ttggtgacag attgggtaga ttgaagatca tgtttattgg ctgtgtaatt	360
gtctgtattg gtgctgcttt gcaagcttct gctttcacta ttgctcactt gactgttget	420
agaattatca ctggtttagg tacaggtttc atcacttcta ctgttccagt ttaccaatcg	480
gagtgctctc cagccaagaa aagaggacag ttgatcatga tggaaaggttc tcttatgcc	540
cttggcattg ccactctcata ctggattgac tttggatttt actttttgag aaacgatggt	600
ttgcaactct cggcttcttg gagagcacct atcgcgcttc aatgtgtctt cgctgtcttg	660
ttgatttcca cagtcttctt ctcccagaa tctccaagat ggttgetcaa caaaggtagg	720
accgaagaag ctagagaagt tttttctgct ctttacgact tgccagccga ctctgaaaag	780
atttctattc aaattgaaga aattcaagct gctatagatt tagaaagaca agcgggagaa	840
ggtttcgtac ttaaggaatt gttcactcag ggcccagcca gaaacttgca gcgtgtggcc	900
ttgtcatggt ggtctcaaat aatgcaacaa atcaactggta ttaacattat tacgtactat	960
gctggtaacta tttttgaatc atacattgggt atgagtccat ttatgtcaag aatcttggct	1020
gccttgaagc gtactgaata tttccttctc tctcttattg ctttctacac cgtcgaagaa	1080

-continued

```

ttaggtagaa gattcctttt gttctggggt gccatcgcca tggetcttgt catggctggt 1140
ttaactgtta ccgtaaact tgccggtgaa ggcaacaccc atgctggtgt cgggtgctgt 1200
gttcttttgt ttgcatcaa ctcattcttc ggcgtctcct ggtaggtgg atcctggttg 1260
ttaccacctg aattggtgtc ttgaaattg agagctctg gtgctgcttt gtegaccgt 1320
tctaactggg ctttaactt catggtgtc atgatcactc ctgtcggttt ccaaagtatt 1380
ggttcctaca cctaccttat ctttgctgcc atcaatttgt tgatggctcc ggcatctac 1440
ttcttgatc ccgaaccaa gggtagatcg ttggaagaaa tggatatcat tttcaaccaa 1500
tgtctgttt gggagccatg gaaggtgtc caaattgcca gagacctccc tattatgcac 1560
tcagaagttc ttgaccacga aaaggatgtc attattgaaa aatctagaat agagcatgtc 1620
gaaaacatca gctaa 1635

```

```

<210> SEQ ID NO 23
<211> LENGTH: 1251
<212> TYPE: DNA
<213> ORGANISM: Scheffersomyces stipitidis

```

```

<400> SEQUENCE: 23

```

```

acttttgcag ttaacttgta tgtgtttgca gttggtagag tgctttctgg ggtgggtgta 60
ggagtctat cgactatggt gccgtctat caatgcgaaa ttagtcccag cgaagaaaga 120
ggcaagtgg tgtgtggaga gttcacggga aatatactg gttatgctct cagtgtatgg 180
gccgattact tctgctactt tattcaagat ataggtgatg caagggagaa gcctcatagc 240
ttctttgccc acttgctctg gcgattgcct ctattcatcc aggtgggtgat agcggctggt 300
ctctttgttg ggggattttt tattgtogag tcacctcgtt ggttattaga tgtagaccag 360
gaccaacaag gattccatgt attagcgttg ctctatgatt cacatctaga tgataacaaa 420
ccacgtgaag agttctttat gatcaaaaac tccatcttgt tagaaagaga aactacacct 480
aagagcgaac gaacttgga acatatgttc aagaactaca tgacccgagt gcttatagct 540
tgttcagcac ttggctttgc acagttcaac ggcataaata tcatttcgta ctatgcccc 600
atggtatttg aagaagcagg cttcaacaac tccaaggctt tacttatgac aggcataaac 660
tctatagtat attggttcag tacgattcct ccgtggtttc tcgtggatca ttggggtaga 720
aagccaatth tgatatacgg gggtttatct atgggaatat gtattggttt gattgcggtg 780
gtaattctac tagacaagtc gttcacaccg tctatggttg cggattgggt gataatctac 840
aatgcatctt ttggctacag ttgggtcctc atcggattct tgatcccgcc ggaggtgatg 900
ccattggcag ttgatcgaa aggtgtttct atttctacgg ctacaaaactg gtttgccaat 960
tttgttggg gtcagatgac gccaatctc cagcagagat tgggctgggg aacttatcta 1020
ttcccggctg gtagttgtat catctcgggt atagtggta tttcttcta tccagagaca 1080
aaggtgacag agctagagga tatggactct gtgttcgaga gcttttcaa ctacaagtct 1140
ccgttcaaga tttcacgaaa gagacaccag aatgatggcc aggcgtacca aaggttagag 1200
aacgatatcc gccacaacga ttagaagtg gacgatttgg acgatttga c 1251

```

```

<210> SEQ ID NO 24
<211> LENGTH: 1638
<212> TYPE: DNA
<213> ORGANISM: Yarrowia lipolytica

```

```

<400> SEQUENCE: 24

```

-continued

atgattggaa acgctcaaat taaccaggtg ggagccttac agcaccgggt ccccaaaactc	60
cacaatcctt acttaactgc ggcgtggcc accatgggtg gcctgctttt cggctttgat	120
atctcgtctg tttctgcctt tgcgacacc aaacctaca aggagtactt tgggtacccc	180
acctccatcc agcagggcgg tatcaactgcc tcaatggcgg gtggatcctt cctgctgtct	240
ctggtggcgg gctggatttc cgaccgactt ggtcgacggt tcgccatcca ctttgccttc	300
ttttggggg tggttggagc tgcctccag tctcagccc aaaacaaggg ccaattgatc	360
gccggtcgac tcatttcogg ccttggatc ggtctgggct cctcggttat ccccgctac	420
atctccgagc tgtctcccaa gaagattcga ggtcggcttg tcggtctctt ccaatgggcc	480
gttaacctgg gtatcctcat catgtctac atttctctcg gtctcagtaa catccacgga	540
gtcggcgat tcagagtcgc ctggggtctg cagatcatcc ccggtctgct catgtctctc	600
ggtgtttgt tcctggaaga gtctccgca tggctagcca agcaggacaa ctgggacgag	660
tccgtgcgag tgcttcgagc catccaccag ggaggctacg gcaccgaaga agacattttg	720
ttagagattg aagagatccg agaagcagtc cgaatcgagc atgagaccaa aaacctgcca	780
ttctggcacc tgttccaaaa ggactctatc aaccgaacaa tgggtgggtat ctgggccag	840
atctggcagc agctcaccgg catgaacgtc atgatgtatt acattgtgct gattttcacc	900
atggctggat aactggaaa cgccaatctg gtggcctcgt ctatccagta cgtcatcaac	960
atgatcatga ccacccccg tcttctgttc attgaccgag tgggacgacg acccctgctg	1020
ctgttcggat caatcgtcat gatgatctgg ctgtttgccc tcgctggtat cettgcagtg	1080
tacggaactc agatccccgg tggactcagc ggagacgcat tcacaacct tgcattgag	1140
cccactcaca agcctgccc aaagggagtc attgcgtgct cgtacctgtt tgtggccacc	1200
tttgcccta cctggggccc cggtatctgg ctgtactgct ccgagctgtt cctctgaag	1260
cagcgagctg tggctgcogg tgaaccgcc tctgccaaact ggatcttcaa ctttgcctc	1320
gctctattcg tgcctcggc ctcaagaac atcaactgga agacctacat catctttgga	1380
gtctctgta tcgtcatgac catccacgtc tttgtctct tccccgaaac caagggcaag	1440
accctcgaag agattgatat gatgtgggcc gcccgagttc ctgcctggag aaccgcaaac	1500
tgggtgctg accacgttc tggcgcctt cccgaagacg agaaacactc ggaggagatg	1560
gtcagggcgg tcgaatccaa tgaagaggag cccaagatag ccagtgctaa cgtcgacgcc	1620
cctccctctc aattgtaa	1638

<210> SEQ ID NO 25

<211> LENGTH: 1665

<212> TYPE: DNA

<213> ORGANISM: Yarrowia lipolytica

<400> SEQUENCE: 25

atgtacaagg tccataaacc ctacctcag cgggcggtag ccaccatggg cggaatgctc	60
tttggtttgg atatctcgtc cgtgtcggcc tttgtgggag aagataacta catgaactac	120
tttggtcac ccacctcctt ccagcaggga ggtatcaccg cctccatggc cggaggatcc	180
atgctgtcgt gtgcgtttgc cggctacatt tccgaccggg ttggccgaaa gccaccatt	240
caatttgccg ccgcctgggt gatggttggc gctccattc aatgctctgc ccagaatatg	300
ggccaactga ttgcggcgg gccatttcc gggcttggaa tcggcctcgg ctctgccag	360
atccccgtct tcactccga gttgtcccc aagaagatcc gaggcgggt cgtcggctgc	420
ttccaatggt ccgttacctg ggtattctc atcatgtct acatttcggt cggctgctca	480

-continued

tacatcaagg gccactcgtc cttccgactg gcgtggggca tccagctgat tccaggagcc	540
atggtggcgt tcggaatgat gctgctggac gaatcgccgc gatggctggc gtccaaagac	600
cgctgggaag aggccatcca gatcatccgc tccatcaatg ccaactacgg atccgaggag	660
gacattctca tggaaatoga agatctgcga gaggtgggtc gaatcgacca cgagtccaag	720
tcggtcacca tctgggaact gttccgaaaa gactccatca accgaacct ggtcggagtg	780
tgggcccaga tctggcaaca actgactgga atgaacatca tgatgtacta cgtggtcac	840
atcttcaaaa tggccggcta ctccggcaaaa agtgcctca ttgtctccgg ctccatccag	900
tacatcatca acgtggteat gaccatcccg gcgctgcttt tcattgataa aatcggacga	960
agaccctcgc tcctctgtgg aagcatgctc atggccacct ggctgctagc tgcggaggga	1020
atgctaggag cctacggaat ccaaatgccc caaggtctac cggcagtacc ctccaaaaac	1080
caggcagcgg acccctacac caccatctac attcccgaca accaggcgcg gcccccgaag	1140
gccattatcg cctgttgcta tctgtttgtc gcctcttttg caccacactg gggccccggc	1200
atctggctct actgctccga gatcttcccc aacaaacaac gagecgtggc caactcgtg	1260
accgccggcg ccaactgggg cttcaacttt gcctggccc tgtttgtgcc caccgccttc	1320
aagaacatca actggaaggt gtacatcacc ttcggtgtct tctgcatcgt catgtccatc	1380
cacgtcttcc tgctcttccc cgaaccaaac ggcaagatc tggaggtgat tgaccagatg	1440
tgggacgccc gcgtgcccgc ttgaaaaacc gcctcctggg tcccgaacca catgccttct	1500
cattacgcag gggaccagga gaaaaagccc accgacgaac tggccgaggc gccgtttcac	1560
gaggagaatg ccccggtgaa caccgagacc cctcctcatg aggatgagcc cacttttgcg	1620
gagaccgagc ccaagaccca gtatcctgga actgagcatg tctaa	1665

<210> SEQ ID NO 26

<211> LENGTH: 1599

<212> TYPE: DNA

<213> ORGANISM: Yarrowia lipolytica

<400> SEQUENCE: 26

atgttttcgt taacgggcaa accgctgctc tattttacgt cagtgttcgt ctctctgggc	60
gtgttcctgt ttggatacga ccagggagtc atgtccggca tcatcacggg cttctacttc	120
aaggagtact tccatgagcc caccgagcc gaaatcgaa ccatggtgtc gattcttgag	180
gtcggagcgt ttgtctcgtc gctcatggtc ggccgaatcg gtgacattat tggccgacga	240
aaaaccatca tgtacggtgc cttcatcttc atcatcggag gtgccttcca gacatttgcc	300
gtcagcatgt ccgagatgat tttgggccga gtagtggccg gtttcggcgt tggtagctg	360
tcgaccattg tgccagtcta tcagtctgag atctcgcctc ctcaacaacc aggcaagctc	420
gcgtgcatcg agttcacggg aaacattgtg ggctatgcca gctcagtgtg ggtggactac	480
ttctgcagtt tcatcaatc caacatgagc tggcgtatcc cactgtttct gcagtgtgct	540
atgggcgctc tttgtttgg agctcgttc ctaattgccc agactcctcg atggttgctg	600
gataacgacc atgacgagga gggattggtt gtccctggcca acctgcatgg aggaggagac	660
attgactctc ctctggctaa gcaggaatat cgggagatta agcagtcctg tttgatccac	720
cggctcgagg gcgagcgatc atataccgac atgtggaaga agtacaagaa gcgagtgctg	780
attgccatgt cgtegcagat gtttgcccag ctcaacggta tcaacgcat ctcttactac	840
gctcctctgg tgtttgaaga ggcaggatgg gtcggacggt ctgctattct catgaccggt	900

-continued

```

atcaacggta tcgtctacgt gtgttccact attccccctg ggtacctcgt ggacaaatgg 960
ggccgaagac ctattcttct gtcoggtgca gtaattatgg ctatttcctt ggcgtctgtg 1020
gcgttctgga tgcgtctaga ctttgacat acaccggctc tgggtggtgat ttcogtctc 1080
atcttcaacg ctgcttttgg atactcgtgg ggcacctc cctggctcta tccccctgag 1140
attatgcctc ttaccatccg agccaaggga gcttctctgt caaccgccac caactgggcc 1200
ttaaactggc tgggtgggata tatgaccccc attctccagg agaccatcaa gtggcgactg 1260
tatttgatgc atgccgcctt ctgtagtctt tcgtttgttc tcgtctactt cacctaccg 1320
gagacctcgg gaatcaactt ggaagacatg gactcgttgt tcggcgacaa gtctgttgtg 1380
aacacccccg actcgcggtc tttgcttggg gatcgagaca ctccagagcc tgacgtgcct 1440
cacagttata ctgatctgc caccgatcga ctgcctgctg gtatgcaggg ctatggctcc 1500
gtccccagct cgagaggagg cagtgtggtc ggaagtcccc gacgaggaaa cagtgtggtt 1560
gggtctccca agcgggactt cctcaacct ccggtataa 1599

```

<210> SEQ ID NO 27

<211> LENGTH: 1548

<212> TYPE: DNA

<213> ORGANISM: Yarrowia lipolytica

<400> SEQUENCE: 27

```

atgggactcg ctaacatcat caaccgtgga gaaaagcccc agggctcggc ctcatggcg 60
gcctttgtgg ccgtgtttgt cgcgtttgga ggtattctgt ttggatacga cactggaacc 120
atctccggcg tcattggccat gccattctgc aagaagacct ttacagatga cggcctggag 180
ttcacttctg agcagacctc gctcatcaact tccattcttt ctgcaggcac ctcaactgga 240
gccatttctg ctccctgggc ctctgatact ctgggaagac gactgggtct gatcctcttc 300
tgtgtcgtct tctctgttgg cgctattctt cagactgctg ccaccggccg aacgcttttg 360
attgtcggac gagttgttgc tggctcttgg gttggtggag tctcttccat cgttctctt 420
taccagtctg aggttgcccc caagtggatc cgaggtgccc ttgtctccat ctaccagttt 480
gccatcacca ttggtcttct gctggctgcc attgtcaaca acgcaaccaa aaacaaagac 540
aacagtctt cctaccgaat tctctctggc cttcagcttc tgtgggcccgt catcctgagt 600
ggaggtctca tcctgctacc ggagactcct cgattctgga tcaagaaggg cgagtacgac 660
aaggccgcg attccctgcg acgactacga cgacttctg ttgagcacga ggctgtacag 720
aaggagctcc tggagatcca atcttctcac gaccacgaga tgcagatcgg tagcgcacc 780
tgggcccgcct gcttctcccc caaggggtcc cagctgaagc gaatgctgac cggatttgcc 840
attcaggccc tgcagcagct caccggtatc aacttcatct tctactacgg aaccgagttc 900
ttcaagaagt ccaacatctc caacccttc ctcatccaga tgatcaccaa cattgtcaac 960
gtggttatga ccatccccgg tatcatgtt gttgatcgag tcggacgacg aaagctgctg 1020
ttgatcggag ctatcgtcat gtgctcttcc gagtttatcg tggcggctgt tggtaactgcc 1080
attgataacg agacctctc aaaggttctg attgcctca cttgtacctt cattgcccgt 1140
tccgcccaca cctgggggtcc tattgctgg gttgtcattg gagagatctt ccctctacga 1200
atccgagcca aggtgttgc tctatgcgcc gcttccaaact ggcttttcaa ctttgccatt 1260
gcctttgcaa cccctcaact cgtcgaagc gcccctggat cggcgggtct caagaccaag 1320
gtcttcttca tctggggagg ctgcaacttc ctgtgcatcg ccttcaacta cttcttctc 1380
tacgagacca aggttcttac tctggaggag gtggaccaga tgtacgccga gatcaagatt 1440

```

-continued

```

gettctcgat cccaccagtt tgtgectacc actcgagtcg ctgcttacga cgagcacgct 1500
tctgacgaca agaaggacgg acagcacgtc tacattgagt ctgtctag 1548

```

```

<210> SEQ ID NO 28
<211> LENGTH: 1485
<212> TYPE: DNA
<213> ORGANISM: Yarrowia lipolytica

```

```

<400> SEQUENCE: 28

```

```

atggccatta ttgtgctgt atttgggct tttggaggac ttctctacgg ctacgacaca 60
ggaactattg ctggaatcat gaccatgggc tatgtgaaag aacactttac agactttggg 120
aagaacgact tcacctggg ccaatcatct ctcaccacat ctatcctatc tgtgggcaca 180
tttaccggag ccctcgttc tcccttagct gctgacacgg ctggtcgacg tctgggtctt 240
ctggtgtatt gtcttatt ctctgtgggt gctatcttgc agaccgtcac aaccggaaga 300
gtcttgctaa ttgtgggacg ggtgattgct ggtcttggtg tgggaggat ctctgctt 360
gtgctctct atcagtcaga agtgtctccc aaatggatca gaggggccgt tgtttctgtc 420
taccagttg ccatcactgt ggtctctcta ctggcagcta ttgtcaacaa tgccactaag 480
gaccgtccaa atacgtcatc ataccgtatc cctcttgcta ttcaactcat ttgggctctt 540
attctttcag caggacttgt gtttcttct gagactctc gtttctgggt caagaagaac 600
cggccagaga aagccgccga agcactctca cggctaagaa ggctaccaac agactcgaaa 660
ccggtaaaaa aggaactgct tgaactacag aagtcgctcg aaatggaaat ggaggttggg 720
aactcctct ggaagcttg tttcagtcca catggatcac agctcaaaag actgctgaca 780
ggagtctcaa tccaggctct gcaacaacta acaggcatca atttcatatt ctactatgga 840
accaactttt tcaaacagc tggcataaaa gatccctttg tgggtgctcat gatcaacctc 900
gccgtcaatg tggccttcc cctccgggt attctgttg tcgacaaagt gggccgaaga 960
aagctgctct tgattggagc cgtggtaatg tgtgtttcag aattaattgt ggcagctgta 1020
ggagcagctc tggatagcca ggtgtcttcc aaggctctta ttgccttcc ttgtacgttc 1080
atcgaggct ttgcatccac ctggggacct atagcctggg tggttgttgc ggagattttc 1140
ccgcttcgaa tccgggccaa gggagtggct atcagtgtgg ctgccaactg gattttcaac 1200
tttgccattg cctttgcaac tccgtacttg gtcgacaaga aacctgggtc tgccggtctc 1260
gagtccaagg tgtttttcat ctggggagggt tgcaattttc tagccattgc ctttgtgtac 1320
ttgtttgtct atgaaaccaa agggttgtca ctggagcagg tggacgaaat gtactcggag 1380
gtcaagtacg cctggcagag tgataggttc cagaccgaga tcatgtctgg aaagcggag 1440
gtttcgccgg atcagagctg cgattctgga tttgattcgg attag 1485

```

```

<210> SEQ ID NO 29
<211> LENGTH: 2881
<212> TYPE: DNA
<213> ORGANISM: Cryptococcus neoformans

```

```

<400> SEQUENCE: 29

```

```

atgcccgatc cctctattcc ggtagtcggt cacaagactc agcgtaggct ggtcggacat 60
aacctactgt acagtgtttc agtgtttctt agcataggag tctggttatt tgggtagggt 120
ttttgtgaat ccggatgggt ttgtgctgat atgataaccg gttaagatat gaccaggggt 180
aagttcactt caagcctta tccttaaget tcgctgatca tggccttcaa actgtgaaag 240

```

-continued

agtaatgtcc	ggtaagttgg	tatcaccatc	acatatctcg	atagctagtt	gactttctac	300
gtaaggaatt	attaccggcc	catactttaa	gtaggtatga	agtaaact	gagtatcttt	360
cattaacatt	tatggcagct	gtctacagag	cttattgtga	gtaaagccct	aatcaagcaa	420
actaagctga	cgtttttagt	caaccaacca	acgtcaacgc	agattggcaa	gtaagttagt	480
cattcagtg	cagtagcaat	gtctgaacat	tcgatagtat	ggtggccgtt	ttggagattg	540
gtgccttcag	taagattatt	ccatgtcatg	tcattcttc	aagaactcac	ggcaatcacg	600
cagttacttc	tctggctgcc	gctcatattg	cagataatta	tggaaagcgt	atgacccttc	660
gcacaggtgc	aatagtcttc	accattggag	gtgctataca	gactttttgc	gttggatata	720
attccatggt	acttgggaaga	attgtcagcg	gctttgggg	agggatggtg	agtatggctg	780
tgccaatcta	tcaggtatgt	ggtttacaat	agtaaagcgc	cggagcataa	gccaatggt	840
gtcaccctcg	cgcagtcoga	aatatctcct	gcagaccatg	taagaaacac	ttcaatattt	900
ccagaagatt	tttctaataa	aacacgatga	tagcgaggcc	ttttgggctc	tgtcgaattc	960
acaggttaata	tcattggcta	tgccctcctc	gttgtacgat	gtgtcacctg	cctatgtcaa	1020
gcaccatgct	cactcaacgt	tcatagtga	tcgactatgc	ctgttcattc	ttccagtctg	1080
actggtcttg	gcgcctcccg	ctttctgttc	aatgtatagg	cggtctggt	ctcttcacg	1140
gcagcttcgt	cacaccagag	tctcccgggt	aagccttcta	tatgtgcac	atatgtaggt	1200
gcagaactaa	gagctgttca	aaggatctt	gtcgatacag	accaagaggt	ggaaggttta	1260
gcagtcacg	ctgattttca	agggaaagcg	ctggacgata	tttcagtgca	agccgagtac	1320
aaagaaattc	gagatgctgt	tctagccgac	gtgagacaat	cctctcaacg	ttatcccat	1380
acacttgctt	attctgtttt	gtttttttt	tactagagag	ctgtcggaga	tagaagctat	1440
agggctttat	ggaggagata	caaaggacga	gttctgattg	caatgagcag	tcaattgttt	1500
gctcaactgg	tgagtcaatc	tttgcaaaag	tcaaagaaac	atgaaaataa	agcggctctgt	1560
actaatgatt	attggcagaa	tggcatcaat	ggtgagctta	gagagcctg	caaacaatta	1620
ctaaccattt	ggccagtcac	ctcatattat	gcacgtgcgt	cccatctcat	tctgccgatc	1680
attcttgaca	gctgatgtgg	attacagctc	ttgtcttga	acgttagtct	cactgaagac	1740
ctaactactgc	cgtgttacta	atgggtggtga	agaggcgggg	tggattggac	gtgacgctat	1800
ccttatgaca	ggatcaaatg	ccttatttta	tgtggcaagc	tcacttccgc	cgtaagtcca	1860
ggtccatccg	aaattatgca	cattctcaaa	ttattactag	atggtatctc	atggatcgag	1920
cgggtcgaag	gcccattttg	ctctcgggag	cagtgcccat	ggcgattgca	ctgacggcta	1980
caggatggtg	gatatatatt	gatcaagcaa	taacacccaa	tgctggctcg	tcttttgctc	2040
tgccatgctg	gatgaagctg	atggatattg	tcgatagtgg	tcatttgcgt	agtgatttat	2100
aattccgcat	ttggcatgag	ctggggacct	gtcccatggt	atgtgtcatt	gacatatggc	2160
cggttggaag	tgaagttaat	taattaattg	atcccaggct	ttatcctccg	gaaatcatgc	2220
cgttgctcatt	ccgagcaaa	ggagtatcct	tatctactgc	tacagtagct	ccaattttat	2280
cccgaatgca	caacgatggg	ctaatttgag	ttatcagaac	tggatctcag	tgggtctgga	2340
gcgctttaat	cctgcttttt	gctaacgact	gttgtatctg	cagaattggt	gggtaggggt	2400
ttcaacacccg	ctctttcaag	aacttatcgg	atggcgatta	tatccgatgc	acgcattctt	2460
ttgtgcatta	tcattcatcc	tcgtgtactt	cgtgagttg	tcagccgaga	ttcgtaaacc	2520
accactcatg	cacatggaac	tagtctatcc	cgaaaccgga	ggcgtaccgc	ttgaagaaat	2580
ggacaattg	tttgggatg	aaagtgatga	agacgaggtt	gattcggact	tcgatgaagt	2640

-continued

tgaggaagcc gaatcagaaa taccctctct agtcagcaat cctcgacacc gacgcccctc	2700
ggccagctct tcattgggccc catctttgcc gacctcccga aaaccgtcac ccataccctc	2760
tagggaggct tcactagacc gaggactggt tggacgtata actgactcgg tgaatggtct	2820
gattggaagc acaaacacagc aaagcaggag cgtgggggtat actgctgtca acgaggaata	2880
g	2881

<210> SEQ ID NO 30

<211> LENGTH: 1713

<212> TYPE: DNA

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 30

atgtcacaag acgctgctat tgcagagcaa actcctgtgg agcatctctc tgctgttgac	60
tcagcctccc actcggtttt atctacacca tcaaaccaagg ctgaaagaga tgaataaaaa	120
gcttatgggtg aagggtgaaga gcacgaacct gtcggtgaaa ttccaaagag accagcttct	180
gcctatgtca ctgtctctat tatgtgtatc atgatcgccct ttggtgggtt cgttttcggg	240
tgggatactg gtaccatttc tggtttcatc aatcaaaccg atttcatcag aagatttggg	300
atgaagcata aagatggtac taattatttg tctaaggtta gaactgggtt gattgtctcc	360
attttcaaca ttggttgtgc cattggtggt attattcttt ccaaattggg tgatatgtac	420
ggtcgtaagg tgggtttgat tgcggtgtt gtcacttaca tcatcggtat tattattcaa	480
attgcatcta tcaacaaatg gtaccaatat ttcacggta gaattatttc cggtttgggt	540
gttggggta ttgcccgttt atctcctatg ttgatttctg aagtatcccc aaagcattta	600
aggggtactt tagtctcttg ctaccaattg atgattactg ccggtatttt cttgggttac	660
tgtaccaact tcggtactaa gaactactcc aactctgtgc aatggagagt tccattaggt	720
ttgtgttttg cctgggcttt gtttatgatt ggtggtatga catttgttcc agagtctcca	780
cgttatattg ctgaagtogg taagatcgaa gaagccaaac gttctattgc cgtttctaac	840
aaggttctg ttgatgatcc atctgttttg gctgaagtcc aagctgtctt ggctggtgta	900
gaggcagaga aattagctgg taatgcatcc tgggggtaat tgttttagtag caagacaaag	960
gtccttcagc gtttgatcat gggtgctatg attcaatctc tacaacaatt gacaggtgat	1020
aactatttct tctactatgg tactactatt ttcaaggctg ttggtttgag tgactctttc	1080
gaaacctcta ttgtcttggg tattgttaac tttgcttcca cctttgttgg tatttacggt	1140
gttgagagat atggctcgtc tacttgtttg ctatgggggtg ctgcatccat gactgcttgt	1200
atggttgtct atgcttccgt ggggtgcacc agattatggc caaatggtca agaccaacca	1260
tcttccaagg gtgctggtaa ctgtatgatt gtctttgcct gtttctatat tttctgtttt	1320
gctactacat gggctccaat tccttatgtc gttgtttctg aaactttccc attgagagtc	1380
aagtctaagg ctatgtctat tgctacagct gctaattggt tgtgggggtt cttgattggt	1440
ttcttcactc catttattac tgggtctatt aacttctact acggttacgt tttcatgggc	1500
tgtttggct tcattgttct ctatgttttg ttagttgttc cagaaactaa gggtttgact	1560
ttggaagaag tcaacacccat gtgggaagaa ggtgttctac catggaagtc tgcctcatgg	1620
gttccaccat ccagaagagg tgccaactac gacgctgaag aatgactca cgatgacaag	1680
ccattgtaca agagaatggt cagcaccaaa taa	1713

<210> SEQ ID NO 31

-continued

<211> LENGTH: 1695

<212> TYPE: DNA

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 31

```

atgtctagtg cgcaatcctc tattgatagc gatggagatg ttcgagatgc tgatattcat    60
gtcgcaccac ccggtgaaaa agagtggca gatggatttg atgacaacga agtcataaac    120
ggggataacg ttgagccacc aaaaagaggg ctcataggtt atcttgtcat ttacttactg    180
tgttatccaa tacccttgg gggtttctg cctggttggg atagtggat cacagcaggt    240
ttcattaaca tggacaactt taaatgaac ttcggttctt acaagcatag cactggtgaa    300
tattatttga gcaacgtgcg tatgggtctt cttgtggcta tggtcagtat tggatgtgcc    360
ataggtggcc ttatttttgc cgtcttgcct gatactttag gtagaaggct ggcaattgtg    420
atcgtggtgt tggatatatat ggttgggtgca attattcaga tcagttcaaa tcacaaatgg    480
taccagtact ttgtcggtaa gatcatctac ggtcttgggt ctggtggctg ttcggtgttg    540
tgtccaatgc ttttgtctga aatagcccc acagatttga gaggtggact ggtctcattg    600
taccaactga acatgaactt cgttattttc ttgggttatt gtagcgttta tggtagcaga    660
aaatacgata aactgcaca atggagagtc cccctgggc tttgctttt atgggctttg    720
attatcatca ttggtatggt attggttcca gactcccaa gatatctgat tgaatgtgag    780
agacacgaag agggccgtgc ttcattgccc aaaaacaaca aggtttcacc agaggatcca    840
tgggtactca aacaggtgta tgaatcaac gccggtgtcc ttgcccagg ggaactagga    900
gaagcttcat ggaagaact tttctctgta aaaactaaag tccttcaacg tttgatcaca    960
ggtattcttg tgcaaacctt tttgcaactt actggtgaaa actacttctt cttctacgga   1020
actaccattt ttaaatcagt cgttcttact gatgggtttg agacgtcgat cgtcctaggt   1080
acagtgaact tcttctccac tattattgct gttatggtcg tagacaaaat tggccgtcgt   1140
aaatgtctgt tatttgggtc agctgggatg atggcttcta tggtcattat tgcgaagtac   1200
ggggtgaaat gtcttaccoc tcattggccag gacggctcct cttcgaaaagg tgcaggtaat   1260
gccatgattg tggtcacttg tttctatata ttctgctttg caacgacatg ggctcctggt   1320
gcttatattg tggttgcca gtcgttccct tcgaaggcca agtctagagc catgtcgatt   1380
tcaactgcat gcaactggtt atggcaattt ttgatcgggt ttttcacacc attcattact   1440
gggtctatcc acttctatta tggttatgtg ttcgtaggtt gtttgggtgc tatgtttttg   1500
tacgttttct tctttttacc agaaacgatt ggtctatctt tggaggaaat ccaattacta   1560
tacgaagaag gtataaaaacc atggaaatct gcatcttggg tcccaccttc taggagaggt   1620
atctctccg aagaaagtaa gaccgagaag aaggattgga agaaattttt gaagttctca   1680
aagaattctg attga                                     1695

```

<210> SEQ ID NO 32

<211> LENGTH: 1704

<212> TYPE: DNA

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 32

```

atgaattcaa ctccagattt aatattctca caaaagtcaa gtgagaattc gaatgctgac    60
ctgccttcga atagctctca ggtaatgaac atgctggaag aaaaagggtg tcaagatgat    120
ttccaagctg aggccgacca agtacttacc aacccaaata caggtaaagg tgcataatgtc    180
actgtgtcta tctgttgtgt tatggttgc ttcggtggtt tcgttttctg ttgggatact    240

```


-continued

ggtaccattt ctggtttogt cgcccaaact gatttcttga gaagattcgg tatgaagcat	300
aaagatggta gttattatth gtctaagggt agaactgggt taattgtctc ctttttcaac	360
attggttggt ccattgggtg tattattttg gctaaattgg gtgatatgta cggtcgtaaa	420
atgggtttga ttgtcgttgt tgttatctac atcatcggtt ttattattca aattgcatcc	480
atcaacaaat ggtaccaata tttcatcggg agaattatth ccggtttggg tgttggtggg	540
attgcogttt tatctctcat gttgatttct gaagtcgctc ctaaggaaat gagagggtact	600
ttagtctcct gttaccaact gatgattacc ttgggtatth tcttgggtta ctgtaccaac	660
ttcggtaact agaactactc caactctgtg caatggagag ttccattagg tttgtgtttt	720
gcctgggctt tgtttatgat cgggtggtatg actttcgttc cagaatcccc acgttatttg	780
gttgaagctg gtcaaatgta cgaagcaaga gcactctctt ccaaagttaa caaggttgcc	840
ccagaccatc cattcattca acaagagttg gaagtattg aagctagtgt tgaagaagct	900
agagctgctg gttcagcatc atgggggtgag ttgttccactg gtaagccggc catgtttaag	960
cgtactatga tgggtatcat gatccaatct ctacaacaaat tgactggtga taactatttc	1020
ttctactatg gtactacogt ttttaacgct gttggtatga gtgattcttt cgaacttct	1080
attgttttctg gtgtcgtcaa ctctctctct acttgttgtt ctttgtacac tgtcogctgt	1140
tttggaagct gtaactgttt gttatatggt gccattggta tggctctctg ttatgtagtt	1200
tacgcttctg ttggtgtcac cagactatgg ccaaatgggt aaggtaatgg ttcattccaag	1260
ggtgctggta actgtatgat tgtctttgcc tgtttctata ttttctgttt tgctaccact	1320
tgggtctcaa ttgcttatgt tgttatttct gaaactttcc cattgagagt caagtctaag	1380
gctatgtcta ttgctacagc tgctaattgg ttgtgggggt tcttgattgg tttcttact	1440
ccatttatta ctggtctcat taactctctc tacggttacg ttttctatgg ctgtatgggt	1500
ttcgcctact tctacgtttt ctctcttctg ccagaaacta agggtttgac tttggaagaa	1560
gtcaatgata tgtacgctga aggtgttcta ccatggaagt ctgcttcatg ggttccaaca	1620
tctcaaagag gtgctaacta cgatgctgat gcattgatgc atgatgacca gccattctac	1680
aagaaaatgt tcggcaagaa ataa	1704

<210> SEQ ID NO 33

<211> LENGTH: 1713

<212> TYPE: DNA

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 33

atgaattcaa ctcccgatct aatatctcct cagaaatcca attcatccaa ctcatatgaa	60
ttggaatctg gtcgttcaaa ggccatgaat actccagaag gtaaaaatga aagttttcac	120
gacaacttaa gtgaaagtca agtgcaaccc gccgttgccc ctccaaacac cggaaaaggt	180
gtctacgtaa cggtttctat ctgtttgtgt atggttgctt tcggtggttt catatttgga	240
tgggatactg gtaccatttc tggttttgtt gctcaaactg attttctaag aagatttggt	300
atgaagcacc acgacggtag tcattacttg tccaagggtg gaactggttt aattgtctct	360
atthtttaaca ttggttggtc cattggtggg atcgtcttag ccaagctagg tgatattgat	420
ggtcgtagaa tcggtttgat tgcgttggtg gtaactctaca ctatcgggtt cattattcaa	480
atagcctcga tcaacaagtg gtaccaatat ttcattggta gaattatctc tggtttaggt	540
gtcgtgggta tcacagtttt atctcccatg ctaatatctg aggtcgcctc cagtgaatg	600

-continued

agaggcacct tggtttcacg ttaccaagtc atgattactt taggtatfff cttaggttac	660
tgtaccaatt ttggtaccac gaattactca aactctgtcc aatggagagt tccattaggt	720
ttgtgtttcg cctgggcott atttatgatt ggtggtatga tggttgttcc tgaatctcca	780
cgttatattg ttgaagctgg cagaatcgac gaagccaggg cttctttagc taaagttaac	840
aatgcccac ctgaccatcc atacattcaa tatgagttgg aaactatcga agccagtgtc	900
gaagaaatga gagccgctgg tactgcatct tggggcgaat tattcaactgg taaaccagcc	960
atgtttcaac gtactatgat gggatcatg attcaatctc tacaacaatt aactggtgat	1020
aactatttct tctactacgg taccattggt ttccaggctg tcggtttaag tgactctttt	1080
gaaacttcta ttgtctttgg tgcgtcaac ttctctcca cttgtgttcc tctgtacacc	1140
gttgaccgtt ttggccgtcg taactgtttg atgtggggtg ctgtcggtat ggtctgctgt	1200
tatgtgtctc atgctctgtg tgggtttacc agattatggc caaacggta agatcaacca	1260
tcttcaaagg gtgctggtaa ctgtatgatt gttttcgcat gttctacat tttctgttcc	1320
gctactacct gggcccacat tgcttaagtt gttatttcag aatgtttccc attaagagtc	1380
aatccaagt gtatgtctat tgccagtgtc gtaactgga tctggggttt cttgattagt	1440
ttcttcaccc catttattac tggtgccatc aactctact acggttacgt tttcatgggc	1500
tgtatggtt tgcgttaact ttacgtctt ttctctgttc cagaaactaa aggtttatca	1560
ttagaagaag ttaatgatat gtacgccgaa ggtgtttctac catggaaatc agcttctcgg	1620
gttccagtat ccaagagagg cgtgactac aacgctgatg acctaagca tgatgaccaa	1680
ccattttaca agagtttgtt tagcaggaaa taa	1713

<210> SEQ ID NO 34

<211> LENGTH: 1644

<212> TYPE: DNA

<213> ORGANISM: Debaryomyces hansenii

<400> SEQUENCE: 34

atgggttacg aagataaatt agtggctcct gctttaaagt ttagaaaact tttggacaag	60
actccaaata tctataatgt tttgttatt gctctgattt catgtatctc cgggttgatg	120
tttggtattg atatctcacc tatgtcttta tttattgggg atgacaaata cattaaatac	180
ttccataaac ctgacacaac catgcaatcg tttattactt cggctatgtc tttgggttca	240
ttttttggat caatatgttc ttcttttgg tccgaaccat ttggcagaag atcatcgta	300
atggtatgtg ggtttttctg gtgtgttggg gctgctatc aatcatctgc tcaaaatcaa	360
gcgcaattga ttatcgcccg ttttatttct ggttttgggt ttggtttcgg ttcacagtt	420
gctccagttt atggatctga attagctcct agaaagatca gaggtttaat tgggggtcct	480
tttcaattct ctgtcacttt gggatcttta atcatgttct atatctgtta tggtttgaat	540
tttattaacg gtgtcgcacc cttcagagtt gcttgggggt tacaatcat accaggttta	600
gttttgatcc ttggttgttt ctttattcca gaatctccta gatggttagc taaacaaggt	660
tactgggaag atgctgaata tgttgttggc aagatccaag ctaagggtta cagagaagat	720
ccagatgttt taattgaaat gtctgaaatt aaggaacaaa ttatgttaga tgaacacatc	780
aaggctttca catatgctga tttgttccact aagaagtata tattgagaac tgttactgct	840
tgttgggctc aagcctggca acaattaacc ggtatgaata ctttgatgta ttatattggt	900
tatgttttcc aatggccgg ttacgaaggt gatgctaatt tggttgctag ttccattcaa	960
tactgtctta aactggtat gaccattccc gcattatact tcatggataa gcttggtaga	1020

-continued

```

agaccagttt tattaactgg ggctgcattc atgatggcct ggcaattcgc cgttggtgga 1080
ttattggcaa cttacagtgt tgataaccct atcagtgaac ccgtcagaat tcaaattcca 1140
gaagaacatg gcaaggcgcg caaagctgtc attgcttgcg gttatttatt cgttggttct 1200
ttcgcttgta gttggggtgt ttgtatttgg gtttactgtg ctgaagtctg gggtgatagt 1260
gcttcaagac aaagagggtg tgctcttacc acttctgtta attggatttt caatttcgcc 1320
attgctatgt tcaactcaag tgctttcaag aacattactt ggaagactta catgggtctc 1380
gctacattct gtgggtgtat gtttatccat gtcttctctt tcttccaga aactaagggt 1440
aagagattgg aagaaattgg tcaaatgtgg gccgaagggt taccagcttg gaagtctgct 1500
tcttggaac catctattcc aatcgtgtcc gataatgaat tacataacaa gatgaagatt 1560
gaccataatg aagataactt attaaattct tcttcacatt ctgaagtatc ccaagagaag 1620
ggcgatactt ctcacatgac ttaa 1644

```

<210> SEQ ID NO 35

<211> LENGTH: 1638

<212> TYPE: DNA

<213> ORGANISM: *Debaryomyces hansenii*

<400> SEQUENCE: 35

```

atgggatatg aagaaaagt ggctgcacca gctctaaaac taagattatt cttagataaa 60
cttccaaaca tttataatat ttatgtcatt gctactatct cgtgtatttc ggggttgatg 120
ttcggtatcg acatttcatc gatgtcggcg ttccttagta atgacgctta ccttaaatat 180
tttggtaccc ctgaacctga tatgcaagga ttcattactg ctgctatgct tcttgggtca 240
ttctttgggt ctcttgcate ggcattttgt tcagagccat ttggtagaag agcttcattg 300
ttgctttgtg gattcttttg gtctgttggg gctgcaatc aatcttcatc acagaatgct 360
gcccagttaa ttattggtcg ttttatttct ggttttggaa ttggtttcgg ctcatctgtg 420
gcaccctgtg atggatcaga gttggcacct agaaaaatta gagggttaat tgggtgctct 480
ttccaaactt cagttacctt aggtattttg attatgttct atatctgtta cgggttaggt 540
aaaatccaag ccgtcggctc attcagaact gcgtgggggt tgcaaatat cccaggattg 600
atcttaatac ttggatgttt ctttattcca gaatcaccta gatggttagc caaacagaac 660
tactgggaag aggcggagga cattgtcgcg agagtccaag ctaaaggtaa cagagaagac 720
cctgaagtat taattgaaat ggtgaaatt agagaaccaga ttctgacgtt agacaaagt 780
aagtccttta cttatattga tttattcaaa aagaagtatc ttcttagaac tgtaacggca 840
atatttgctc aaatttggtc acaattaaca ggtatgaaca ccttaatgta ttatattgct 900
tatgttttgc aaatggctgg ttatcatggt gatgcaaat ttggttgcctc ttcaattcaa 960
tattgtatca attttgccat gactattcct gcattgtact taatggataa ggtcggtagg 1020
agaccagtgt tgtaactcgg agctgcgctc atgatggcct ggcaatttgc tatagggggc 1080
ttgcttgcca cttatgctga acccacgat atttttgggg gtaataatac cgttaaaatc 1140
agtatcccag aagatgaatc tcttgcggcc aaggctgtta ttgcatgttg ttacttattc 1200
gttgtttcat ttgcgtctac ttgggggtgt ggtatctggg tttactgtgc tgaagtctgg 1260
ggtgatagtg cctcaagaca aagagggtcg tgtgttgcaa ctgctggtaa ttggattttt 1320
aatctcgcaa ttgcaatggt cactccgcac gcttccagta ctattacttg gaaaacatac 1380
atgatttttg ctactttctg tgcattgatg tttcttcatg ttttctctt tttccctgaa 1440

```

-continued

actaagggta aaagattgga agaaattggg caaatgtggg acgagcatgt tccagcttgg	1500
aagtctgctt cgtggcaacc acacgtacca ctctttcag ataataaat acacgggaag	1560
atggattctg cgcattgatga acattcatct cgttcggagt ctactggaga aaaggtggtc	1620
gcagatcaca ttgcttga	1638

<210> SEQ ID NO 36
 <211> LENGTH: 1638
 <212> TYPE: DNA
 <213> ORGANISM: Debaryomyces hansenii

<400> SEQUENCE: 36

atgggatacg aagataaatt actaggcca gccttaaagt ttagaaactt cttggacaaa	60
tttccgaaca tccataatgt ctatattgtt gttggtattt cttgtatatac tggatgatg	120
tttggcattg atattttctc aatgtcttta ttcattggag atgataaata tttagactat	180
ttcaattcac cagattctac acttcaaggg tttattacag catccatgtc tttagggctc	240
ttctttgggg ccttattttc cgcatttata tcagaacctt tcggtagaag aatgtcgttg	300
atgttctctg catttttttg gtgtgttga gctgccatc aatcatcgtc acagaacgctc	360
gtccagttaa ttataggctg ttttatttct ggttttggtg tggatttgg atcctcagta	420
gccccagttt atggtacaga attggcacca agaaaaataa gaggattaat tgggtgatta	480
tttcagcttt cggtcacttt gggaaatttg gttatgttct atgtttgtta tgcattacat	540
tatatcaatg gtgtggcttc ttttagatta tcttgggggt tacaataagt tctgtgtctt	600
cttttgttca ttggctgttt tttcatcctc gaatcaccta gatggttagc aaagcaaggc	660
tgttgggagg aggctgaata cattgtggca atgattcagg ccaaaggcaa tagggaagat	720
cccgatgtca tgattgagat tactgagatc aaagatcaaa ttttgacaga ggaaaatatt	780
aaggctttca cttatgcaga tttattcaag cgcaaatatc tccttagaac tgtcacgcga	840
acatttctc aaatatggca acaattaacg ggtatgaata cattaatgta ttatattgtt	900
tatgttttg atatggctgg ctatcagggc gatgcaaat taattgcatc tccaattcaa	960
tacgttcttt tttcgttat gactgcccct tcattatatt taatggataa acttggtaga	1020
aggcccattt tgtaagcgg tctgtcattc atgatgatat ggcaattcgc agtcggtggt	1080
ttgctttcca cttatgctga gccccaat gatgttggg gaaatgatac tgtcagatta	1140
aggatcccag ctgataactc aaccgctgcc aagggtgtaa ttgcatgttg ctatttgttt	1200
ggtgtatcat ttgcatatag ttgggggtt tgtatctgga tgtattgtgc tgaagtttgg	1260
gggtatagcg cctcaagaca aagaggggccc tgttttacia cttcagctaa ttggattttt	1320
aacttcgcga tcgcaatggt taccctctc gcgttcaaaa acattacttg gaaaacatac	1380
atgatttttg ctacgttctg cggttgtatg ttccttcacg tatttttctt tttcccagaa	1440
actaaaggta aaagattgga ggaaattggc caaatgtggg atgaaggaat tccagcatgg	1500
agaacagcgg catgggagcc atctattcca ttcctatctg ataatgactt gcgtgaaaag	1560
ctagaagtaa aacacgtcga agattcaaac tccagtaatt cggatgctga aaagcctagt	1620
gctgtccata ttgcttag	1638

<210> SEQ ID NO 37
 <211> LENGTH: 1851
 <212> TYPE: DNA
 <213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 37

-continued

```

atgaaaaata tcatttcatt ggtaagcaag aagaaggctg cctcaaaaa tgaggataaa    60
aacatttctg agtcttcaag agatattgta aaccaacagg aggttttcaa tactgaagat    120
tttgaagaag ggaaaaagga tagtgccctt gagctagacc acttagagtt caccaccaat    180
tcagcccagt taggagatc tgacgaagat aacgagaatg tgattaatga gatgaacgct    240
actgatgatg caaatgaagc taacagcgag gaaaaaagca tgactttgaa gcaggcgctg    300
ctaaaatata caaaagcagc cctgtggctc atattagtg ctactaccct ggttatggaa    360
ggttatgata ccgcactact gagcgactg tatgcctgc cagttttca gagaaaattc    420
ggttacttga acggggaggg ttcttaogaa attacttccc aatggcagat tggtttaaac    480
atgtgtgtcc tttgtgtgga gatgattggg ttgcaaatca cgacttatat ggttgaattt    540
atggggaatc gttatacgat gattacagca ctgggttgt taactgctta tatctttatc    600
ctctactact gtaaaagttt agctatgatt gctgtgggac aaattctctc agctatacca    660
tggggttggt tccaaagttt ggctgttact tatgcttcgg aagtttgccc tttagcatta    720
agatattaca tgaccagtta ctccaacatt tgttggttat ttggtcaaat ctgcctctc    780
ggattatga aaaactcaca agagaattta gggaaactccg acttgggcta taaattgcca    840
tttgctttac aatggatttg gcctgctcct ttaatgatcg gtatctttt cgctcctgag    900
tcgccttggg ggttgggtgag aaaggatagg gtcgctgagg caagaaaatc ttttaagcaga    960
atthttgagtg gtaaaggcgc cgagaaggac attcaagttg atcttacttt aaagcagatt   1020
gaattgacta ttgaaaaaga aagactttta gcatctaaat caggatcatt ctttaattgt   1080
ttcaaggagg ttaatggaag aagaacgaga ctgcatggt taacttgggt agctcaaat   1140
agtagcggtg ccgttttact tggttactcg acatattttt ttgaaagagc aggtatggcc   1200
accgacaagg cgtttacttt ttctctaatt cagtactgtc ttgggttagc gggtacactt   1260
tgctcctggg taatatctgg ccgtgttggg agatggacaa tactgacctg tggctctgca   1320
tttcaaatgg tctgcttatt tattattggg ggaatgggtt ttggttctgg aagcagcctc   1380
agtaatggtg ccggtggttt attgctggct ttatcattct tttacaatgc tggtatcggg   1440
gcagttggtt actgtatcgt tgetgaaatt ccatcagcgg agttgagaac taagactata   1500
gtgctggccc gtatttgcta caatctcatg gccgttatta acgctatatt aacgccctat   1560
atgctaaaac tgagcgattg gaactggggg gccaaaactg gtctatactg gggtggtttc   1620
acagcagtcg ctttagcttg ggtcatcctc gatctgctcg agacaactgg tagaaccttc   1680
agtgaaatta atgaactttt caaccaaggg gttcctgcca gaaaatttgc atctactgtg   1740
gttgatccat tcggaaaagg aaaaactcaa catgattcgc tagctgatga gagtatcagt   1800
cagtcctcaa gcataaaaaca gcgagaatta aatgcagctg ataatgtta a           1851

```

<210> SEQ ID NO 38

<211> LENGTH: 1710

<212> TYPE: DNA

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 38

```

atgaaggatt taaaattatc gaatttcaaa ggcaaattha taagcagaac cagtcactgg    60
ggacttacgg gtaagaagtt gcggtatttc atcactatcg catctatgac gggcttctcc   120
ctggttggat acgaccaagg gttgatggca agtctaatta ctggtaaaca gttcaactat   180
gaatttcag caaccaaaga aatggcgat catgacagac acgcaactgt agtgcagggc   240

```

-continued

gctacaacct cctgttatga attaggttgt ttcgcaggtt ctctattcgt tatgttctgc	300
ggtgaaagaa ttggtagaaa accattaatc ctgatgggtt ccgtaataac catcattggt	360
gccgttattt ctacatgcgc atttcgtggt tactgggcat taggcagtt tatcatcgga	420
agagtcgca cccgtgttgg aacagggttg aatacatcta ctattcccgt ttggcaatca	480
gaaatgtcaa aagctgaaaa tagagggttg ctggcaatt tagaaggttc cacaattgct	540
tttggtacta tgattgctta ttggattgat ttgggttctt cttataccaa cagttctggt	600
cagtgagat tccccgtgc aatgcaaatc gtttttctc tcttctgct tgctttcatg	660
attaactac ctgaatgcc acgttggctg atttctcaa gtcgaacaga agaagctcgc	720
tacttgtag gaacactaga cgcgcggat ccaaatgatg aggaagtat aacagaagt	780
gctatgctc acgatgctgt taacaggacc aaacacgaga aacattcact gtcaagttg	840
ttctccagag gcaggtecca aaatcttcag agggcttga ttgcagctc aacgcaatt	900
ttccagcaat ttactggtg taacgtgcc atatactact ctactgtatt attcaacaaa	960
acaattaaat tagactatag attatcaatg atcataggtg gggctctgc aacaatctac	1020
gccttatcta ctattggttc attttttcta attgaaaagc taggtagacg taagctggtt	1080
ttattagggt ccacaggtea agcagtttca ttcacaatta catttgcag cttggtaaaa	1140
gaaaataaag aaaacgcaag aggtgctgcc gtcggcttat tttgttcat tacattctt	1200
ggtttgtctt tgctatcatt accatggata taccaccag aaattgcatc aatgaaagt	1260
cgtgcatcaa caaacgttt ctccacatgt actaattggt tgtgtaactt tgcggttgc	1320
atgttcacc caatatttat tggacagtcc ggttgggtt gctacttatt tttgctgtt	1380
atgaattatt tatacattcc agttatcttc ttttctacc ctgaaaccgc cggaagaagt	1440
ttggaggaaa tcgacatcat ctttgtctaa gcatacagag atggcactca accatggaga	1500
gttgctaacc atttgcccaa gttatcccta caagaagtcg aagatcatgc caatgcattg	1560
ggctcttatg acgacgaaat gaaaaagag gactttggtg aagatagagt agaagacacc	1620
tataacaaa ttaacggoga taattcgtct agttcttcaa acatcaaaaa tgaagataca	1680
gtgaacgata aagcaaat ttagggttga	1710

<210> SEQ ID NO 39

<211> LENGTH: 1626

<212> TYPE: DNA

<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 39

atgtctgaat tcgctactag cccggttga agtggtctc aacaaacttc tatccactct	60
actccgatag tgcagaaatt agagacggat gaatctccta ttcaaacc aaatctgaatac	120
actaacgctg aactcccagc aaagccaatc gccgcattat ggactgttat ctgtttatgt	180
ctaatgattg catttgggtg gtttctctt ggttgggata ctggtaccat ctctggtttt	240
gttaatcaaa ccgatttcaa aagaagattt ggtcaaatga aatctgatgg tacctattat	300
ctttcggagc tccggactgg tttgatcgtt ggtatcttca atattggtt tgcctttggt	360
gggttaacct taggacgtct ggtgatgatg tatggacgta gaattggtt gatgtgcgtc	420
gttctggtat acatcgttgg tattgtgatt caaattgctt ctagtgacaa atggtaacaa	480
tatttcattg gtagaattat ctctggtatg ggtgtcggtg gtattgctgt cctatctcca	540
actttgattt ccgaaacagc accaaaacac attagaggtta cctgtgttct tttctatcag	600
ttaatgatca ctctaggtat tttcttaggt tactgtacca actatggtac taaagactac	660

-continued

tccaattcag	ttcaatggag	agtgcctttg	ggtttgaact	ttgccttcgc	tattttcatg	720
atcgctggta	tgctaattgt	tccagaatct	ccaagattct	tagtcgaaaa	aggcagatac	780
gaagacgcta	aacgttcttt	ggcaaaatct	aacaaagtca	ccattgaaga	tccaagtatt	840
gttgctgaaa	tggatacaat	tatggccaac	gttgaaactg	aaagattagc	cggtaacgct	900
tcttggggtg	agttattctc	caacaaaggt	gctattttac	ctcgtgtgat	tatgggtatt	960
atgattcaat	ccttacaaca	attaactggg	aacaattact	tcttctatta	tggtactact	1020
atthtcaacg	ccgtcgggat	gaaagattct	ttccaaactt	ccatcgtttt	aggatatagtc	1080
aacttcgcat	ccactttcgt	ggccttatac	actgttgata	aatttggctg	tcgtaagtgt	1140
ctattgggtg	gttctgcttc	catggccatt	tgttttgta	tcttctctac	tgctcgtgtc	1200
acaagcctat	atccaaatgg	taaagatcaa	ccatcttcca	aggctgcggg	taacgctatg	1260
attgtcttta	cctgtttatt	cattttcttc	ttcgtctatta	gttgggcccc	aattgcctac	1320
gttattgttg	ccgaatccta	tcctttgcgt	gtcaaaaatc	gtgctatggc	tattgctgtt	1380
ggtgccaact	ggatttgggg	tttcttgatt	ggtttcttca	ctcccttcat	tacaagtgca	1440
attggatttt	catacgggta	tgtcttcatg	ggctgtttgg	tattttcatt	cttctacgtg	1500
tttttctttg	tctgtgaaac	caagggtcta	acattagagg	aagttaatga	aatgtatgtt	1560
gaaggtgtca	aaccatggaa	atctggtagc	tggatctcaa	aagaaaaaag	agtttccgag	1620
gaataa						1626

<210> SEQ ID NO 40

<211> LENGTH: 1710

<212> TYPE: DNA

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 40

atgactgata	gtaaaaccaa	cttgccagaa	gaaccgattt	tcgaagaggg	agaagatgat	60
ggctgccctt	cgatagaaaa	ttcttcacat	ctgtcagtac	ctacagtgga	ggaaaacaag	120
gacttttccg	agtataatgg	ggaagaggca	gaggaaagtg	ttgttccaga	aaagcctgct	180
tcagcctatg	ctactgtttc	tatcatgtgt	ttatgtatgg	ctttcgggtg	atthtatgtcc	240
ggttgggaca	caggtacgat	ttctggtttc	gtcaatcaga	ctgatttttt	aagaagattt	300
ggtaattata	gccattccaa	gaacacttac	tacttatcta	atgtgagaac	tgggttgatt	360
gtgtccatct	tcaatgtggg	aagcggcatt	ggctgtcttt	tcttgcctaa	attgggtgat	420
atthacggcc	gctgcatggg	tttgattata	gttattgtcg	tttatatggg	tggtattgtc	480
attcaaatgg	cctctataga	taagtgggat	cagtatttta	ttggaagaat	tatcgtcggg	540
ataggtgctg	gttccattag	tgttcttgcc	ccgatgctta	tttcggaaac	tgcgccaag	600
catatcagag	gtacgttgct	agcttgttgg	caattgatgg	tgactttcgc	aattttcttg	660
ggttattgta	ccaattatgg	taccaagact	tactcgaatt	ctgtccagtg	gcgtgttccg	720
cttggcttat	gttttgcatg	ggctattatt	atgattggtg	gtatgacggt	tgthccggaa	780
tctctcggg	ttttgggtca	agtcggtaag	attgagcaag	ctaaagcttc	ttttgccaag	840
tcgaacaagc	ttagtgttga	cgatcctgct	gtggttgcag	agattgatct	tcttgttgct	900
gggtggagg	cagaagaagc	aatgggaact	gcttcatgga	aggaattatt	ttcgagaaag	960
actaaagttt	ttcaacgttt	aacgatgaca	gtcatgatta	actctctgca	gcaactaacc	1020
ggtgacaact	atthcttcta	ctacggtaact	actatcttca	aatctgtcgg	tatgaatgac	1080

-continued

tcttttgaga cttcaattgt cttgggtatt gtgaattttg cttcttgctt cttttcactt	1140
tattctgttg ataagttggg ccgctcgtaga tgtcttttac ttggagcagc caccatgacg	1200
gcgtgcatgg ttatttaagc ctcggtggc gtcacaagac tatatccgaa cggtaaaagt	1260
gaaccatcat ctaaagggtgc tgtaattgt acgattgttt tcacgtgttt ttacattttc	1320
tgcttttctc gcacctgggg acctgtatgt tatgtgatta tttctgaaac atttccatta	1380
agggtgagat ccaagtgtat gtcggtgca acagcggcca acttattgtg ggggttctca	1440
atcgggtttt tcaactcttt tattacttcg gcaattaatt tctactacgg ttacgttttc	1500
atgggttgct tagcgttttc atatttttac gtcttttctt ttgttccaga aacaaaaggt	1560
ctaactttag aagaagtga tgagatgtgg atggacggtg tattaccttg gaaatctgaa	1620
tctgggtac cagcttctag aagggatggt gattatgata acgaaaaatt acagcatgac	1680
gagaaacct tctacaaaag aatgttttag	1710

<210> SEQ ID NO 41

<211> LENGTH: 1704

<212> TYPE: DNA

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 41

atgtccggtg ttaataatac atccgcaaat gatttatcca ctaccgagtc taactctaac	60
tcagtagcaa atgcaccatc tgtaaaaact gagcataatg actctaaaaa ctccctcaac	120
ctggatgcca ctgaaccacc tattgactta cctcaaaaac ccctctctgc atataccacc	180
gtcgcfaatc tgtgtttgat gattgcattt ggcggttca tctttggttg ggataccggt	240
accatttctg gttttgttaa cctttctgat ttcacagaa gggtcggta aaaaaatgac	300
aagggaaact actacttacc gaaagtaaga atgggtttga tcgtctcaat attcaacatt	360
ggctgcgcca taggcggaat tgtctgtca aaagtcggtg atatatatgg tcgtcgtatt	420
ggattgatta cagtactgc catttacgtt gtaggcaccc taatccaaat aacttccata	480
aacaagtgg accaataact cattggaaga attatttctg gcctaggagt gggaggcatt	540
gctgtccttt ccccaatggt gatatctgaa gttgctccca acaaatcag aggaaccctg	600
gtccaattgt accagctgat gtgtacgatg ggtatttttc taggatactg taccaattac	660
ggtaccaaga actatcacia cgcactcaa tggagagtcg gccttggctt ttgctttgcc	720
tggactacat tcattggttag tggaaatgat tttgtaccag aatcaccacg ttacctgatt	780
gaggttggtg aagatgagga agcgaaacgt tcaacttcca aatccaacia agtctcagtc	840
gacgatccag ccttggttagc agaataatgac actataaagg cgggaatcga acttgaaaag	900
ctggcaggta acgcatcatg gtctgaaact ctctccacta aaacaaaggt ctttcagcgt	960
gttctcatgg gagtcatgat ccaatcgtg cagcaattaa ccggtgataa ctacttcttt	1020
tactacggca ccaccatctt caaatctgtc ggtctaaaagg actcctttca gacttcgatc	1080
attatcggtg tggtaatttt tttctctca ttcatacagg tataccatc tgagaggttt	1140
ggacgccgta cgtgtctatt gtgggtgct gcttctatgc tatgctgctt tgctgtgttt	1200
gcctccgctg gtgtgacaaa gttgtggcct caaggaagca gtcaccaaga cattaactct	1260
cagggcgcgc gtaactgtat gattgtgttt actatgtctt tcattttttc gttcgcacc	1320
acttgggcag gcgctgtta cgttattgtc tcagagacgt ttctcttag ggtcaaatca	1380
agaggaatgg caatcgcaac agctgcaaac tggatgtggg gtttctctgat tagtttcttt	1440
acccattca ttaccggggc aatcaacttt tactacggtt atgtattctt aggotgtctg	1500

-continued

```

gtttttgcat acttttatgt ctttttcttt gtcccagaaa caaaaggcct gacgctggag 1560
gagggtgaata ctatgtggct ggaaggtgtg ccagcatgga aatcagcctc atgggtgcca 1620
ccagaaagaa gaaccgcaga ttacgatgct gacgccatag accatgacga tagaccaatc 1680
tacaagaggt tcttttccag ctaa 1704

```

```

<210> SEQ ID NO 42
<211> LENGTH: 1641
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

```

```

<400> SEQUENCE: 42

```

```

atggttagtt caagtgtttc ctttttgggg actagcgcca aggcattccac ttctctaagt 60
agaaaggatg aaattaaact aacctctgaa acaagggaag ctagcttggga cattccatac 120
aaaccatta ttgcatactg gacggtgatg ggtctctgtc tgatgattgc ctttgggtgga 180
ttcatttttg gttgggatac aggaaccatt tcagggttta ttaaccaaac agatttcaag 240
agaaggtttg gtgagttaca aagggacggc agttttcaac tatcagatgt caggacaggg 300
ctaattgtcg gtatcttcaa cataggttgt gctttagggt gcctaacgct gggacgctg 360
ggcgatattt atgggctgaa aatcggtcta atgtgtgta tactggtgta tgttgttgg 420
atcgtgatcc agattgcttc ctctgacaaa tggatcaat atttattgg tagaattgtt 480
tctggaatgg gtgttgagg tgttgcgtg ctgtcgccaa ctttgatctc agaaatttc 540
ccaaagcacc taagggcac ttgtgtctct ttttaccagc taatgattac ccttgaatt 600
ttcttgggct actgtaccaa ttatgttaca aagaaatatt caaattcaat acagtggcgg 660
gttccttgg gtttgtgtt tgcgtgggca atctttatgg tgattggaat ggttatggtt 720
ccggaatcgc ccagatattt agtagaaaaa ggtaagtag aagaagctag aaggctcttg 780
gccaaatcaa acaaggctac agttactgat ccaggcgttg tttttgagt tgatactata 840
gttgcaaata tggaaataga aagggctgtt ggaaatgcca gttggcacga actcttctca 900
aataaaggag caattctacc aagggttaata atgggaatcg ttatccagtc actgcaacag 960
cttactggct gtaattattt tttctactac ggcacgacca ttttcaatgc tgttggaatg 1020
caagactctt tcgagacttc cattgtcctt ggggctgta attttgcctc tacatttgtt 1080
gcactataca ttgtggataa atttggcgt cgaaaaatgtt tattgtgggg gtctgctcgc 1140
atggcaattt gtttgcctat attcgcacc gttggcgtca ctagattatg gccacaaggg 1200
aaagaccaac cttcttcgca aagtgcctgt aatgttatga tcgtttttac ttgtttcttc 1260
atcttctctt ttgccattac ttgggtcctc atcgcctatg tcattgtggc agaaacttat 1320
ccattaagag ttaaaaaatcg tgccatggcc attgcggttg gtgcgaaactg gatgtgggg 1380
ttcttgattg gatttttcac accctttatc actagatcca taggattttc ttatggctat 1440
gttttcatgg gttgcttaat cttttcgtac ttctacgttt tcttctttgt ttgcgaaaca 1500
aagggattaa ctctggagga agttaatgaa atgtacgaag aaagaataaa gccatggaag 1560
tccggagggt ggattcccag ttctagaaga acaccacaac caacaagcag tacaccatta 1620
gttattgttg atagtaaata a 1641

```

```

<210> SEQ ID NO 43
<211> LENGTH: 1623
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

```

-continued

<400> SEQUENCE: 43

atgactgctc agattccgta tcaacatagc tcgggataca tctctcattt tcacaataat	60
gagcttgatg caggcagggg aagggtattat aatgtaacca ttaagtatct agatgataaa	120
gaagaaaata tagaaggcca agcagcaaag attagtccca atgcgagtct gcatattccc	180
gttttattgt gcttggaat ctcgcttggg ggctttattt ttggatggga cattggaacc	240
atcggtgga tgacaaatat ggttagcttt caagaaaaat ttggcacaac taatattatc	300
catgacgatg aaacaatttt tgtatctact aagaaactta ctgatctgca aataggccta	360
attatcagta tttttaacat cagtttggtc gtaggggcct taactctgtc aaaaatcggg	420
gattggattg gtaggaaagg tggatatagg ttgctcttag tagtgactg catcggtata	480
accattcaaa ttctctccta tggaaagtgg tatttttga cattgggaag agccgtaacg	540
ggaatcggtg tggagtaac cactgtcttg gtgccaatgt ttctctccga gaattctcca	600
ctaaaaatca gaggtcccat ggtatctacg tatcaattga ttgtaacatt tggcactacta	660
atgggaaaca ttttaaattt catatgcgaa agatgttata aagatcctac acaaaatata	720
gcctggcaat tgccattggt cttgggatac atttgggcaa ttataattgg aatgtcactt	780
gtttacgttc ctgaatctcc acagtacttg gcaaaaaatca aaaatgatgt gccctctgct	840
aaatactctt ttgcgaggat gaatggcatc cctgcgacgg atagcatggt aattgaattc	900
atcgatgatt tgctggaaaa taactataat aatgaggaaa ctaacaacga atcaaaaaag	960
caaagcttag ttaaaagaaa cacatttgaa tttattatgg gaaagccaaa gttatgggtg	1020
agactgatta ttggtatgat gataatggca tttcaacagc tgtccggaat aaattatttc	1080
ttttattacg gaacgtctgt tttcaaaggt gtcgggatta aggatcctta tattacttca	1140
atcactactg caagtgttaa cttcctttct acgatattag gcatatatta cgtggagaaa	1200
tggggccaca agacatgttt attatatggt tcaacaaatt tattatttta tatgatgaca	1260
tatgctactg tggggacatt tggaaagaaa acggacttct caaatattgt ttaattatc	1320
gtgacttggt gttttatttt ttggtttgca ataacattgg gccagttac atttgacta	1380
gtgtccgaat tgttccctct aagaacgagg gccatatcaa tggctatttg cacatttatc	1440
aattggatgt tcaatttctt aatatcactt ttaacaccaa tgattgtatc caaaattgat	1500
ttcaaaactag gatacatatt tgetgcttgc ctattagcgt tgataatatt cagttggata	1560
ctagttcctg aaacgaggaa aaagaatgag caagagatca ataaaatatt tgaaccggag	1620
tag	1623

<210> SEQ ID NO 44

<211> LENGTH: 2067

<212> TYPE: DNA

<213> ORGANISM: *Scheffersomyces stipitis*

<400> SEQUENCE: 44

ctacaagtga cagtcagtcg attagacttt gcatccactt gagtttgaca attgatatat	60
tccactagag acaatgagtg ctgacgaaaa agtcgctgct gccggccagg acggcttggt	120
tgaacacaac agttccactt cgagcatcga ggacaagaag ccctccaaga gctccgatgt	180
cgattccgtg aactcgcaat tagtagacaa ctcggtagag ggcaacatct tgtcccagta	240
caccgaaagt caggtgatgc agatgggtag aagctatgcc accaagcacg gcttggacce	300
agaattgttc gccaaaggcag ctgctgttgc cagaactcct cttggtttca actccatgcc	360
cttcttgaca gaggaagaga aggttggttt gaatgccgaa gccactaata agtggcacat	420

-continued

tccaccaga ttgateggg ttattgocct gggttctatg gcegetgctg tgcagggtat	480
ggacgaatcg gtcattaacg gtgccaactt gttctacccc aaggctttcg gagtcgacac	540
catgcacaat tccgacttga ttgaaggttt gatcaatggt gctccttacc tttgctgtgg	600
tattctttcc tgttggttgt ctgaecgttg taaccgtcgt cttggtagaa aatggaccat	660
tttctgggtg tgtgtcattt ctgccatcac ctgtgtctgg caaggcttg tcaacaactg	720
gtaccatttg ttcattgctc gtttcttccct tggatttggg gttggatca agtccgccac	780
tgttctgccc tactctgccc aatgtactcc taaacacatc agaggttcgt tagtcatggt	840
gtggcaatc ttcacagctg ttggattat gtttggttat gttgcttcc tggctttcta	900
caatgctcga gatagaggaa tccattacgg gttgaactgg agattgatgc ttggttcggc	960
cgctattcct gctgtcatca tcttgttcca aattccttcc gctcctgaat ctccacgttg	1020
gttaatgggt aaggacagac accttgaagc ctttgagtc ttgaagcaat tgagatacga	1080
agaacttgct gctgctcgtg actgtttcta ccagtacgtc ttgttagctg aagaaggttc	1140
ttacaagatc ccaacctca ccagatttaa ggaaatgttc accaagagaa gaaacagaaa	1200
cggtgccatc ggtgcattta ttgtcatggt catgcaacag ttctgtggta tcaacgcat	1260
tgttacttac tcttcgtcta tctttgtcca atctggtttc tctcaactt ctgctttgat	1320
cgcttcttgg ggttctggta tgettaactt cacctttgcc attcctgctt tcttcacaat	1380
cgatcgttcc ggtagaagat ccttattggt ggttaccttc cccttgatgg ctattttctt	1440
attgattgcc ggttctcggt tcttgataaa cgaagaaaca aactccaagg gaagattggg	1500
aatgatcacc atcggtatct atagtctcac catctgttac tcttccgggt aaggctccagt	1560
tctttcacc tactctgccc aagccttccc attgtacatc agagacttgg gtatgtcttt	1620
tgetactgcc acctgttggg ctttcaactt catcttggcc ttcacctgga acagattggt	1680
caatgcattc acatctactg gtgccttcgg cttctacgct gcttggaaaca tcattggttt	1740
cttcttggtc ttatggttct tgccagaaac caagggttg accttggag aattggacga	1800
agtcttcgcc gtttccgccc tccaacacgc caagtaccaa accaagatt tgatcaactt	1860
catccaaaga tacgttttac gttccaaggt ggctccattg cctccattgt acgaccacca	1920
gagattggct gtcaccaacc cagaatggaa cgacaagcca gaagtctctt atgtcgagta	1980
ggctccttga taacacatc atttatttcc tctttataat taatagttaa cttagtgtt	2040
caattcttca catcgctag atagtaa	2067

<210> SEQ ID NO 45

<211> LENGTH: 1629

<212> TYPE: DNA

<213> ORGANISM: Scheffersomyces stipitis

<400> SEQUENCE: 45

atgagctacg aagataaact cgttcaacct gccttgaagt tcaggacctt cttggacaga	60
cttccaaaca tttacaatgt gtacattatt gcatctattt cctgtatttc aggtatgatg	120
ttcggttttg atatttcatc tatgtctgct tttataggty aagatgacta caagaacttt	180
ttcaataatc caggctcaga catccaaggt tttatcactt cctgtatggc tttaggttct	240
ttcttcggtt ccatcgtctc ttccttcatt tccgaacctt ttggtagaag agcatccttg	300
ttgttgtggt cattctctctg gatggtcggg gctgctgtac aatcatcttc tcaaaacaga	360
gcccaattga tgatcggacg tatcatcgtc ggttccggtg ttggttttgg ttctctctgt	420

-continued

gctccagttt acggttccga attggtccca agaaagatta gaggttttgt tgggtgtatt	480
ttccaattct gtgttacctt gggatatctg attatgttct acatttgta cggtttgc	540
ttcattaacg gtgttgctc tttcagaatt gcttggggtt taaaaattgt cccaggttg	600
gttttatttg tcggttgttt ctttattcca gaatcccaa gatggttagc caaacatggt	660
tactgggatg aagcagaatt catcggtgcc caaattcaag ctaagggtaa tagagaagac	720
ccagacgtgt tgattgaaat ctccgaaatc aaggacaaa ttttgattga agaaaacctc	780
aagagtttcg gttacgttga cttattcacc aagaagtata tcagaagaac ttaactgcc	840
atatttgctc aaatctggca acaattgacc ggtatgaatg tcatgatgta ctatattgtc	900
tacattttca acatggcggg ttactctaac aacgcaaaact tggttgcctc tccatccaa	960
tacgtcttga aactgctgc aactgttcca gctttgtttt taatggatta cattggtaga	1020
agaagattgt tgattggtgg tgccatcatg atgatgattt tccaatttgg tgttgetggt	1080
atcttaggta aatactcogt ccccgttcca ggcggtcttc caggtaaacc aactgtcacc	1140
atccaaatcc cagaagataa caagtcagct gctagaggtg ttattgcttg ttgttactta	1200
ttcgttgtat cattcgctct gagttggggt gtcggtatct gggctactctg ttcagaagtt	1260
tggggtgact ctgctccag acagagaggt gctgctgttt caactgctgc caactggatt	1320
cttaactttg ctattgcoat gtacactcca tcttccttca agaatacac ctggaagact	1380
tacatcatct acgccgtctt ctgtcttgtt atggcaatcc atgtctactt tggattccca	1440
gaaaccaagg gcaagcgttt ggaagaagtc ggacaaatgt gggacgaaaa tgttcccgca	1500
tggagatctt ccagctggca accaactgtt ccattgttgt cagatgccga cttggcacac	1560
aagatggatg tttcccacaa ggaagagcaa tctccagatg ccgagtcaag ttctgaggaa	1620
aagccttaa	1629

<210> SEQ ID NO 46

<211> LENGTH: 1673

<212> TYPE: DNA

<213> ORGANISM: Scheffersomyces stipitidis

<400> SEQUENCE: 46

acaaaaatgg gtttgaaga cagtgtcttc ttgcaaaagt acatcaactt cggtgaaaag	60
aaggctggtt ccaccaccat gggatatctgt gttggtttgt tgcgagcctt cggtggtatc	120
cttttcggtt atgacactgg taccatctcc ggtatcatgg ccatggacta cgtoactgcc	180
agattcccat ccaaccacca atctttcagt tcttctgaat ctcccttat tgtttccatt	240
ttgtctgttg gtacctctt tggttctctt tctgcatctt tcatctccga cagattgggt	300
cgtagattga ctttaatgat ctccacctg atcatcttca atgtcggat tatcttgcaa	360
actgccteta ctagcattcc acttttgtgt gttggtagag ttcttctggt tcttgggtgt	420
ggctctcatt ccgctgttat tccattgtac caagctgaaa cagttccaaa gtggatcaga	480
ggtgctgttg tctcctgtta ccaatgggcc attacccttg gtttgtgtt ggctgctgtt	540
gttaaccaag gtaccacaa cagaaatgac tctggttctt acagaatccc aattgctatc	600
caattcttgt gggctttgat tttgggaggt ggtatgtgtt tggttccaga aacccaaga	660
ttctgggttt ctaaagggtga caacgacaga gccaaaggact ccttgagaag attgagaag	720
tgccctctg accatccga cttgattgaa gaatacgaag aaatcaaggc taactacgaa	780
tacgaagctc aatacggctc aggttcttgg agtcaagttt ttgctaacaa gaaccaccaa	840
agaaagagat tggccatggg tgttggatc caagccttgc aacaattgac cggtattaac	900

-continued

```

tttatcttct actatggtac taacttcttc aagggttctg gtatcaaaaa cgaattcctt   960
atccaaatgg ccactaacat tgtcaacttc ggttctactg tcccagggtat tcttttggtt  1020
gaaattattg gtagaagaaa gttggtgttg ggtggttctg cagttatgtc cattttctcaa  1080
ttgattgttg ctattgtcgg tgttgccgct ggtgaaggtt caacttctgc caacaagtgt  1140
ttggttgcc tctgttggat cttcattgct gctttcgcag ccacttgggg tctcttttgt  1200
tgggctgtca ttgccgaatg ttaccactt acagttagac aaaagtccat ctctctgtgt  1260
acagcttcca actggtgtg gaactggggt attgcctacg ctactcetta catggccaac  1320
tccggtccag gtaacgcaa cttgggttcc aaggtttct tcatctgggg tggttgtaat  1380
atcattggtg gtcttttctg gtggtaacct gtctacgaaa ctaagggctt gacctagaa  1440
caaatcgatg aaatgtacga aaaggttcca aagccttggc aatctaccag attcattcca  1500
tccgaacatg cttcactca accatccgca gctgcctctg tctcttctgg taaggctgaa  1560
ggtgtttctg aagttgaaga agcttctgta tagatacctc atgcacatta atcgacttac  1620
gtctttatga attttacttt ctcttaatat aatattatgg ccaattaacg ttg       1673

```

<210> SEQ ID NO 47

<211> LENGTH: 1650

<212> TYPE: DNA

<213> ORGANISM: Debaryomyces hansenii

<400> SEQUENCE: 47

```

atggctttaa aaatcttttc tagaaccaac actatggggt taaggggtaa acgtcttaga   60
gtaatgttca ctggtgtggc tactcttggg ttctccttat tgggttacga tcaaggttta  120
atgtctggtc ttattactgg tgagcaattc aatgctgaat tccctccaac agcaggtaa  180
gatcactggg cttctgttaa tcaaggtgcc gttaccgctt gttacgaaat tggatgtttg  240
tttggtgctt tatttgtttt attctatggt gataaaacag gtagaagaat tttagtgttt  300
tgcggttctt tgattattat tattggtacc gtcatttcta ctgcccatt tgggtccaaa  360
tggggtttag gtcagtttgt tgtcggaga gtggttacag gtgtaggtaa tggtttgaat  420
accgctacaa ttccagtttg gcaatctgaa atgtctaaag ccgaaaacag aggtctttta  480
gtcaacttgc aaggttctgt cattgcagtg ggtacctttg ttgcttactg gatcgatttc  540
ggattatctt atgttgatag ttctgtocaa tggagattcc cagttgcttt ccaagcatta  600
tttgcactct tcttgttatt tggagctatt gaaatgccag aatctccaag atggatgttc  660
gctcacgata tgaagctga aggtatggaa gtttttagctg caatgaaaga tatttctcca  720
gatgatgatg aaatctacgc agaataacg ttcattactg actcgattaa gagattcgat  780
aataaccaag ccggattcaa ggaattattc aaaggtggta aagagcaata ctttgctaga  840
atgataattg gttcatctgg tcaatttttc caacaattta ctggttgtaa tgcggcaatt  900
tattattcta ccgtgttatt tgaagatact attcatttag aaagaagatt ggctttgatt  960
ttaggtggtg tttttgcaac cgtctatgcc ttateccaaa ttccttcggt cttcttggtt  1020
gatacacttg gtagaagaaa cttattctta attggtgcaa ttggtcaagc tatctcattt  1080
acaattacat ttgcctgttt gattccagaa gatggagaaa aactcaaga tgccaaaggt  1140
gctgctgttg gtcttttctt gtttattgct tttttcgggt ttactatctt accaatgcct  1200
tggatttatc caccagaaat caatccaatg aagacaagaa ccgtagcttc agctgtttcc  1260
acctgtacta actggttaac taatttcggg gttgttatgt tcacccaat tttcattgca  1320

```

-continued

```

caaagtacct ttggatgta ttattcttt gcgcttatga attatacttt cattccaatt 1380
atthttctct tttatccaga aactgcoggt cgttccttag aagagattga tatcattttc 1440
gccaaagcac atgttgacaa tagattacca ttcagagttg cagctactat gccacgttta 1500
tctgtcaaag atattgaaga atacaatgtt caattaggtc ttgatgatga ctctgataag 1560
gaacagaatg aattgcaaga gaatgcctcc tccaattcag aaaaatcacc tgatgacact 1620
ccagaaggta tcttaactcc taatgcttaa 1650

```

```

<210> SEQ ID NO 48
<211> LENGTH: 1650
<212> TYPE: DNA
<213> ORGANISM: Debaryomyces hansenii

```

```

<400> SEQUENCE: 48

```

```

atgtataaaa tatgggtcaaa aactaacact atgggactca gaggtaaacc tcttagagtt 60
gccatcacta tatgctgtac tattggattt tcattattcg gttacgacca aggattaatg 120
tcgggaatta ttactggtaa acaattcaat gaggaatttc cacccactca cggtagagat 180
cagcatgcta ctgttataca aggtgcagtc acgtcgtgct acgagttagg ctgttttttt 240
gggtctttgt ttgcggtggt tcaagtgat aagtatggta gaaggccaat gattattggt 300
ggatctctgc tcattgtcat aggtactgtt attgcagttt ctgcatttgg tccgcaatgg 360
ggattgggcc aatttgttat tggtagagtt atcacaggat taggtaacgg catggatacg 420
gctactattc cgggttgcca atcggaaatt tctaaggcag aaaatagagg tcttttggtt 480
aacttgaag gttccatggt tgcggttggg acatttattg catattggtt agatttggg 540
ttatcatatg tcgatacttc agttcagtg agattcccag ttgcttttca gatagtattt 600
gctctttct tgtttcttgg cgttgcaaaa ttaccggaat cacctagatg gtttaattgct 660
catggcctca aggatgaagc tcattatgta ttggcaactt taaatgacgt tgatattgat 720
gatgagttcg taattgaaga aagcgcatt ataactgatg gtgtcaatag atttgccaga 780
acccaaattg ggttcaaaga actatthtcc ggcggcaagc aacagaactt cgctagaatg 840
ataattggtg catctacaca attctttcag caatttactg gatgtaacgc tccatctat 900
tattcaactg ttttatttga aaatagtatt ggattgaccg gtaaattgcc cttaatttca 960
ggaggtgttt ttgctaccat ttatgcttta tctactatc catctttctt cttgattgat 1020
aggctaggca gaagagcttt gtttttgatc ggtgccaccg gtcaaggcat atcattcacc 1080
atcacttttg catgcttgat ccttgataat ggacaaaata aggaacacgc caaaggtgct 1140
gccgttggtt tcttcttgtt tatcgttttt tttgcgttca ctattttacc attgccatgg 1200
atthaccctc cagaaatcaa ccccttgaga acaagaacag ttgctactgc ggtatcaaca 1260
tgtactaatt ggtaacaaaa ctttgcagtc gttatgttta ctctattht tattggtgca 1320
agtagttatg gctgttactt attctttgct ataatgaatt tcctthtcat tctgttcatt 1380
ttctggtttt atcccagac tgctggtcgt gagttggaag agattgatat tathttcgca 1440
aaggcatacg ttgataatag actaccatgg agagttgctg ctactttacc acactgaaat 1500
ttcaaggaac aagaacaaga ggaatcaaaa ttgggtcttt atgatgactt cgaaggaa 1560
gcacctgaac atgtaaacac tttgtccgac gctcaggct cagaagatgg ctcaaatgaa 1620
gccgcttcag ttaagcctgc ggaagttaa 1650

```

```

<210> SEQ ID NO 49
<211> LENGTH: 1980

```

-continued

<212> TYPE: DNA

<213> ORGANISM: *Yarrowia lipolytica*

<400> SEQUENCE: 49

```

atggtttttg gacgagaaaa agacgactct gagggtatcg aacatgtgcc ctccccccaa    60
gacaaccctg cagatcagac gtcagacatc atcgctctga acgagaaggc ttctaataaa    120
catgatgatc tccccactat ccccaaaccc gagggagatg ccccagtcac ctctgaactg    180
gaccccgaca atccgctgat tcgatacagt cgcgctgagc tccctcgagat tgccaccag    240
tttgcggtgg ataacgaact tgccgacaaa gcggaagcat tcagaaaggg cgccttagtg    300
gcccagaacc cttccggttt cgagaacatc gacatacttg atgacgacga cagatactgg    360
cttaaccctg agatcaccaa taagtgggat catcccatga aggtttacta ccttgttgtg    420
tgctgctcct tggtgctgc tgtccagggt atggatgaga ctgttatcaa cggcgccaac    480
atcattttcc ctgctcagtt tggtatcaag gaggattccg gtgttgtgtc tcgaaagagt    540
tggtgctctg gtcttgtcaa ctctgctcct tatttgtgtt gtgcatgcat ctctgctctg    600
atgactgacc ctatcaacaa agtacttggc cgaaaatgga cagtgttctg gacctgttc    660
tgggccggag ccacctgttt ctggtctggt ttcgtcaaca cctggtggca tctgtttatt    720
gccccgttct tcctgggatt cgttattggt cccaagtccg ccacagtcc cgtgtacgct    780
gcccagtggt ctctctccag gattcgaggt gccatgggta tgatgtggca gatgtggact    840
gcttttggtg tcatgatggg ctatgttatg gatcttgcac tctactacgt caaggataga    900
ggaactattg tcggcctaaa ttggagactg atgcttggtt ctgctttgat tcctgctcct    960
ctggtctgta tttttattgt caaatgtcct gagtctccca gatggcacct cgtctgagga   1020
gagatccgaa agtcggttga gtgcatgcca gaaattcgac aactgacat acaagccgct   1080
cgagatacct tctacgcccc cgttcttctg atcgaagaga acgagatgaa gaaaggaaa    1140
aacgatttgg tggagctctt tacctgtcct cgaaaaccggc gagcagcctg ggcttctttc   1200
attgtcatgt tcatgcagca gttctgtggt atcaatgtta ttgcctacta ctctccaac    1260
atcttcatgg agtctggttt tggtgctatc caagctcttc tggtctcgtt tggttttggt   1320
gctatcaact ttgtgtttgc gttgccagct gtttacacta tcgacacatt tggtagacgg    1380
gcactattac tggcgacctt cctctgatg gctatattct tgctatttgc tggtttctgt    1440
ttctacattg gccagaacga tcccaccacc tctcatgctc gtgtcgggtt aattgctcta   1500
ggatctatc tcttctcagc agtttactcc tgtggagaag gtccagtgcc cttcacctac   1560
tcggctgaag ccttcccttt gtacgttcga gatttgggta tgcgtttgc caccgcccgt    1620
tgctggctct ttaattttgt tctagccgct acgtggccct ctctcctggc agccttcaact   1680
ccgcaggggt ccttcggatg gtatgctgcg tggaatgttg tcggattctt cttggtctctg   1740
tgtttcttgc ccgagaccaa gaacctgact ctggaagagc ttgacaaggt cttcagtgtc    1800
cccaccgag tccacatgaa gtaccagttc aacgccttca aaatcaacat tcagcgaaca   1860
atattacgaa aggatgtgcc caagcctcct ccgctctatg cccacgaagc cggatttgggt   1920
ggtaactctc actggagctc caagcctcag cccaacgcga aactgcccga gttcgtttaa   1980

```

<210> SEQ ID NO 50

<211> LENGTH: 1587

<212> TYPE: DNA

<213> ORGANISM: *Scheffersomyces stipitidis*

<400> SEQUENCE: 50

-continued

```

atggcatatc ttgattgggt aacagctaga accaacaactt tcgggttgag gggcaagaag    60
ttgagagcct tcatcactgt agtgggtgtc actggtttct cattattcgg atatgatcaa    120
gggttgatgt ccggaattat tactgctgat caattcaact ctgagtttcc cgccactaga    180
aataacagta ctatccaagg tgccgtcacc tctctttaag agcttggttg tttctttggt    240
gctgtgtttg ccttgtaag aggtgaaaga attggaagaa gacctctgt gctttgtggc    300
tcgcttatta tcactctggg aacagttatt tctgtaaccg ccttccatcc aactgggtca    360
ttaggtcagt ttgttattgg tagagttatc actggtattg gtaatggtat gaatactgcc    420
accattccag tttggcaatc ggaatgtca agagctgaaa acagaggaag attggccaac    480
ttggaaggtt ccgttgctgc tgtgggtaca tgtattgcct actggttggg tttcggtttg    540
tcttatgtcg acaattcagt ttctggaga tttccagttg ctttccaaat agtgtttgct    600
tccgttttat ttgtgggaat gttgcaattg cccgactctc caagatggtt ggttgetaac    660
cacagaagag cagaggtctc tcaagtgtg tctgcttga aagacttgc cgaagacgac    720
gaagaaatc ttaatgaagc tgaagtatt caggaaagt tagacaagtt tgctggacat    780
gcttccgcca aggaagtgt tactgtgtg aagaccagc actggcaaag aatggttatt    840
ggatccagca cccaattctt tcagcagttc actggttgta acgctgccat ttactattcc    900
actgtcttgt ttcaagacac tattggttta gaaagaagaa tggcattgat tatcggtggt    960
gttttcgcaa ctgtctacgc cattttcaca attccttctt tcttcttggg cgatactctt   1020
ggacgtagaa acttgttctt gattggtgct atgggacaag gtattgcatt cactatcacc   1080
tttgctgtt tgattgacga tactgaaaac aacgccaagg gtgcccagc tggtttattc   1140
ttgtttattt gtttcttgc cttcaccatc ttgccattgc catgggtata cccaccagaa   1200
atcaatcctt tgagaactag aactatagct tctgcaattt ccacttgta ccaactggatc   1260
tgtaactttg ctgttggtat gttcaccctt gtctttgtca ctaacaccag atggggagcc   1320
tatctttctt ttgctgtgat gaacttctt ttcgttccta ttattttctt cttctaccca   1380
gaaacagctg gaagatcgtt ggaagaaatc gatatcatct ttgcgaaggc attcgttgac   1440
aaaagacagc catggagagt tgctgcaacc atgccaaagt tgtccaacca cgaattgaa    1500
gacgaagcca acagattggg cttgtttgac gatggtacat tcgacaagga agcatttgaa    1560
accaaagaaa acgcatccag cagctct                                     1587

```

```

<210> SEQ ID NO 51
<211> LENGTH: 51
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

```

```

<400> SEQUENCE: 51

```

```

Lys Phe Arg Asn Phe Leu Asp Lys Thr Pro Asn Ile Tyr Asn Val Phe
 1             5             10             15
Val Ile Ala Ser Ile Ser Cys Ile Ser Gly Leu Met Phe Gly Ile Asp
 20             25             30
Ile Ser Ser Met Ser Leu Phe Ile Gly Asp Asp Lys Tyr Ile Lys Tyr
 35             40             45
Phe His Lys
 50

```

```

<210> SEQ ID NO 52
<211> LENGTH: 51

```


-continued

<212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 52

Lys Leu Arg Leu Phe Leu Asp Lys Leu Pro Asn Ile Tyr Asn Ile Tyr
 1 5 10 15
 Val Ile Ala Thr Ile Ser Cys Ile Ser Gly Leu Met Phe Gly Ile Asp
 20 25 30
 Ile Ser Ser Met Ser Ala Phe Leu Ser Asn Asp Ala Tyr Leu Lys Tyr
 35 40 45
 Phe Gly Thr
 50

<210> SEQ ID NO 53
 <211> LENGTH: 51
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 53

Lys Phe Arg Asn Phe Leu Asp Lys Phe Pro Asn Ile His Asn Val Tyr
 1 5 10 15
 Ile Val Val Gly Ile Ser Cys Ile Ser Gly Met Met Phe Gly Ile Asp
 20 25 30
 Ile Ser Ser Met Ser Leu Phe Ile Gly Asp Asp Lys Tyr Leu Asp Tyr
 35 40 45
 Phe Asn Ser
 50

<210> SEQ ID NO 54
 <211> LENGTH: 51
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 54

Lys Phe Arg Thr Phe Leu Asp Arg Leu Pro Asn Ile Tyr Asn Val Tyr
 1 5 10 15
 Ile Ile Ala Ser Ile Ser Cys Ile Ser Gly Met Met Phe Gly Phe Asp
 20 25 30
 Ile Ser Ser Met Ser Ala Phe Ile Gly Phe Asp Asp Tyr Lys Asn Phe
 35 40 45
 Phe Asn Asn
 50

<210> SEQ ID NO 55
 <211> LENGTH: 51
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 55

Ser Leu Asn Lys Phe Leu Asp Lys Phe His Thr Thr Tyr Asn Ile Tyr
 1 5 10 15
 Val Ile Ala Met Ile Thr Thr Ile Ser Gly Met Met Phe Gly Phe Asp
 20 25 30

-continued

Val Ser Ser Ile Ser Ala Phe Ile Ser Glu Pro Ser Tyr Arg Arg Phe
 35 40 45

Phe Asn Tyr
 50

<210> SEQ ID NO 56
 <211> LENGTH: 51
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 56

Gln Val Gly Ala Leu Gln His Arg Phe Pro Lys Leu His Asn Pro Tyr
 1 5 10 15

Leu Thr Ala Ala Val Ala Thr Met Gly Gly Leu Leu Phe Gly Phe Asp
 20 25 30

Ile Ser Ser Val Ser Ala Phe Val Asp Thr Lys Pro Tyr Lys Glu Tyr
 35 40 45

Phe Gly Tyr
 50

<210> SEQ ID NO 57
 <211> LENGTH: 43
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 57

Met Tyr Lys Val His Asn Pro Tyr Leu Thr Ala Ala Val Ala Thr Met
 1 5 10 15

Gly Gly Met Leu Phe Gly Phe Asp Ile Ser Ser Val Ser Ala Phe Val
 20 25 30

Gly Phe Asp Asn Tyr Met Asn Tyr Phe Gly His
 35 40

<210> SEQ ID NO 58
 <211> LENGTH: 52
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 58

Met Gly Arg Ile Thr Asn Pro Tyr Val Leu Thr Ala Leu Ala Cys Thr
 1 5 10 15

Gly Gly Leu Leu Phe Gly Phe Asp Ile Ser Ser Met Ser Ala Ile Ile
 20 25 30

Ser Ser Pro Asn Tyr Leu Thr Tyr Phe Gly Pro Lys Asp Leu Thr Val
 35 40 45

Glu Cys Pro Asp
 50

<210> SEQ ID NO 59
 <211> LENGTH: 55
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 59

-continued

Leu Ala Ser Asp Ala Pro Glu Ser Phe Ser Trp Ser Ser Val Ile Leu
 1 5 10 15
 Pro Phe Ile Phe Pro Ala Leu Gly Gly Leu Leu Phe Gly Tyr Asp Ile
 20 25 30
 Gly Ala Thr Ser Gly Ala Thr Leu Ser Leu Gln Ser Pro Ala Leu Ser
 35 40 45
 Gly Thr Thr Trp Phe Asn Phe
 50 55

<210> SEQ ID NO 60
 <211> LENGTH: 56
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 60

His Val Pro Glu Asn Tyr Ser Val Val Ala Ala Ile Leu Pro Phe Leu
 1 5 10 15
 Phe Pro Ala Leu Gly Gly Leu Leu Tyr Gly Tyr Glu Ile Gly Ala Thr
 20 25 30
 Ser Cys Ala Thr Ile Ser Leu Gln Glu Pro Met Thr Leu Leu Ser Tyr
 35 40 45
 Tyr Ala Val Pro Phe Ser Ala Val
 50 55

<210> SEQ ID NO 61
 <211> LENGTH: 56
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 61

Leu Asn Ala Glu Ala Thr Asn Lys Trp His Ile Pro Pro Arg Leu Ile
 1 5 10 15
 Gly Val Ile Ala Leu Gly Ser Met Ala Ala Ala Val Gln Gly Met Asp
 20 25 30
 Glu Ser Val Ile Asn Gly Ala Asn Leu Phe Tyr Pro Lys Ala Phe Gly
 35 40 45
 Val Asp Thr Met His Asn Ser Asp
 50 55

<210> SEQ ID NO 62
 <211> LENGTH: 60
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 62

Leu Asn Arg Phe Ile Thr Asn Lys Trp Asp His Pro Met Lys Val Tyr
 1 5 10 15
 Tyr Leu Val Val Cys Cys Ser Leu Ala Ala Ala Val Gln Gly Met Asp
 20 25 30
 Glu Thr Val Ile Asn Gly Ala Asn Ile Ile Phe Pro Ala Gln Phe Gly
 35 40 45
 Ile Lys Glu Asp Ser Gly Val Val Ser Arg Lys Ser
 50 55 60

-continued

<210> SEQ ID NO 63
 <211> LENGTH: 55
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 63

Phe Leu Gly Met Arg Gly Ile Lys Leu Asn Trp Ala Ile Gly Phe Ala
 1 5 10 15
 Ala Ser Ala Gly Phe Leu Leu Phe Gly Tyr Asp Gln Gly Val Leu Gly
 20 25 30
 Ser Leu Tyr Thr Leu Pro Ser Trp Asn Ala Gln Phe Pro Glu Ile Asn
 35 40 45
 Thr Ala Ala Val Gly Asp Ser
 50 55

<210> SEQ ID NO 64
 <211> LENGTH: 53
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 64

Ala Lys Thr Asn Ser Tyr Leu Gly Leu Arg Gly His Lys Leu Asn Phe
 1 5 10 15
 Ala Val Ser Cys Phe Ala Gly Val Gly Phe Leu Leu Phe Gly Tyr Asp
 20 25 30
 Gln Gly Val Met Gly Ser Leu Leu Thr Leu Pro Ser Phe Glu Asn Thr
 35 40 45
 Phe Pro Ala Met Lys
 50

<210> SEQ ID NO 65
 <211> LENGTH: 53
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 65

Lys Thr Asn Thr Met Gly Leu Arg Gly Lys Pro Leu Arg Val Ala Ile
 1 5 10 15
 Thr Ile Cys Cys Thr Ile Gly Phe Ser Leu Phe Gly Tyr Asp Gln Gly
 20 25 30
 Leu Met Ser Gly Ile Ile Thr Gly Lys Gln Phe Asn Glu Glu Phe Pro
 35 40 45
 Pro Thr His Gly Thr
 50

<210> SEQ ID NO 66
 <211> LENGTH: 53
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 66

Arg Thr Asn Thr Met Gly Leu Arg Gly Lys Arg Leu Arg Val Met Phe
 1 5 10 15
 Thr Val Val Ala Thr Leu Gly Phe Ser Leu Phe Gly Tyr Asp Gln Gly

-continued

20 25 30
 Leu Met Ser Gly Leu Ile Thr Gly Glu Gln Phe Asn Ala Glu Phe Pro
 35 40 45
 Pro Thr Ala Gly Lys
 50

<210> SEQ ID NO 67
 <211> LENGTH: 51
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 67

Arg Thr Asn Thr Phe Gly Leu Arg Gly Lys Lys Leu Arg Ala Phe Ile
 1 5 10 15
 Thr Val Val Ala Val Thr Gly Phe Ser Leu Phe Gly Tyr Asp Gln Gly
 20 25 30
 Leu Met Ser Gly Ile Ile Thr Ala Asp Gln Phe Asn Ser Glu Phe Pro
 35 40 45
 Ala Thr Arg
 50

<210> SEQ ID NO 68
 <211> LENGTH: 55
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 68

Arg Thr Ser His Trp Gly Leu Thr Gly Lys Lys Leu Arg Tyr Phe Ile
 1 5 10 15
 Thr Ile Ala Ser Met Thr Gly Phe Ser Leu Phe Gly Tyr Asp Gln Gly
 20 25 30
 Leu Met Ala Ser Leu Ile Thr Gly Lys Gln Phe Asn Tyr Phe Phe Pro
 35 40 45
 Ala Thr Lys Glu Asn Gly Asp
 50 55

<210> SEQ ID NO 69
 <211> LENGTH: 49
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 69

Ile Asp Val Gly Leu Arg Gly Asn Trp Leu Leu Thr Val Ile Thr Ala
 1 5 10 15
 Ser Cys Ala Ala Gly Phe Leu Leu Phe Gly Tyr Asp Asn Gly Val Met
 20 25 30
 Gly Gly Val Val Gly Leu Gly Glu Phe Asn Lys Thr Phe Asn Asn Pro
 35 40 45
 Asp

<210> SEQ ID NO 70
 <211> LENGTH: 44
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:

-continued

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 70

Gly Lys Gln Val Ser Tyr Ala Val Thr Phe Thr Cys Glu Leu Ala Phe
 1 5 10 15
 Ile Leu Phe Gly Ile Glu Gln Gly Ile Ile Gly Asn Leu Ile Asn Asn
 20 25 30
 Gln Asp Phe Leu Asn Thr Phe Gly Asn Pro Thr Gly
 35 40

<210> SEQ ID NO 71

<211> LENGTH: 52

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 71

His Lys Thr Gln Arg Arg Leu Val Gly His Asn Leu Leu Tyr Ser Val
 1 5 10 15
 Ser Val Phe Leu Ser Ile Gly Gly Val Leu Phe Gly Tyr Asp Gln Gly
 20 25 30
 Val Met Ser Gly Ile Ile Thr Gly Pro Tyr Phe Lys Ala Tyr Phe Asn
 35 40 45
 Gln Pro Thr Ser
 50

<210> SEQ ID NO 72

<211> LENGTH: 49

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 72

Met Phe Ser Leu Thr Gly Lys Pro Leu Leu Tyr Phe Thr Ser Val Phe
 1 5 10 15
 Val Ser Leu Gly Val Phe Leu Phe Gly Tyr Asp Gln Gly Val Met Ser
 20 25 30
 Gly Ile Ile Thr Gly Phe Tyr Phe Lys Phe Tyr Phe His Glu Pro Thr
 35 40 45

Arg

<210> SEQ ID NO 73

<211> LENGTH: 55

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 73

Val Gly Ala Thr Gly Ala Lys Gly Leu Ile Lys Asn Ala Arg Thr Phe
 1 5 10 15
 Ala Ile Ala Val Phe Ala Ser Met Gly Gly Leu Ile Tyr Gly Tyr Asn
 20 25 30
 Gln Gly Met Phe Gly Gln Ile Leu Ser Met His Ser Phe Gln Phe Ala
 35 40 45
 Ser Gly Val Lys Gly Ile Thr
 50 55

-continued

<210> SEQ ID NO 74
 <211> LENGTH: 56
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 74

Ala Gly Lys Ser Gly Val Ala Gly Leu Val Ala Asn Ser Arg Ser Phe
 1 5 10 15
 Phe Ile Ala Val Phe Ala Ser Leu Gly Gly Leu Val Tyr Gly Tyr Asn
 20 25 30
 Gln Gly Met Phe Gly Gln Ile Ser Gly Met Tyr Ser Phe Ser Lys Ala
 35 40 45
 Ile Gly Val Glu Lys Ile Gln Asp
 50 55

<210> SEQ ID NO 75
 <211> LENGTH: 54
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 75

Ala His Gly Asn Val Val Thr Ile Met Met Lys Asp Pro Val Val Phe
 1 5 10 15
 Leu Val Ile Leu Phe Ala Ser Leu Gly Gly Leu Leu Phe Gly Tyr Asp
 20 25 30
 Gln Gly Val Ile Ser Gly Ile Val Thr Met Glu Ser Phe Gly Ala Lys
 35 40 45
 Phe Pro Arg Ile Phe Met
 50

<210> SEQ ID NO 76
 <211> LENGTH: 54
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 76

Ala His Gly Asn Val Val Thr Ile Met Met Lys Asp Pro Val Val Phe
 1 5 10 15
 Leu Val Ile Leu Phe Ala Ser Leu Gly Gly Leu Leu Phe Gly Tyr Asp
 20 25 30
 Gln Gly Val Ile Ser Gly Ile Val Thr Met Glu Ser Phe Gly Ala Lys
 35 40 45
 Phe Pro Arg Ile Phe Met
 50

<210> SEQ ID NO 77
 <211> LENGTH: 54
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 77

Ala His Gly Asn Val Val Thr Ile Met Met Lys Asp Pro Val Val Phe
 1 5 10 15
 Leu Val Ile Leu Phe Ala Ser Leu Gly Gly Leu Leu Phe Gly Tyr Asp

-continued

20 25 30
 Gln Gly Val Ile Ser Gly Ile Val Thr Met Glu Ser Phe Gly Ala Lys
 35 40 45
 Phe Pro Arg Ile Phe Met
 50

<210> SEQ ID NO 78
 <211> LENGTH: 54
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 78

Ser Lys Gly Asn Ile Ile Thr Val Met Ser Lys Asp Pro Leu Val Phe
 1 5 10 15
 Cys Ile Ile Ala Phe Ala Ser Ile Gly Gly Leu Leu Phe Gly Tyr Asp
 20 25 30
 Gln Gly Val Ile Ser Gly Ile Val Thr Met Glu Ser Phe Ala Ala Lys
 35 40 45
 Phe Pro Arg Ile Phe Ser
 50

<210> SEQ ID NO 79
 <211> LENGTH: 57
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 79

Pro Ile Glu Ile Pro Lys Lys Pro Met Ser Glu Tyr Val Thr Val Ser
 1 5 10 15
 Leu Leu Cys Leu Cys Val Ala Phe Gly Gly Phe Met Phe Gly Trp Asp
 20 25 30
 Thr Gly Thr Ile Ser Gly Phe Val Val Gln Thr Asp Phe Leu Arg Arg
 35 40 45
 Phe Gly Met Lys His Lys Asp Gly Thr
 50 55

<210> SEQ ID NO 80
 <211> LENGTH: 58
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 80

Glu Val Val Val Pro Glu Lys Pro Ala Ser Ala Tyr Ala Thr Val Ser
 1 5 10 15
 Ile Met Cys Leu Cys Met Ala Phe Gly Gly Phe Met Ser Gly Trp Asp
 20 25 30
 Thr Gly Thr Ile Ser Gly Phe Val Asn Gln Thr Asp Phe Leu Arg Arg
 35 40 45
 Phe Gly Asn Tyr Ser His Ser Lys Asn Thr
 50 55

<210> SEQ ID NO 81
 <211> LENGTH: 57
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence

-continued

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 81

Ala Val Ala Pro Pro Asn Thr Gly Lys Gly Val Tyr Val Thr Val Ser
 1 5 10 15
 Ile Cys Cys Val Met Val Ala Phe Gly Gly Phe Ile Phe Gly Trp Asp
 20 25 30
 Thr Gly Thr Ile Ser Gly Phe Val Ala Gln Thr Asp Phe Leu Arg Arg
 35 40 45
 Phe Gly Met Lys His His Asp Gly Ser
 50 55

<210> SEQ ID NO 82

<211> LENGTH: 57

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 82

Val Leu Thr Asn Pro Asn Thr Gly Lys Gly Ala Tyr Val Thr Val Ser
 1 5 10 15
 Ile Cys Cys Val Met Val Ala Phe Gly Gly Phe Val Phe Gly Trp Asp
 20 25 30
 Thr Gly Thr Ile Ser Gly Phe Val Ala Gln Thr Asp Phe Leu Arg Arg
 35 40 45
 Phe Gly Met Lys His Lys Asp Gly Ser
 50 55

<210> SEQ ID NO 83

<211> LENGTH: 57

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 83

Val Val Glu Ile Pro Lys Arg Pro Ala Ser Ala Tyr Val Thr Val Ser
 1 5 10 15
 Ile Met Cys Ile Met Ile Ala Phe Gly Gly Phe Val Phe Gly Trp Asp
 20 25 30
 Thr Gly Thr Ile Ser Gly Phe Ile Asn Gln Thr Asp Phe Ile Arg Arg
 35 40 45
 Phe Gly Met Lys His Lys Asp Gly Thr
 50 55

<210> SEQ ID NO 84

<211> LENGTH: 57

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 84

Pro Ile Asp Leu Pro Gln Lys Pro Leu Ser Ala Tyr Thr Thr Val Ala
 1 5 10 15
 Ile Leu Cys Leu Met Ile Ala Phe Gly Gly Phe Ile Phe Gly Trp Asp
 20 25 30
 Thr Gly Thr Ile Ser Gly Phe Val Asn Leu Ser Asp Phe Ile Arg Arg
 35 40 45

-continued

Phe Gly Gln Lys Asn Asp Lys Gly Thr
50 55

<210> SEQ ID NO 85
<211> LENGTH: 57
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 85

Asn Ala Glu Leu Pro Ala Lys Pro Ile Ala Ala Tyr Trp Thr Val Ile
1 5 10 15
Cys Leu Cys Leu Met Ile Ala Phe Gly Gly Phe Val Phe Gly Trp Asp
20 25 30
Thr Gly Thr Ile Ser Gly Phe Val Asn Gln Thr Asp Phe Lys Arg Arg
35 40 45
Phe Gly Gln Met Lys Ser Asp Gly Thr
50 55

<210> SEQ ID NO 86
<211> LENGTH: 57
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 86

Ser Leu Asp Ile Pro Tyr Lys Pro Ile Ile Ala Tyr Trp Thr Val Met
1 5 10 15
Gly Leu Cys Leu Met Ile Ala Phe Gly Gly Phe Ile Phe Gly Trp Asp
20 25 30
Thr Gly Thr Ile Ser Gly Phe Ile Asn Gln Thr Asp Phe Lys Arg Arg
35 40 45
Phe Gly Glu Leu Gln Arg Asp Gly Ser
50 55

<210> SEQ ID NO 87
<211> LENGTH: 57
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 87

Gln Val Asp Ala Pro Gln Lys Gly Phe Lys Asp Tyr Ile Val Ile Ser
1 5 10 15
Ile Phe Cys Phe Met Val Ala Phe Gly Gly Phe Val Phe Gly Phe Asp
20 25 30
Thr Gly Thr Ile Ser Gly Phe Val Asn Met Ser Asp Phe Lys Asp Arg
35 40 45
Phe Gly Gln His His Ala Asp Gly Thr
50 55

<210> SEQ ID NO 88
<211> LENGTH: 58
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 88

-continued

Asn Val Glu Pro Pro Lys Arg Gly Leu Ile Gly Tyr Leu Val Ile Tyr
 1 5 10 15
 Leu Leu Cys Tyr Pro Ile Ser Phe Gly Gly Phe Leu Pro Gly Trp Asp
 20 25 30
 Ser Gly Ile Thr Ala Gly Phe Ile Asn Met Asp Asn Phe Lys Met Asn
 35 40 45
 Phe Gly Ser Tyr Lys His Ser Thr Gly Glu
 50 55

<210> SEQ ID NO 89
 <211> LENGTH: 58
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 89

Met Val Phe Gln Val Arg Gly Thr Pro Ile Gly Ala Leu Thr Leu Phe
 1 5 10 15
 Ile Ala Met Leu Ala Ser Met Gly Gly Phe Leu Phe Gly Trp Asp Thr
 20 25 30
 Gly Gln Ile Ser Gly Leu Thr Gln Met Ala Asp Phe Arg Gln Arg Phe
 35 40 45
 Ala Thr Val Asp Asn Pro Asp Ala Ile Gly
 50 55

<210> SEQ ID NO 90
 <211> LENGTH: 60
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 90

Gly Gln Ala Ala Lys Ile Ser His Asn Ala Ser Leu His Ile Pro Val
 1 5 10 15
 Leu Leu Cys Leu Val Ile Ser Leu Gly Gly Phe Ile Phe Gly Trp Asp
 20 25 30
 Ile Gly Thr Ile Gly Gly Met Thr Asn Met Val Ser Phe Gln Glu Lys
 35 40 45
 Phe Gly Thr Thr Asn Ile Ile His Asp Asp Glu Thr
 50 55 60

<210> SEQ ID NO 91
 <211> LENGTH: 59
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 91

Gly Pro Val Ala Arg Pro Ala Ser Val Lys Gln Ser Leu Pro Ala Ile
 1 5 10 15
 Leu Val Ala Ala Ala Ser Ala Phe Gly Gly Val Leu Phe Gly Tyr Asp
 20 25 30
 Thr Gly Thr Ile Ser Gly Leu Ile Val Met Pro Asn Phe Gln Thr Glu
 35 40 45
 Gly Lys Pro Val Pro Gly Ser Thr Thr Gly Ala
 50 55

-continued

<210> SEQ ID NO 92
 <211> LENGTH: 60
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 92

Gly Pro Val Ala Arg Pro Ala Ser Val Lys Gln Ser Leu Pro Ala Ile
 1 5 10 15
 Leu Val Ala Ala Ala Ser Ala Phe Gly Gly Val Leu Phe Gly Tyr Asp
 20 25 30
 Thr Gly Thr Ile Ser Gly Leu Ile Val Met Pro Asn Phe Gln Glu Thr
 35 40 45
 Phe Gly Lys Pro Val Pro Gly Ser Thr Thr Gly Ala
 50 55 60

<210> SEQ ID NO 93
 <211> LENGTH: 53
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 93

Phe Val Asn Val Gly Glu Lys Lys Ala Gly Ser Thr Ala Met Ala Ile
 1 5 10 15
 Ile Val Gly Leu Phe Ala Ala Ser Gly Gly Val Leu Phe Gly Tyr Asp
 20 25 30
 Thr Gly Thr Ile Ser Gly Val Met Thr Met Asp Tyr Val Leu Ala Arg
 35 40 45
 Tyr Pro Ser Asn Lys
 50

<210> SEQ ID NO 94
 <211> LENGTH: 53
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 94

Phe Val Asn Val Gly Glu Lys Lys Ala Gly Ser Thr Ala Met Ala Ile
 1 5 10 15
 Ile Val Gly Leu Phe Ala Ala Ser Gly Gly Val Leu Val Gly Tyr Asp
 20 25 30
 Thr Gly Thr Ile Ser Gly Val Met Thr Met Asp Tyr Val Leu Ala Arg
 35 40 45
 Tyr Pro Ser Asn Lys
 50

<210> SEQ ID NO 95
 <211> LENGTH: 53
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 95

Phe Val Asn Val Gly Glu Lys Lys Ala Gly Ser Thr Ala Met Ala Ile
 1 5 10 15

-continued

Ile Val Gly Leu Phe Ala Ala Phe Gly Gly Val Leu Ser Gly Tyr Asp
 20 25 30

Thr Gly Thr Ile Ser Gly Val Met Thr Met Asp Tyr Val Leu Ala Arg
 35 40 45

Tyr Pro Ser Asn Lys
 50

<210> SEQ ID NO 96
 <211> LENGTH: 53
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 96

Tyr Val Asn Val Gly Glu Lys Arg Ala Gly Ser Ala Ser Met Gly Ile
 1 5 10 15

Phe Val Gly Ala Phe Ala Ala Phe Gly Gly Val Leu Phe Gly Tyr Asp
 20 25 30

Thr Gly Thr Ile Ser Gly Ile Met Ala Met Asn Tyr Val Lys Gly Glu
 35 40 45

Phe Pro Ala Asn Lys
 50

<210> SEQ ID NO 97
 <211> LENGTH: 53
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 97

Tyr Val Asn Val Gly Glu Lys Arg Ala Gly Ser Ala Ser Met Gly Ile
 1 5 10 15

Phe Val Gly Ala Phe Ala Ala Phe Gly Gly Val Leu Phe Gly Tyr Asp
 20 25 30

Thr Gly Thr Ile Ser Gly Ile Met Ala Met Asn Tyr Val Lys Gly Glu
 35 40 45

Phe Pro Ala Asn Lys
 50

<210> SEQ ID NO 98
 <211> LENGTH: 53
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 98

Tyr Ile Asn Phe Gly Glu Lys Lys Ala Gly Ser Thr Thr Met Gly Ile
 1 5 10 15

Cys Val Gly Leu Phe Ala Ala Phe Gly Gly Ile Leu Phe Gly Tyr Asp
 20 25 30

Thr Gly Thr Ile Ser Gly Ile Met Ala Met Asp Tyr Val Thr Ala Arg
 35 40 45

Phe Pro Ser Asn His
 50

<210> SEQ ID NO 99
 <211> LENGTH: 53
 <212> TYPE: PRT

-continued

<213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 99

Ile Ile Asn Arg Gly Glu Lys Pro Glu Gly Ser Ala Phe Met Ala Ala
 1 5 10 15
 Phe Val Ala Val Phe Val Ala Phe Gly Gly Ile Leu Phe Gly Tyr Asp
 20 25 30
 Thr Gly Thr Ile Ser Gly Val Met Ala Met Pro Phe Val Lys Lys Thr
 35 40 45
 Phe Thr Asp Asp Gly
 50

<210> SEQ ID NO 100
 <211> LENGTH: 41
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 100

Met Ala Ile Ile Val Ala Val Phe Val Ala Phe Gly Gly Leu Leu Tyr
 1 5 10 15
 Gly Tyr Asp Thr Gly Thr Ile Ala Gly Ile Met Thr Met Gly Tyr Val
 20 25 30
 Lys Phe His Phe Thr Asp Phe Gly Lys
 35 40

<210> SEQ ID NO 101
 <211> LENGTH: 52
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 101

Tyr Tyr Lys Lys Met Gln Gln Lys Ser Ser Ser Ser Ala Ile Thr
 1 5 10 15
 Val Gly Leu Val Ala Ala Val Gly Gly Phe Leu Tyr Gly Tyr Asp Thr
 20 25 30
 Gly Leu Ile Asn Asp Ile Met Phe Met Thr Tyr Val Lys Asp Asn Phe
 35 40 45
 Pro Ala Asn Gly
 50

<210> SEQ ID NO 102
 <211> LENGTH: 54
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 102

Met Asn Thr Gln Tyr Asn Ser Ser Tyr Ile Phe Ser Ile Thr Leu Val
 1 5 10 15
 Ala Thr Leu Gly Gly Leu Leu Phe Gly Tyr Asp Thr Ala Val Ile Ser
 20 25 30
 Gly Thr Val Glu Ser Leu His Thr Val Phe Val Ala Pro Gln Asn Leu
 35 40 45
 Ser Glu Ser Ala Ala Asn

-continued

50

<210> SEQ ID NO 103
 <211> LENGTH: 50
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 103

Arg Ser Ile Gly Pro Leu Ile Pro Arg Asn Lys His Leu Phe Tyr Gly
 1 5 10 15

Ser Val Leu Leu Met Ser Ile Val His Pro Thr Ile Met Gly Tyr Asp
 20 25 30

Ser Met Met Val Gly Ser Ile Leu Asn Leu Asp Ala Tyr Val Asn Tyr
 35 40 45

Phe His
 50

<210> SEQ ID NO 104
 <211> LENGTH: 57
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 104

Lys Ser Met Thr Leu Lys Gln Ala Leu Leu Lys Tyr Pro Lys Ala Ala
 1 5 10 15

Leu Trp Ser Ile Leu Val Ser Thr Thr Leu Val Met Glu Gly Tyr Asp
 20 25 30

Thr Ala Leu Leu Ser Ala Leu Tyr Ala Leu Pro Val Phe Gln Arg Lys
 35 40 45

Phe Gly Thr Leu Asn Gly Glu Gly Ser
 50 55

<210> SEQ ID NO 105
 <211> LENGTH: 6
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 105

Gly Gly Leu Ile Phe Gly
 1 5

<210> SEQ ID NO 106
 <211> LENGTH: 6
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 106

Gly Gly Phe Ile Phe Gly
 1 5

<210> SEQ ID NO 107
 <211> LENGTH: 6
 <212> TYPE: PRT
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic polypeptide

-continued

<400> SEQUENCE: 107

Gly Gly Phe Ile Met Gly
1 5

<210> SEQ ID NO 108

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 108

Gly Phe Phe Ile Met Gly
1 5

<210> SEQ ID NO 109

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 109

Gly Gly Phe Ile Ser Gly
1 5

<210> SEQ ID NO 110

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 110

Gly Phe Phe Ile Ser Gly
1 5

<210> SEQ ID NO 111

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 111

Gly Gly Phe Ile Thr Gly
1 5

<210> SEQ ID NO 112

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 112

Gly Phe Phe Ile Thr Gly
1 5

<210> SEQ ID NO 113

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

-continued

<400> SEQUENCE: 113

Gly Gly Phe Leu Met Gly
1 5

<210> SEQ ID NO 114

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 114

Gly Phe Phe Leu Met Gly
1 5

<210> SEQ ID NO 115

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 115

Gly Gly Phe Leu Ser Gly
1 5

<210> SEQ ID NO 116

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 116

Gly Phe Phe Leu Ser Gly
1 5

<210> SEQ ID NO 117

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 117

Gly Gly Phe Leu Thr Gly
1 5

<210> SEQ ID NO 118

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 118

Gly Phe Phe Leu Thr Gly
1 5

<210> SEQ ID NO 119

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 119

-continued

Gly Gly Phe His Met Gly
1 5

<210> SEQ ID NO 120
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 120

Gly Phe Phe His Met Gly
1 5

<210> SEQ ID NO 121
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 121

Gly Gly Phe His Ser Gly
1 5

<210> SEQ ID NO 122
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 122

Gly Phe Phe His Ser Gly
1 5

<210> SEQ ID NO 123
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 123

Gly Gly Phe His Thr Gly
1 5

<210> SEQ ID NO 124
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 124

Gly Phe Phe His Thr Gly
1 5

<210> SEQ ID NO 125
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 125

-continued

Gly Gly Leu Val Tyr Gly
1 5

<210> SEQ ID NO 126
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 126

Gly Gly Phe Val Phe Gly
1 5

<210> SEQ ID NO 127
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 127

Gly Gly Arg Pro Thr Gly
1 5

<210> SEQ ID NO 128
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 128

Gly Phe Arg Pro Thr Gly
1 5

<210> SEQ ID NO 129
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 129

Gly Gly Thr Pro Thr Gly
1 5

<210> SEQ ID NO 130
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic polypeptide

<400> SEQUENCE: 130

Gly Phe Thr Pro Thr Gly
1 5

237

What is claimed is:

1. A recombinant xylose transporter protein comprising a xylose transporter motif sequence and at least one glucose mitigation mutation; wherein said xylose transporter motif sequence corresponds to amino acid residue positions 36, 37, 38, 39, 40, and 41 of SEQ ID NO: 1, and wherein the xylose transporter motif comprises the sequence -G-G/F-X¹-X²-X³-G-;

wherein,

X¹ is D, C, G, H, I, L, or F;

X² is A, D, C, E, G, H, or I; and

X³ is N, C, Q, F, G, L, M, S, T, or P;

and further wherein said glucose mitigation mutation is at a position corresponding to N326, T170, I171, K155, N225, S354, A361, L407, and/or N446 of SEQ ID NO: 1 wherein the xylose transporter protein is at least 95% identical to the sequence of SEQ ID NO: 1 and has xylose transporter protein activity.

2. The recombinant xylose transporter protein of claim 1, wherein said xylose transporter motif sequence is G-G-F-I-M-G- (SEQ ID NO:107), -G-F-F-I-M-G- (SEQ ID NO:108), -G-G-F-I-S-G- (SEQ ID NO:109), or -G-F-F-I-S-G- (SEQ ID NO: 110).

238

3. The recombinant xylose transporter protein of claim 1, wherein said glucose mitigation mutation is a N326H mutation.

4. The recombinant xylose transporter protein of claim 1, wherein said glucose mitigation mutation is a N326S mutation.

5. The recombinant xylose transporter protein of claim 1 further comprising an amino acid deletion.

6. The recombinant xylose transporter protein of claim 5, wherein said deletion is within a protein domain corresponding to residue 497-522 of SEQ ID NO: 1.

7. A recombinant yeast cell comprising a recombinant xylose transporter protein of claim 1.

8. The recombinant yeast cell of claim 7, wherein the growth rate of said recombinant yeast cell in a xylose-glucose growth media is at least about 50% of the growth rate of said recombinant yeast cell in a xylose growth media.

9. A nucleic acid encoding the recombinant xylose transporter protein of claim 1.

* * * * *