



US 20110129874A1

(19) **United States**

(12) **Patent Application Publication**
Tsutsumi et al.

(10) **Pub. No.: US 2011/0129874 A1**

(43) **Pub. Date: Jun. 2, 2011**

(54) **PICHIA PASTORIS DAS PROMOTER
VARIANTS**

(75) Inventors: **Noriko Tsutsumi**, Ichikawa-sh
(JP); **Shinobu Takagi**, Funabashi
(JP)

(73) Assignee: **Novozymes A/S**, Bagsvaerd (DK)

(21) Appl. No.: **13/002,631**

(22) PCT Filed: **Jul. 10, 2009**

(86) PCT No.: **PCT/EP2009/058859**

§ 371 (c)(1),
(2), (4) Date: **Jan. 4, 2011**

Related U.S. Application Data

(60) Provisional application No. 61/080,549, filed on Jul.
14, 2008.

(30) **Foreign Application Priority Data**

Jul. 11, 2008 (EP) 08160226.0

Publication Classification

(51) **Int. Cl.**

C12P 21/02 (2006.01)

C07H 21/04 (2006.01)

C12N 15/63 (2006.01)

C12N 1/19 (2006.01)

(52) **U.S. Cl.** **435/69.1**; 536/24.1; 435/320.1;
435/254.23

(57) **ABSTRACT**

The present invention relates to promoter variants in the form of an isolated polynucleotide comprising: i) a nucleotide sequence consisting of the DAS promoter sequence from *Pichia* or a functional part thereof, wherein the said DAS promoter is contained in SEQ ID NO: 1; and ii) at least one additional UAS, wherein the said UAS is comprised in SEQ ID NO: 2.

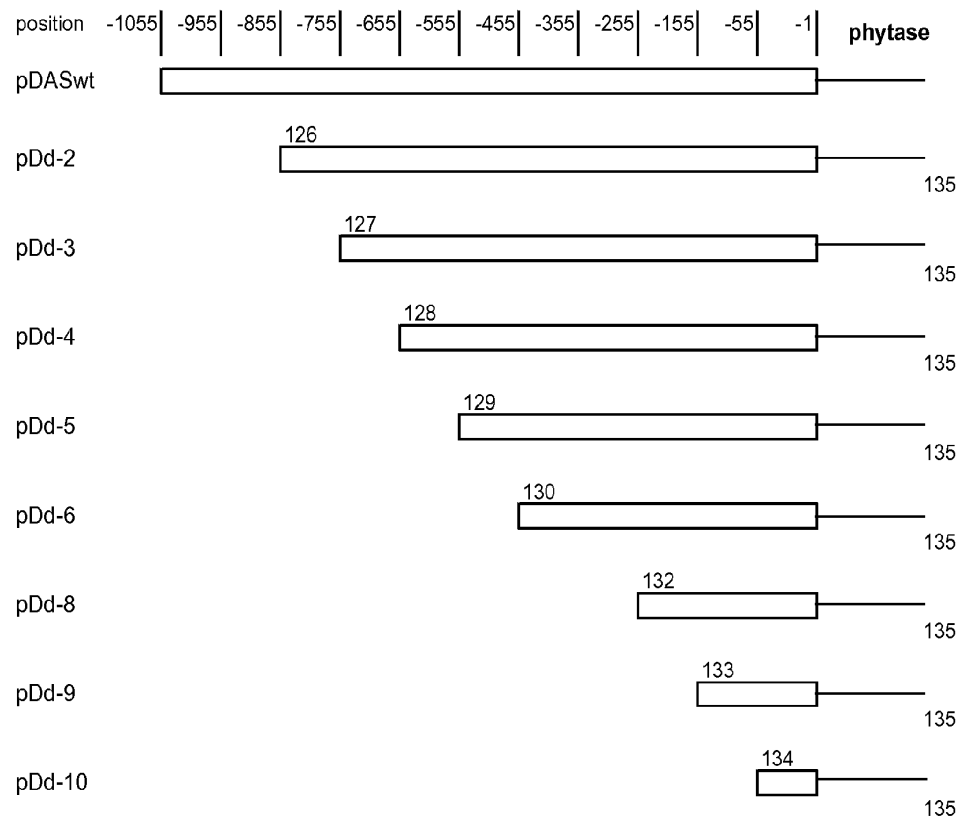


Figure 1

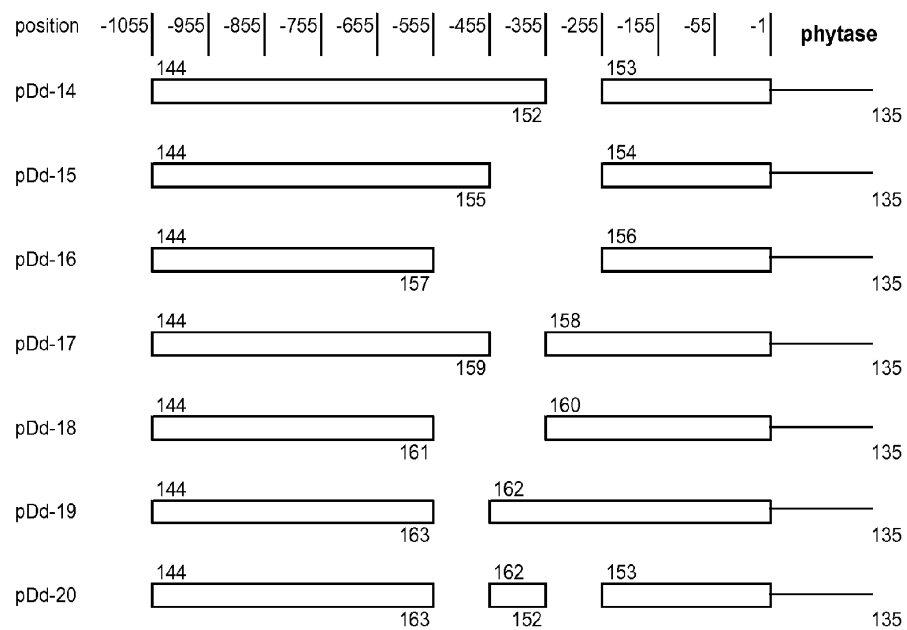


Figure 2

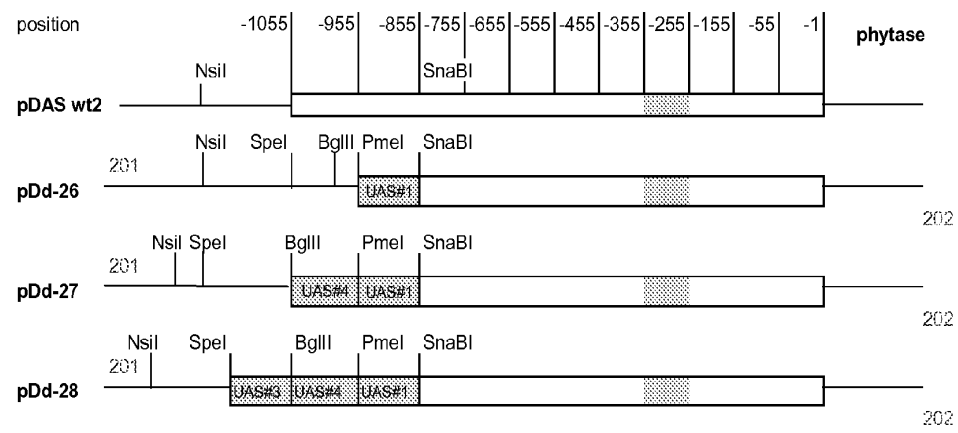


Figure 3

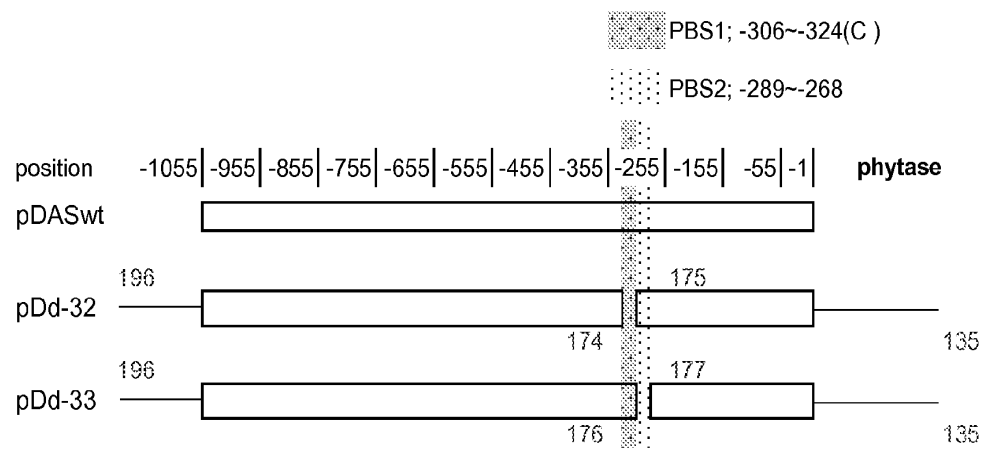


Figure 4

US 2011/0129874 A1

Jun. 2, 2011

1

PICHIA PASTORIS DAS PROMOTER VARIANTS

REFERENCE TO SEQUENCE LISTING

[0001] This application contains a Sequence Listing in computer readable form. The computer readable form is incorporated herein by reference.

FIELD OF THE INVENTION

[0002] The present invention relates to an isolated polynucleotide comprising a *Pichia* DAS promoter variant, a DNA construct comprising the promoter variant operably linked to a polynucleotide encoding a polypeptide of interest, an expression vector comprising the DNA construct, an host cell comprising the DNA construct or the expression vector, a method of producing a polypeptide of interest, a promoter comprising an UAS, and to a use of an UAS for increasing transcription.

BACKGROUND OF THE INVENTION

[0003] Eukaryotic organisms are widely used in industry as host cells for producing polypeptides for, e.g., pharmaceutical and industrial applications. The ability to manipulate gene transcription and expression gives the basis for providing higher production yields.

[0004] Conventionally, maximal expression of a gene in a eukaryotic organism is achieved by amplifying in the chromosome an expression cassette containing a single promoter operably linked to a gene encoding the polypeptide of interest and an amplifier selective marker.

[0005] Controlled expression is often desirable. In methylotrophic yeast it has been known for long that certain promoters are dependent on the presence of methanol in the growth medium for the induction of transcription. This induction by methanol requires the presence of additional factors, however, the exact mechanism of action for such factors have not been elucidated. Examples of positive factors known from yeast include Mxr1 p, described as a key positive regulator required for methanol utilization in *Pichia pastoris* (Lin-Cereghino et al., 2006, Mol Cell Biol 26(3): 883-897).

[0006] Examples of these methanol dependent promoters have been described in several yeast cells belonging to the group of yeast known as methylotrophic yeast. The promoters controlling expression of the enzymes involved in methanol metabolism in these organisms are particularly strong, and these promoters are generally used to control the heterologous expression of proteins in yeast. However, the specific carbon source used for the cultivation of these host cells has an enormous influence on the regulation of methanol metabolism promoters. Methanol and glycerol are considered as adequate substrates for methylotrophic yeast expression systems, while glucose has been considered inadequate (EP 299108). It is therefore desirable if expression from the known methanol metabolism promoters can be made less dependent on the substrate.

SUMMARY OF THE INVENTION

[0007] The invention provides improved variants of the *Pichia* DAS promoter for increased expression of a polypeptide of interest.

[0008] In a first aspect the present invention relates to an isolated polynucleotide comprising: i) a nucleotide sequence consisting of the DAS promoter sequence from *Pichia* or a

functional part thereof, wherein the said DAS promoter is comprised in SEQ ID NO: 1; and ii) at least one additional UAS, wherein the said UAS is comprised in SEQ ID NO: 2.

[0009] In a second aspect the invention relates to a DNA construct comprising a polynucleotide sequence of the invention (modified DAS promoter) operably linked to a structural gene encoding a polypeptide of interest and a terminator.

[0010] In a third aspect the invention relates to an expression vector comprising a DNA construct of the invention, further comprising a signal peptide coding region.

[0011] In a fourth aspect the present invention relates to a *Pichia* host cell comprising an expression vector of the invention.

[0012] In a fifth aspect the invention relates to a method of producing a polypeptide of interest comprising:

[0013] (a) cultivating the host cell of the invention, under conditions conducive for the production of the polypeptide of interest; and

[0014] (b) recovering the polypeptide.

[0015] In a sixth aspect the invention relates to a promoter comprising an UAS selected from the group consisting of:

[0016] i) (a) a polynucleotide comprising or consisting of SEQ ID NO: 2; or

[0017] (b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 2; or

[0018] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 2 or a full-length complementary strand thereof; or

[0019] ii) (a) a polynucleotide comprising or consisting of SEQ ID NO: 3; or

[0020] (b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 3; or

[0021] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 3 or a full-length complementary strand thereof; and

wherein the UAS according to (i) or (ii) is either foreign to the promoter or present in more than one copy.

[0022] In a seventh aspect the invention relates to a use of an UAS for increasing transcription from a promoter, wherein the UAS is selected from the group consisting of:

[0023] i) (a) a polynucleotide comprising or consisting of SEQ ID NO: 2; or

[0024] (b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 2; or

[0025] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 2 or a full-length complementary strand thereof; or

[0026] ii) (a) a polynucleotide comprising or consisting of SEQ ID NO: 3; or

[0027] (b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 3; or

[0028] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 3 or a full-length complementary strand thereof.

US 2011/0129874 A1

Jun. 2, 2011

2

gency conditions with SEQ ID NO: 3 or a full-length complementary strand thereof; and wherein the UAS according to (i) or (ii) is either foreign to the promoter or present in more than one copy.

BRIEF DESCRIPTION OF DRAWINGS

[0029] FIG. 1 shows a deletion analysis of the DAS promoter shown in SEQ ID NO: 1.

[0030] FIG. 2 shows the position of DAS promoter variants having internal deletions and the numbers of the applied primers.

[0031] FIG. 3 shows DAS promoter variants according to the invention having multiple UASs. The DNA sequence of the pDAS wt2 as shown in the figure from the NsiI site and including the region encoding the N-terminal of the phytase corresponds to SEQ ID NO: 6.

[0032] FIG. 4 shows further deletion variants of the DAS promoter having internal deletions around the -255 position.

DEFINITIONS

[0033] Identity: For purposes of the present invention, the degree of identity between two deoxyribonucleotide sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *supra*) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, *supra*), preferably version 3.0.0 or later. The optional parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled "longest identity" (obtained using the -nobrief option) is used as the percent identity and is calculated as follows:

$$\frac{(\text{Identical Deoxyribonucleotides} \times 100)}{(\text{Length of Alignment} - \text{Total Number of Gaps in Alignment})}$$

Hybridization: For purposes of the present invention, hybridization indicates that the nucleotide sequence hybridizes to a labeled nucleic acid probe corresponding to sequence in question e.g. SEQ ID NO: 7; its full-length complementary strand; or a subsequence thereof; under very low to very high stringency conditions. Molecules to which the nucleic acid probe hybridizes under these conditions can be detected using, for example, X-ray film.

[0034] For long probes of at least 100 nucleotides in length, very low to very high stringency conditions are defined as prehybridization and hybridization at 42° C. in 5×SSPE, 0.3% SDS, 200 µg/ml sheared and denatured salmon sperm DNA, and either 25% formamide for very low and low stringencies, 35% formamide for medium and medium-high stringencies, or 50% formamide for high and very high stringencies, following standard Southern blotting procedures for 12 to 24 hours optimally.

[0035] For long probes of at least 100 nucleotides in length, the carrier material is finally washed three times each for 15 minutes using 2×SSC, 0.2% SDS preferably at 45° C. (very low stringency), more preferably at 50° C. (low stringency), more preferably at 55° C. (medium stringency), more preferably at 60° C. (medium-high stringency), even more preferably at 65° C. (high stringency), and most preferably at 70° C. (very high stringency).

[0036] For short probes that are about 15 nucleotides to about 70 nucleotides in length, stringency conditions are defined as prehybridization, hybridization, and washing post-

hybridization at about 5° C. to about 10° C. below the calculated T_m using the calculation according to Bolton and McCarthy (1962, *Proceedings of the National Academy of Sciences USA* 48:1390) in 0.9 M NaCl, 0.09 M Tris-HCl pH 7.6, 6 mM EDTA, 0.5% NP-40, 1×Denhardt's solution, 1 mM sodium pyrophosphate, 1 mM sodium monobasic phosphate, 0.1 mM ATP, and 0.2 mg of yeast RNA per ml following standard Southern blotting procedures for 12 to 24 hours optimally.

[0037] For short probes that are about 15 nucleotides to about 70 nucleotides in length, the carrier material is washed once in 6×SSC plus 0.1% SDS for 15 minutes and twice each for 15 minutes using 6×SSC at 5° C. to 10° C. below the calculated T_m .

Subsequence: The term "subsequence" is defined herein as a nucleotide sequence having one or more (several) nucleotides deleted from the 5' and/or 3' end of the sequence of SEQ ID NO: 1; wherein the subsequence has promoter activity (thus a functional part thereof). In a preferred aspect, a subsequence contains at least 755 nucleotides, more preferably at least 555 nucleotides, even more preferably at least 455 nucleotides, and most preferably at least 355 nucleotides of the sequence of SEQ ID NO: 1 corresponding to positions 301-1055, 501-1055, 601-1055 and 701-1055 of SEQ ID NO: 1 respectively. Isolated polynucleotide: The term "isolated polynucleotide" as used herein refers to a polynucleotide that is isolated from a source. In a preferred aspect, the polynucleotide is at least 1% pure, preferably at least 5% pure, more preferably at least 10% pure, more preferably at least 20% pure, more preferably at least 40% pure, more preferably at least 60% pure, even more preferably at least 80% pure, and most preferably at least 90% pure, as determined by agarose electrophoresis.

Substantially pure polynucleotide: The term "substantially pure polynucleotide" as used herein refers to a polynucleotide preparation free of other extraneous or unwanted nucleotides and in a form suitable for use within genetically engineered protein production systems. Thus, a substantially pure polynucleotide contains at most 10%, preferably at most 8%, more preferably at most 6%, more preferably at most 5%, more preferably at most 4%, more preferably at most 3%, even more preferably at most 2%, most preferably at most 1%, and even most preferably at most 0.5% by weight of other polynucleotide material with which it is natively or recombinantly associated. A substantially pure polynucleotide may, however, include naturally occurring 5' and 3' untranslated regions, such as promoters and terminators. It is preferred that the substantially pure polynucleotide is at least 90% pure, preferably at least 92% pure, more preferably at least 94% pure, more preferably at least 95% pure, more preferably at least 96% pure, more preferably at least 97% pure, even more preferably at least 98% pure, most preferably at least 99%, and even most preferably at least 99.5% pure by weight. The polynucleotides of the present invention are preferably in a substantially pure form, i.e., that the polynucleotide preparation is essentially free of other polynucleotide material with which it is natively or recombinantly associated. The polynucleotides may be of genomic, cDNA, RNA, semisynthetic, synthetic origin, or any combinations thereof.

cDNA: The term "cDNA" is defined herein as a DNA molecule that can be prepared by reverse transcription from a mature, spliced, mRNA molecule obtained from a eukaryotic cell. cDNA lacks intron sequences that are usually present in the corresponding genomic DNA. The initial, primary RNA transcript is a precursor to mRNA that is processed through a

US 2011/0129874 A1

Jun. 2, 2011

3

series of steps before appearing as mature spliced mRNA. These steps include the removal of intron sequences by a process called splicing. cDNA derived from mRNA lacks, therefore, any intron sequences.

Nucleic acid construct: The term “nucleic acid construct” as used herein refers to a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or which is modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature or which is synthetic. The term nucleic acid construct is synonymous with the term “expression cassette” when the nucleic acid construct contains the control sequences required for expression of a coding sequence of the present invention.

Control sequences: The term “control sequences” is defined herein to include all components necessary for the expression of a polynucleotide encoding a polypeptide of the present invention. Each control sequence may be native or foreign to the nucleotide sequence encoding the polypeptide or native or foreign to each other. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the nucleotide sequence encoding a polypeptide.

Operably linked: The term “operably linked” denotes herein a configuration in which a control sequence is placed at an appropriate position relative to the coding sequence of the polynucleotide sequence such that the control sequence directs the expression of the coding sequence of a polypeptide.

Expression: The term “expression” includes any step involved in the production of the polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

Expression vector: The term “expression vector” is defined herein as a linear or circular DNA molecule that comprises a polynucleotide encoding a polypeptide of the present invention and is operably linked to additional nucleotides that provide for its expression.

Host cell: The term “host cell”, as used herein, includes any cell type that is susceptible to transformation, transfection, transduction, and the like with a nucleic acid construct or expression vector comprising a polynucleotide of the present invention.

Foreign: The term “foreign” means herein that the upstream activating sequence (UAS) according to the invention is derived from a different origin, where “origin” may refer to the gene or cell. Thus for example the UAS is normally found in the promoter region of the *Pichia pastoris* DAS promoter, however, it may according to the invention be used in a different promoter that naturally does not contain the UAS. Thus the UAS may derive from different genes from the same cell or it may derive from functionally equivalent genes from genetically different cells/species.

DETAILED DESCRIPTION OF THE INVENTION

[0038] The present invention relates to the controlled expression of polypeptides from methanol inducible promoters. Examples of these promoters have been described in several yeast cells belonging to the group of yeast known as

methylophilic yeast. In the context of the present invention a methylophilic yeast is defined as a group of yeast which can utilize methanol as a sole carbon source for their growth. The promoters for the enzymes involved in methanol metabolism in these organisms are particularly strong, and these promoters (methanol metabolism promoters) are generally used to control the heterologous expression of proteins in yeast.

[0039] Known members of methylophilic yeast host cells belong to the genera selected from the group consisting of *Pichia*, *Hansenula*, *Candida*, *Torulopsis*. According to the invention the *Pichia* host cell can in one embodiment be selected from the group consisting of *P. pastoris*, *P. methanolica*, *P. angusta*, *P. thermomethanolica*. The *Hansenula* or *Candida* host cells can be selected from the group consisting of *H. polymorpha*, and *C. boidinii*.

[0040] Several promoters have previously been isolated and described in the literature from which the expression of heterologous polypeptides can be controlled by the addition of methanol to the growth medium. Such promoters include but are not limited to e.g. the AOX1 promoter (Alcohol Oxidase promoter), DHAS promoter (or DAS promoter) (dihydroxyacetone synthase promoter), FDH promoter (or FMDH promoter) (formate dehydrogenase promoter), MOX promoter (Methanol Oxidase promoter), AOX2 promoter, ZZA1, PEX5-, PEX8-, PEX14-promoter. Particularly the promoter relevant for the present invention is the dihydroxyacetone synthase (DAS or DHAS) promoter.

[0041] Normally all of the above promoters require the presence of methanol for their induction. This induction by methanol requires the presence of additional factors (such as transcription factors), however, the exact mechanism of action for such factors have not been elucidated. In yeast e.g. Mxr1p, has been described as a key positive regulator required for methanol utilization in *Pichia pastoris* (Lin-Cereghino et al., 2006, Mol Cell Biol 26(3): 883-897).

[0042] The inventors of the present invention have previously discovered that the controlled expression of a single positive factor, encoded by the Pm1 gene from *Pichia pastoris*, as described elsewhere (co-pending application WO 2008/090211; priority date 26 Jan. 2007), can be sufficient in order to induce transcription from several methanol inducible promoters without the need for methanol in the growth medium. This was demonstrated using the Pm1 protein as a model protein for the positive activator and using the AOX1 or the DAS promoters for the controlled expression of a reporter polypeptide. The results obtained have shown that it is possible to induce the AOX1 or the DAS promoters simply by controlling the expression of the pm1 gene and without the presence of methanol in the growth medium.

[0043] The mechanism of action of the positive regulator, Pm1, has not been elucidated but one possibility is that the regulator binds to the promoter region of the methanol inducible promoter. The inventors of the present invention have identified one such region within the DNA sequence comprising the DAS promoter from *Pichia pastoris* that could possibly contain a binding region for Pm1.

[0044] As described elsewhere (WO 2008/090211/PCT/EP2008/050870) a fragment comprising the DAS promoter from *Pichia pastoris* can be obtained on a 1055 bp fragment (SEQ ID NO: 1).

[0045] In order to test promoter activity for analysis of modified promoter variants the promoter needs to be operably linked to a reporter gene. Any gene the expression of which

can be easily determined may be used. A suitable reporter gene may be the *Citrobacter braakii* phytase gene which has been codon optimized for expression in *Pichia* fused in frame to the alpha factor signal peptide from *S. cerevisiae*. The nucleic acid sequence encoding the signal peptide may also advantageously be codon optimized for *Pichia* expression. The complete DNA sequence for such a reporter gene construct is shown in SEQ ID NO: 4. (For details on how to construct this reporter gene construct see the co-pending application WO 2008/090211). The alpha factor signal peptide is encoded for in positions 1 to 255 of SEQ ID NO: 4. The reporter gene can then be fused to the fragment (SEQ ID NO: 1) comprising the DAS promoter. One such construct is shown in SEQ ID NO: 5 comprising the promoter and the start of the phytase gene. The start codon of the signal peptide can be found in position 1065 in SEQ ID NO: 5.

[0046] Using the above construct inserted into an appropriate expression vector (any vector that support expression in *Pichia*) deletion analyses were performed. From these analyses, as explained in details in the examples, it can be concluded that the 1055 bp DAS promoter fragment (SEQ ID NO: 1) contains a region which appears to be an Upstream Activating Sequence (UAS). This UAS sequence seems to be contained in a 100 bp fragment (position 701 to 800 in SEQ ID NO: 1). When this fragment is added in 1, 2, or 3 copies to a subsequence of the 1055 bp fragment a significant increase in promoter activity could be observed. The *Pichia pastoris* DAS promoter is therefore comprised in the 1055 bp fragment and the UAS is comprised in the 100 bp fragment.

[0047] In one embodiment the present invention therefore relates to an isolated polynucleotide comprising:

[0048] i) a nucleotide sequence consisting of the DAS promoter sequence from *Pichia* or a functional subsequence thereof, wherein the said DAS promoter is comprised in SEQ ID NO: 1; and

[0049] ii) at least one additional UAS, wherein the said UAS is comprised in SEQ ID NO: 2.

[0050] From the performed analysis it can be seen that the DAS promoter in another embodiment is comprised in a 855 bp subsequence corresponding to position 201 to 1055 in SEQ ID NO: 1, particularly in a 755 bp subsequence corresponding to position 301 to 1055 in SEQ ID NO: 1, more particularly in a 655 bp subsequence corresponding to position 401 to 1055 in SEQ ID NO: 1, more particularly in a 555 bp subsequence corresponding to position 501 to 1055 in SEQ ID NO: 1, even more particularly in a 455 bp subsequence corresponding to position 601 to 1055 in SEQ ID NO: 1, most particularly in a 355 bp subsequence corresponding to position 701 to 1055 in SEQ ID NO: 1.

[0051] By analysing the promoter sequence further 3 possible TATA boxes at positions 882, 955, and 1002 in SEQ ID NO: 1 respectively are revealed. In one embodiment the promoter comprises at least the TATA box at position 882. In another embodiment the promoter comprises at least the TATA box at position 955. In still another embodiment the promoter comprises at least the TATA box at position 1002.

[0052] The UAS is comprised in the 100 bp subsequence corresponding to position 701 to 800 in SEQ ID NO: 1. However the UAS may be smaller than 100 bp. Within the 100 bp subsequence approximately 20 bp from position 767 to 788 in SEQ ID NO: 1 appears to be essential for proper function.

[0053] In a further embodiment the UAS therefore comprises at least the subsequence from position 767 to 788 in SEQ ID NO: 1.

[0054] Adding even further UASs will increase the promoter activity even more. The highest activity was seen with three additional UASs. In one embodiment therefore the promoter according to the invention comprises at least two additional UASs, particularly at least three additional UASs, more particularly at least four additional UASs, and even more particularly at least five additional UASs.

[0055] It could be envisioned that the number cannot be increased indefinitely however, in case the number of additional UASs becomes too high expression of the positive activator Pm1 can be increased as well, or alternatively that the Mxr1 positive activator level may be increased.

[0056] In one embodiment the modified DAS promoter according to the invention comprises one additional UAS positioned upstream of a 855 bp subsequence of SEQ ID NO: 1. Therefore one embodiment relates to an isolated polynucleotide according to the invention, wherein the promoter is chosen from the group consisting of:

[0057] a) a polynucleotide comprising or consisting of position 77 to 828, particularly 77 to 901, more particularly 77 to 948 of SEQ ID NO: 7;

[0058] b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with position 77 to 828, particularly 77 to 901, more particularly 77 to 948 of SEQ ID NO: 7;

[0059] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with position 77 to 828, particularly 77 to 901, more particularly 77 to 948 of SEQ ID NO: 7 or a full-length complementary strand thereof.

[0060] In another embodiment the modified DAS promoter according to the invention comprises two additional UASs positioned upstream of a 855 bp subsequence of SEQ ID NO: 1. Therefore one embodiment relates to an isolated polynucleotide according to the invention, wherein the promoter is chosen from the group consisting of:

[0061] a) a polynucleotide comprising or consisting of position 60 to 920, particularly 60 to 993, more particularly 60 to 1040 of SEQ ID NO: 8;

[0062] b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with position 60 to 920, particularly 60 to 993, more particularly 60 to 1040 of SEQ ID NO: 8;

[0063] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with position 60 to 920, particularly 60 to 993, more particularly 60 to 1040 of SEQ ID NO: 8 or a full-length complementary strand thereof.

[0064] In another embodiment the modified DAS promoter according to the invention comprises three additional UASs positioned upstream of a 855 bp subsequence of SEQ ID NO: 1. Therefore one embodiment relates to an isolated polynucleotide according to the invention, wherein the promoter is chosen from the group consisting of:

[0065] a) a polynucleotide comprising or consisting of position 48 to 1015, particularly 48 to 1088, more particularly 48 to 1135 of SEQ ID NO: 9;

[0066] b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at

US 2011/0129874 A1

Jun. 2, 2011

5

least 95%, more preferably at least 97%, even more preferably at least 99% identity with position 48 to 1015, particularly 48 to 1088, more particularly 48 to 1135 of SEQ ID NO: 9;

[0067] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with position 48 to 1015, particularly 48 to 1088, more particularly 48 to 1135 of SEQ ID NO: 9 or a full-length complementary strand thereof.

[0068] The modified DAS promoters of the invention will be useful for expression of any polypeptide of interest in *Pichia*, and e.g. in *Hansenula polymorpha* and *Candida boidinii* or other methylotrophic yeast. In another aspect the invention thus relates to a DNA construct comprising a polynucleotide sequence (modified DAS promoter) of the invention operably linked to a structural gene encoding a polypeptide of interest and a terminator.

[0069] The modified DAS promoters of the invention may also advantageously be used in any suitable *Pichia* expression plasmid. The skilled person will know how to clone the promoter into such a construct. In a further embodiment the invention therefore relates to an expression vector comprising a DNA construct of the invention, further comprising a signal peptide coding region.

[0070] In a still further aspect the invention relates to a *Pichia* host cell comprising an expression vector of the invention. Particularly the *Pichia* host cell is a *Pichia pastoris* host cell.

[0071] In an even further aspect the present invention relates to a method of producing a polypeptide of interest comprising:

[0072] (a) cultivating the host cell of the invention, under conditions conducive for the production of the polypeptide of interest; and

[0073] (b) recovering the polypeptide.

[0074] In the above production method the positive regulator Prm 1 will be produced by the host cell since the prm 1 gene is endogenous to *Pichia pastoris*. However, overproducing Prm 1 can further increase promoter activity, especially when the UAS is present in multiple copies. Even over-expressing the Mxr1 protein may have an effect on the modified DAS promoter activity.

[0075] In one embodiment the invention therefore relates to a method according to the invention for producing a polypeptide of interest, wherein expression of the positive regulator Prm1 is increased in the host cell by controlling the expression of Prm1 or by increasing the copy number of the gene encoding Prm1. In an even further embodiment expression of the positive regulator Mxr1 is increased by controlling the expression of Mxr1 or by increasing the copy number of the gene encoding Mxr1. In one additional embodiment both regulators Prm1 and Mxr1 are expressed at increased levels.

[0076] In one embodiment the positive regulator is expressed constitutively from a suitable promoter. Preferably the promoter is not the native promoter meaning that the promoter controlling the expression of the positive regulator is different from the promoter normally controlling the expression. In the context of the present invention such preferred promoters are termed "non-native". The promoter could still be native to the host organism but it will be foreign in the context of the gene in question, e.g. the prm1 gene. In one particular embodiment the promoter is selected from the group consisting of the GAP promoter (glyceraldehyde-3-phosphate dehydrogenase promoter), the TEF1 promoter

(Translational elongation factor EF-1 alpha promoter), and the PGK promoter (phosphoglycerate kinase promoter). The host cell according to the invention would normally express the positive regulator from an endogenous gene present on the chromosome in addition to the expression controlled by the non-native promoter as described above. In a further embodiment the endogenous copy of the gene encoding the positive regulator could be inactivated, e.g. by deletion, or the normal promoter controlling the endogenous copy of the gene could be replaced by the chosen non-native promoter.

[0077] In another embodiment the expression of the positive regulator is controlled from an inducible promoter which is not methanol inducible.

[0078] As described above the positive regulator according to the invention may also be a functional homologue of Prm1 isolated from other yeast cells. According to one embodiment of the invention one such candidate could be Mut3 encoded by the mut3 gene from *Hansenula polymorpha* (syn. *Pichia angusta*). In the examples provided herein Prm1 or Mxr1 have been overproduced in *Pichia pastoris*. It is however possible that the same effect can be obtained by overproducing Mut3 in *Pichia* or Prm1 in *Hansenula* or Mut3 in *Hansenula*. This has not been tested.

[0079] Therefore in a further embodiment of the invention the positive regulator is Mut3.

[0080] An increase in the level of positive regulator present in the host cell can also be provided by simply having multiple copies of the gene encoding the regulator present in the host cell.

[0081] A further aspect of the invention relates to the UAS comprised in SEQ ID NO: 2 or at least comprising SEQ ID NO: 3.

[0082] The present invention therefore relates to an isolated polynucleotide selected from the group consisting of:

[0083] (a) a polynucleotide comprising or consisting of SEQ ID NO: 2; or

[0084] (b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 2; or

[0085] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 2 or a full-length complementary strand thereof.

[0086] In a further embodiment the present invention relates to an isolated polynucleotide selected from the group consisting of:

[0087] (a) a polynucleotide comprising or consisting of SEQ ID NO: 3; or

[0088] (b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 3; or

[0089] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 3 or a full-length complementary strand thereof.

[0090] The UAS according to the invention can be used for activating foreign promoters as well as the DAS promoter. In this case one or more copies of the UAS according to the invention is combined with a promoter that in its natural form does not contain the UAS as explained herein. In a further aspect the present invention therefore relates to a promoter

US 2011/0129874 A1

Jun. 2, 2011

6

comprising an UAS according to the invention, wherein the UAS is foreign to the promoter or wherein the UAS is present in more than one copy.

[0091] In another embodiment the present invention relates to a use of the UAS for increasing transcription from a promoter. The promoter may be a foreign promoter in which case it does not already contain a copy of the UAS or it may be the DAS promoter in which case additional copies can be added as described above.

[0092] Thus in a further embodiment the invention relates to a promoter comprising an UAS selected from the group consisting of:

[0093] i) (a) a polynucleotide comprising or consisting of SEQ ID NO: 2; or

[0094] (b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 2; or

[0095] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 2 or a full-length complementary strand thereof; or

[0096] ii) (a) a polynucleotide comprising or consisting of SEQ ID NO: 3; or

[0097] (b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 3; or

[0098] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 3 or a full-length complementary strand thereof; and

[0099] wherein the UAS according to (i) or (ii) is either foreign to the promoter or present in more than one copy.

[0100] In still another aspect the invention relates to a use of an UAS for increasing transcription from a promoter, wherein the UAS is selected from the group consisting of:

[0101] i) (a) a polynucleotide comprising or consisting of SEQ ID NO: 2; or

[0102] (b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 2; or

[0103] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 2 or a full-length complementary strand thereof; or

[0104] ii) (a) a polynucleotide comprising or consisting of SEQ ID NO: 3; or

[0105] (b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 3; or

[0106] c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 3 or a full-length complementary strand thereof; and

[0107] wherein the UAS according to (i) or (ii) is either foreign to the promoter or present in more than one copy.

[0108] Promoters suitable for the above activation may in particular be promoter induced by methanol.

[0109] In the production methods of the present invention, the cells are cultivated in a nutrient medium suitable for production of the polypeptide using methods well known in

the art. For example, the cell may be cultivated by shake flask cultivation, and small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the polypeptide to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (e.g., in catalogues of the American Type Culture Collection). If the polypeptide is secreted into the nutrient medium, the polypeptide can be recovered directly from the medium. If the polypeptide is not secreted, it can be recovered from cell lysates.

[0110] The polypeptides may be detected using methods known in the art that are specific for the polypeptides. These detection methods may include use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, an enzyme assay may be used to determine the activity of the polypeptide as described herein.

[0111] The resulting polypeptide may be recovered using methods known in the art. For example, the polypeptide may be recovered from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation.

[0112] The polypeptides produced by the present invention may be purified by a variety of procedures known in the art including, but not limited to, chromatography (e.g., ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (e.g., preparative isoelectric focusing), differential solubility (e.g., ammonium sulfate precipitation), SDS-PAGE, or extraction (see, e.g., Protein Purification, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

[0113] In one particular embodiment the polypeptide produced from the host cell is heterologous to the host cell. In another embodiment the polypeptide is homologous to the host cell.

EXAMPLES

Materials and Methods

Strains

[0114] *E. coli* TOP10 (invitrogen) and DH5alpha (TOYOBO) were used for a plasmid construction. LB (1% Trypton (Difco), 0.5% Yeast extract (Difco), 0.5% NaCl) was used as base medium after supplement of relevant antibiotics. *Pichia pastoris* his4 mutant, GS115, was used for the expression test.

Pichia pastoris COIs702 (Muts) is a AOX1 gene disrupted strain of *Pichia pastoris* NRRL Y-15851 and is described in example 5.

The used media for its growth are as following:

YPD (2% Pepton (Difco), 1% Yeast extract (Difco), 2% Glucose)

RD Agar; 1M sorbitol, 2% Glucose, 1.34% Yeast nitrogen base (Difco), $4 \times 10^{-5}\%$ biotin, 0.005% amino acids (L-glutamic acid, L-methionine, L-lysine, L-leucine, and L-isoleucin), 2% Agar Noble (Difco)

US 2011/0129874 A1

Jun. 2, 2011

7

MD Agar; 1.34% Yeast nitrogen base, $4 \times 10^{-5}\%$ Biotin, 2% Glucose

Transformation of *Pichia Pastoris*:

[0115] *Pichia pastoris* strains were transformed by electroporation. Fresh competent cells were prepared by the following procedure. The host strain, was inoculated to 100 ml of YPD and grown till OD₆₆₀ is 1.2–1.4. The cells were washed with ice-cold water twice (100 ml, and 50 ml), and with 4 ml of ice-cold 1M sorbitol. Then the cells were suspended 0.2 ml of ice-cold 1M sorbitol. Linearized plasmid DNA (1–2 µg) was mixed with 80 µl of fresh competent cells and stored on ice for 5 min. Cells were transferred to an ice-cold 0.2 cm electroporation cuvette. Transformation was performed using a BioRad™ GenePulser II. Parameters used were 1500 V, 25 µF and 200Ω. Immediately after pulsing, cells are suspended in 1 ml of ice cold 1 M sorbitol. The mixtures were plated on the relevant selection plates.

Colony PCR for Screening of His4 Locus Integration

[0116] The cell was picked with sterilized tooth picks to a 0.2 ml tube then baked in a microwave oven. The dried cell was suspended in 50 µl of sterilized water and it was subjected to PCR using Expand High Fidelity plus (Roche). The reaction mixture was 20 µl including 2 mM dNTP, 10 microM of each primer, 1 unit of Expand high fidelity plus (Roche), 1× Expand high fidelity plus buffer (Roche), and 1 µl of cell suspension mentioned above. The PCR primers were primer139; 5'-ctgctctagccagtttgcg-3' (SEQ ID NO: 10) (upstream of His4 maker in genome) and S98; 5'-gccgccagtcctgctcgt-3' (SEQ ID NO: 11) (between His4 marker and AOX terminator in expression plasmids). The PCR program is as below.

temp	time	cycle
94° C.	2 min	1
94° C.	10 sec	11
55° C.	30 sec	
68° C.	3 min	
94° C.	10 sec	20
55° C.	30 sec	
68° C.	3 min + 20 sec/cycle	
68° C.	7 min	1

[0117] The strains in which the expression cassette was integrated at His4 locus showed about 2.9 kb band.

Shake Flask Evaluation with Methanol Induction

[0118] Cell cultivated on the MD agar for 3 days was inoculated in 100 ml of YPD in 500 ml of SF and cultivated at 30° C. with shaking. Then 1 ml of the seed culture was inoculated in 100 ml of YPD and cultivated at 30° C. for 2 days. During cultivation, 5 ml of 40% methanol was added for induction on day 2. Sampling was carried out on day 3.

Phytase Assay Method:

[0119] 7.5 mM of sodium phytate dissolved in the acetate buffer pH 5.5 is mixed with ½ volume of enzyme sample solution in the same acetate buffer containing 0.01% Tween 20. After the incubation at 37° C. for 30 minutes, the stop reagent containing 20 mM ammonium heptamolybdate and 0.06% ammonium vanadate dissolved in 10.8% nitric acid was added to generate yellow complex with released inorganic phosphate. The amount of released phosphate is measured photometrically as the absorbance at 405 nm. One phytase unit is defined as the amount of enzyme to release 1 µmol inorganic phosphate per minute.

TABLE 1

Primers and primer combinations				
Primer	sequence	direction	plasmid	
primer126	5' gaaaagtgccacctgagtcggttactcaaaaagtc 3'	forward	pDd-2	
primer127	5' gaaaagtgccacctgcggtatctgaaacacaagac 3'	forward	pDd-3	
primer128	5' gaaaagtgccacctgtgaaaattactggtaacttc 3'	forward	pDd-4	
primer129	5' gaaaagtgccacctgtctataagaacgaaggattt 3'	forward	pDd-5	
primer130	5' gaaaagtgccacctgaaagaaattcgacaaaatag 3'	forward	pDd-6	
primer132	5' gaaaagtgccacctgttctccacattcagtcatag 3'	forward	pDd-8	
primer133	5' gaaaagtgccacctgcagtttaagtttttctttc 3'	forward	pDd-9	
primer134	5' gaaaagtgccacctgttataacccctctaaacact 3'	forward	pDd-10	
primer135	5' aaccagcccgatgggacgtccattgtggccact-gatctgggg 3'	reverse	pDd-2~20	
primer144	5' ccccgaaaagtgccacctgacgtc 3'	forward	pDd-11~20	
primer152	5' actgaatgtggagaacgatggtgttgcctcaa- caaagggatg 3'	reverse	pDd-14, pDd-20	
primer153	5' ttgagcaacaccatcggttctccacattcagtcata- gatggg 3'	forward	pDd-14, 20	

US 2011/0129874 A1

Jun. 2, 2011

8

TABLE 1-continued

Primers and primer combinations			
Primer	sequence	direction	plasmid
primer154	5' aacagataatagaaattctccacattcagtcata-gatggg 3'	forward	pDd-15
primer155	5' actgaatgtggagaatttctattatctgttcgtttggcg 3'	reverse	pDd-15
primer156	5' ttagtatacttgcccttctccacattcagtcatagatggg 3'	forward	pDd-16
primer157	5' actgaatgtggagaagggaagtatac-taaatctcggaac 3'	reverse	pDd-16
primer158	5' aacagataatagaaattagccagtcagaaagag-gaaactt 3'	forward	oDd-17
primer159	5' ttctgactggctaatttctattatctgttcgtttggcg 3'	reverse	pDd-17
primer160	5' ttagtatacttgcccttagccagtcagaaagag-gaaactt 3'	forward	pDd-18
primer161	5' ttctgactggctaagggaagtatac-taaatctcggaac 3'	reverse	pDd-18
primer162	5' ttagtatacttgcccaaagaaattcggacaata-gaacac 3'	forward	pDd-19, 20
primer163	5' tgtccgaatttctttgggcaagtatac-taaatctcggaaca 3'	reverse	pDd-19, 20
primer174	5' taggctaaact-cattcgctcggttaagtttctctttcg 3'	reverse	pDd32
primer175	5' gcgcgaatgagtttagcctagatatacctt 3'	forward	pDd32
primer176	5' aagtatcggaacaactaaactcattcgcgctta-gatttc 3'	reverse	pDd33
primer177	5' gtttagtgttccgatacttctccacattc 3'	forward	pDd33
primer196	5' aagatgcttttctgtgactggtgagtactcaaccaagt-ca 3'	forward	pDd32, 33
primer201	5' ttccccgaaaagtgccacctgacgtcggtggcgcc aaggactagt 3'	forward	pDd26, 27, 28
primer202	5' agacaagttctccaccagtgaggtaac-cagcccagtggaacgtc 3'	reverse	pDd26, 27, 28

TABLE 2

Primer combination to make the promoter variants of pDd2~20, pDd32, and pDd33				
Plasmid	1st PCR (forward)	1st PCR (reverse)	2nd PCR (forward)	2nd PCR (reverse)
pDd-2	primer126	primer135		
pDd-3	primer127	primer135		
pDd-4	primer128	primer135		
pDd-5	primer129	primer135		
pDd-6	primer130	primer135		
pDd-8	primer132	primer135		
pDd-9	primer133	primer135		
pDd-10	primer134	primer135		
pDd-14	primer144	primer152	primer144	primer135
	primer153	primer135		
pDd-15	primer144	primer155	primer144	primer135
	primer154	primer135		
pDd-16	primer144	primer157	primer144	primer135
	primer156	primer135		

TABLE 2-continued

Primer combination to make the promoter variants of pDd2~20, pDd32, and pDd33				
Plasmid	1st PCR (forward)	1st PCR (reverse)	2nd PCR (forward)	2nd PCR (reverse)
pDd-17	primer144	primer159	primer144	primer135
	primer158	primer135		
pDd-18	primer144	primer161	primer144	primer135
	primer160	primer135		
pDd-19	primer144	primer163	primer144	primer135
	primer162	primer135		
pDd-20	primer144	primer163	primer144	primer135
	primer162	primer152		
	primer153	primer135		
pDd-32	primer196	primer174	primer196	primer135
	primer175	primer135		
pDd-33	primer196	primer176	primer196	primer135
	primer177	primer135		

US 2011/0129874 A1

Jun. 2, 2011

9

Example 1

Construction and Expression Evaluation DAS1 Promoter Variants

[0120] Cloning of DAS promoter of *Pichia pastoris* has been described previously (WO 2008/090211). An expression cassette consisting of the wt DAS promoter from *Pichia pastoris* controlling the expression of a phytase gene, optimized for expression in *Pichia*, was used for the construction of promoter variants (deletion variants). The complete sequence of the promoter fragment is shown in SEQ ID NO: 1, the reporter gene in SEQ ID NO: 4 and the fusion construct (expression cassette) is shown in SEQ ID NO: 5 (DAS wt promoter and 5'-end of the phytase coding sequence including a codon optimized alpha factor signal peptide encoding sequence). Promoter variants of different length (as shown in FIG. 1 in which position -1 corresponds to position 1055 in SEQ ID NO: 1 and position -1055 corresponds to position 1 in SEQ ID NO: 1) were prepared by PCR using the primers shown in Table 1 and the primer combination sets shown in Table 2. The template DNA was the wild type DAS promoter. The expression cassette can be inserted in any appropriate *Pichia* expression plasmid having an appropriate selection marker. Preferably the selection marker is HIS4. One such construct is pDAS1 wt.

[0121] PCR was carried out using the 50 microL of reaction including 2 mM dNTP, 10 microM of each primer, 2.8 unit of Expand high fidelity plus (Roche), 1× Expand high fidelity plus buffer (Roche), and 2 ng of the template plasmid DNA. The PCR program is as below.

Temp	time	cycle
94° C.	2 min	1
94° C.	10 sec	11
55° C.	30 sec	
68° C.	3 min	
94° C.	10 sec	20
55° C.	30 sec	
68° C.	3 min + 20 sec/cycle	
68° C.	7 min	1

[0122] The amplified DNA fragment was purified by gel extraction kit (Qiagen) and used for the construction of phytase expression plasmids.

[0123] For making expression plasmids, the amplified fragment was sub-cloned by In-Fusion PCR cloning kit (Clontech) into the template phytase expression vector (pDAS1 wt) digested with the AatII restriction enzyme.

[0124] Resulting expression plasmids (pDd-2 through pDd-10) were linearized by digestion with SalI and transformed into *Pichia pastoris* his4 strain. The strains in which the expression cassette was integrated at the HIS4 locus were screened by colony PCR. The selected strains were cultivated in liquid medium with methanol induction and phytase activity in the culture broth was measured. The results are shown in Table 3.

TABLE 3

Strains (series)	plasmid	Promoter length (bp)	Relative activity (%)
DASwt	pDAS1 wt	1050	100
2	pDd-2	854	155
3	pDd-3	754	102
4	pDd-4	654	128

TABLE 3-continued

Strains (series)	plasmid	Promoter length (bp)	Relative activity (%)
5	pDd-5	554	102
6	pDd-6	454	78
8	pDd-8	254	24
9	pDd-9	154	13
10	pDd-10	54	14

Example 2

Construction and Expression Evaluation DAS1 Promoter Variants (2)

[0125] DAS promoter variants which possess internal deletions as shown in FIG. 2 were constructed by SOE PCR. The 1st PCR was carried out using the 50 microL of reaction including 2 mM dNTP, 10 microM of each primer shown in Table 2, 2.8 unit of Expand high fidelity plus (Roche), 1× Expand high fidelity plus buffer (Roche), and 2 ng of the plasmid DNA of pDAS1 wt. The PCR program is as below.

temp	time	cycle
94° C.	2 min	1
94° C.	10 sec	11
55° C.	30 sec	
68° C.	3 min	
94° C.	10 sec	20
55° C.	30 sec	
68° C.	3 min + 20 sec/cycle	
68° C.	7 min	1

[0126] The purified fragments from 1st PCR were subjected to the 2nd PCR using the 50 microL of reaction including 2 mM dNTP, 10 microM of primer 144 and primer 135, 2.8 unit of Expand high fidelity plus (Roche), 1× Expand high fidelity plus buffer (Roche), and purified PCR product of the 1st run. The PCR program is as below.

Temp	time	cycle
94° C.	2 min	1
94° C.	10 sec	11
60° C.	30 sec	
68° C.	3 min	
94° C.	10 sec	20
60° C.	30 sec	
68° C.	3 min + 20 sec/cycle	
68° C.	7 min	1

[0127] Structure of each promoter variants and the used primers are shown in FIG. 2. The sequence corresponding to the pDAS1 wt fragment shown in FIG. 1 is shown in SEQ ID NO: 1 (promoter) and SEQ ID NO: 5 (promoter+alpha leader+N-terminal phytase). The phytase expression plasmids carrying the DAS1 promoter variants (pDd-14 through pDd20) were generated, and transformants of *Pichia pastoris* were isolated as previously shown. The selected strains were cultivated in YPD medium with methanol induction. The phytase activity in the culture broth was measured and results were shown in Table 4.

US 2011/0129874 A1

10

Jun. 2, 2011

TABLE 4

Strains (series)	plasmid	Deleted region	Relative activity (%)
DASwt	pDAS1 wt	no	100
14	pDd-14	-355~-255	5
16	pDd-16	-555~-255	2
17	pDd-17	-455~-355	68
18	pDd-18	-555~-355	75
19	pDd-19	-555~-455	77
20	pDd-20	-555~-455 -355~-255	2

[0128] It was found that the deletion between -355 and -255 (containing in pDd-14, 16, 20) drastically reduced the expression level. Therefore, it was concluded that the region is an Upstream Activation Region (UAS) of DAS1 wt promoter.

Example 3

Amplification of UAS in DAS Promoter

[0129] DAS promoter variants in which the number of the UAS is amplified were constructed. The resulting constructs are illustrated in FIG. 3. The UAS was amplified by PCR and the constructs were based on the wild type DAS promoter shown as pDAS wt2 in the figure and the complete sequence from the NsiI site and including the phytase gene encoding the N-terminal is shown in SEQ ID NO: 6. The amplification resulted in constructs having multiple copies of the UAS as shown in FIG. 3 and by using the primers 201 and 202 these constructs were sub-cloned into pDAS1 wt.

[0130] The promoter region and a part of phytase gene were amplified by PCR using primer 201 and primer 202 with the 50 microL of reaction mixture including 2 mM dNTP, 10 microM of each primer, 2.8 unit of Expand high fidelity plus (Roche), 1× Expand high fidelity plus buffer (Roche), and 2 ng of the plasmid of template DNA. The PCR program is as below.

temp	time	cycle
94° C.	2 min	1
94° C.	10 sec	11
55° C.	30 sec	
68° C.	3 min	
94° C.	10 sec	20
55° C.	30 sec	
68° C.	3 min + 20 sec/cycle	
68° C.	7 min	1

[0131] The amplified DNA fragment was purified and sub-cloned into pDAS1 wt digested with PmlI and AatII using the In-fusion cloning kit. The amplification resulted in pDd-26, pDd-27, and pDd-28 and the complete sequences corresponding to the region marked by the primers 201 and 202 are shown in SEQ ID NO: 7, 8, and 9 respectively. These promoter cassettes can then be used in an appropriate expression plasmid of choice.

[0132] Each expression plasmid was integrated at HIS4 locus of *Pichia pastoris*, and the expression level with methanol induction was evaluated in YPD medium. The results are shown in Table 5.

TABLE 5

Strains (series)	plasmid	Number of additional UAS	Relative activity (%)
2	pDd-2	0	100
26	pDd-26	1	108
27	pDd-27	2	139
28	pDd-28	3	181

[0133] It can be seen from the results that increasing the number of the UAS will result in an increase in promoter activity.

Example 4

Construction and Expression Evaluation DAS1 Promoter Variants (3)

[0134] In order to narrow down the essential region in the UAS, further deletion variants were created (FIG. 4). Two region, -306 to -324 (PBS1) and -289 to -268 (PBS2), corresponding to position 732-750 and 767-788 in SEQ ID NO: 1 respectively, were predicted as potential binding sites of transcription factor, and variants having these regions deleted were constructed by SOE-PCR as described in Example 2. The amplified DNA fragment was purified by gel extraction kit (Qiagen) and used for the construction of expression plasmid of phytase.

[0135] For making expression plasmids, the amplified fragment was sub-cloned by In-Fusion PCR cloning kit (Clontech) into the template phytase expression vector pDAS1 wt digested by the SacI and SnaBI restriction sites. The primers are shown in Table 1 and 2.

[0136] Each expression plasmid was integrated at HIS4 locus of *Pichia pastoris*, and the expression level with methanol induction was evaluated in YPD medium. The results are shown in Table 6.

TABLE 6

The result of SF evaluation (pDd-32 and pDd-33)			
Strains (series)	plasmid	deletion	Relative activity (%)
DP3	pNo-DP3	no	100
32	pDd-32	PBS1	75
33	pDd-33	PBS2	24

Example 5

Expression of *Humicola insolens* Cutinase from Modified DAS Promoter with Prm1 and/or Mxr1 Over-Expression

[0137] The aim of this experiment was to check whether the expression of a protein of interest, exemplified by the *Humicola insolens* cutinase gene controlled by an improved DAS promoter according to the invention, having four repeats of the UAS, could be improved by co-expression of transcription factors, such as Prm1 and/or Mxr1 in *Pichia pastoris*. The host strain used was COLs702 (Mut^o).

[0138] COLs702 was constructed from *Pichia pastoris* NRRL Y-15851, which has a mutation in the his4 gene to make the gene inactive. NRRL Y-15851 was transformed with pCOLs693 (SEQ ID NO: 41) in standard manner. An

US 2011/0129874 A1

Jun. 2, 2011

11

aox1 deleted strain, COLs702, was obtained using a traditional approach, by transformation with the marker gene, SUC2, flanked by locus specific deletion fragments. The plasmid pCOIs693 has a SUC2 gene from *S. cerevisiae* as the marker gene and flanking sequences from the aox1 gene. Transformants were isolated using sucrose as the sole carbon source. Due to the his4 negative genotype of the mother strain, histidine was supplemented to the selection agar medium. Fast growing transformants on the selection plate were isolated.

[0139] Isolated transformants were studied by PCR to confirm the aimed insertion of SUC2 gene into AOX1 locus. Resulting strains have the AOX1 gene disrupted by this event. One of the strains was named as *Pichia pastoris* COLs702 for further use.

[0140] A re-transformation of low-expressing *Humicola insolens* cutinase transformants comprising pNori58-HIC (wt *Humicola insolens* cutinase controlled by the improved DAS promoter), pNori58-RSII0014 (*H. insolens* cutinase variant controlled by improved DAS promoter), or pNori58-RSII0007 (*H. insolens* cutinase variant controlled by improved DAS promoter), was carried out with the plasmids harboring Prm1 (pGPm, SEQ ID NO: 42) or Mxr1 (pGMxr, SEQ ID NO: 43) or both, and their expression was analyzed. Both plasmids have the Zeocin resistant gene as a selection marker gene and the regulator gene is controlled by the GAP promoter. A homologous recombination event at the GAP promoter region of the respective vectors pNori58 and pGPm/pGMxr was expected. pGPm comprises the expression cassette (SEQ ID NO: 44) having the GAP promoter in position 1-483, Prm1 CDS in position 490-3549, terminator in position 3487-3827. pGMxr comprises the expression cassette (SEQ ID NO: 45) having the GAP promoter in position 1-483, Mxr1 CDS in position 493-3957, terminator in position 4028-4368.

[0141] For transformations, plasmid DNA (pGMxr and pGPm) was linearized with AvrII and cleaned-up using Bio-rad clean-up kit.

[0142] After heat shock at 42° C., the cells were harvested and re-suspended in 1 ml YPDS w/o Zeocin and incubated with shaking at 30° C., 200 rpm for 3 h. The cells were then harvested, resuspended in 100 µl of the supernatant and plated on YPDS plates containing 100 µg/ml Zeocin.

[0143] One transformant from each of the above transformations that showed expression in small scale was chosen for expression in large scale: inoculation in shake flasks with 150 ml Buffered Sorbitol Complex media (BSCM) in 500 ml baffled flasks. After ~24 h of growth, the cultures were induced with 100% methanol to a final concentration of 1%. Once induction was started, the cotton plugs were replaced with four layers of gauze cloth+1 layer of Cheese cloth. Induction was carried out every 24 h at 22° C., 150 rpm for 3 days with induction and without induction for 2.5 days (over the weekend). An aliquot of culture (500 µl) was pelleted by centrifugation at 13000 rpm for 5 min and the supernatants (12 µl) analyzed for protein expression on a 12% SDS-PAGE with coomassie stain.

[0144] From the gel-data the following expression levels could be seen:

Enzyme expression plasmid	Prm1/Mxr1 plasmid	Expression level
pNori58-HIC	none	+++
	pGPm	+++
	pGMxr/pGPm	+++
pNori58-RSII0007	None	—
	pGMxr	+
	pGMxr/pGPm	—
pNori58-RSII0014	None	++
	pGMxr	+++
	pGPm	++
	pGMxr/pGPm	+++

[0145] The obtained data indicate an improved expression of the test protein when Mxr protein was over-expressed.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 45

<210> SEQ ID NO 1

<211> LENGTH: 1055

<212> TYPE: DNA

<213> ORGANISM: *Pichia pastoris*

<400> SEQUENCE: 1

atagggagaa aaaccgagac aacgatggaa ctcccatgta gattccaccg cccaattac 60

tgttttgggc aatcctgttg ataagacgca ttctagagtt gtttcatgaa agggttacgg 120

gtgttgattg gtttgagata tgccagagga cagatcaatc tgtggtttgc taaactggaa 180

gtctggtaag gactctagca agtcctgtac tcaaaaagtc ataccaagta agattacgta 240

acacctgggc atgactttct aagttagcaa gtcaccaaga gggtcctatt taacgtttgg 300

cggtatctga aacacaagac ttgcctatcc catagtacat catattacct gtcaagctat 360

gtacccccac agaaataccc caaaagtga agtgaaaaaa tgaaaattac tggttaacttc 420

accccataac aaacttaata atttctgtag ccaatgaaag taaaccccat tcaatgttcc 480

US 2011/0129874 A1

Jun. 2, 2011

12

-continued

gagatttagt atacttgccc ctataagaaa cgaaggattt cagcttcctt accccatgaa	540
cagaaatctt ccatttacct ccactggag agatccgccc aaacgaacag ataatagaaa	600
aaagaaatc ggacaaatag aacactttct cagccaatta aagtcattcc atgcactccc	660
tttagctgcc gttccatccc ttgttgagc aacaccatcg ttagccagta cgaagagga	720
aacttaaccg ataccttgga gaaatctaag gcgcgaatga gtttagccta gatatcctta	780
gtgaagggtt gttccgatac ttctccacat tcagtcatag atgggcagct ttgttatcat	840
gaagagacgg aaacgggcat taagggttaa ccgccaaatt atataaagac aacatgtccc	900
cagtttaaag tttttcttct ctattcttgt atcctgagtg accgttgtgt ttaataaac	960
aagttcgttt taacttaaga ccaaaaccag ttacaacaaa ttataacccc tctaaacact	1020
aaagttcact cttatcaaac tatcaaacat caaaa	1055

<210> SEQ ID NO 2
<211> LENGTH: 100
<212> TYPE: DNA
<213> ORGANISM: Pichia pastoris

<400> SEQUENCE: 2

ttagccagta cgaagagga aacttaaccg ataccttgga gaaatctaag gcgcgaatga	60
gtttagccta gatatcctta gtgaagggtt gttccgatac	100

<210> SEQ ID NO 3
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Pichia pastoris

<400> SEQUENCE: 3

cctagatatc cttagtgaag gg	22
--------------------------	----

<210> SEQ ID NO 4
<211> LENGTH: 1491
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: alpha factor signal peptide fused to
Citrobacter braakii phytase both optimized for expression in
Pichia

<400> SEQUENCE: 4

atgagattcc catccatctt cactgctgtt ttgttcgctg ctctctccgc ttggctgct	60
ccagttaaca ctactactga agatgaaact gctcaaatcc cagctgaagc tgttatcggt	120
tactccgact tggaagggtga ttccgacgtt gctgttttgc cattctccaa ctccaccaac	180
aacggattgt tgttcattaa caccaccatt gcttccatcg ctgctaagga agaaggtgtt	240
tccttggaag agagagagga acagaatgga atgaagttgg agagagttgt catcgtttct	300
agacacgggt ttagagctcc caccaaatc actccaatca tgaagaacgt caccacagat	360
cagtggccac aatgggacgt ccactgggc tggttgactc cactggttgg agaactgttc	420
tctgaattgg gtcagtacca gagactgtgg ttcacctcca aaggacttct gaataaccaa	480
acttgcccat cccaggaca agtcgctgtt attgccgaca ccgatcaaag aaccagaaaa	540
accggagagg cctttttggc aggacttgct ccaaaatgcc agattcaagt ccactaccaa	600
aaagacgaag agaagaacga tccattgttc aatcccgta agatgggaaa atgctccttt	660

US 2011/0129874 A1

Jun. 2, 2011

13

-continued

aacaccttgc	aagtcaaaaa	cgccattttg	gaaagagcag	gtggcaatat	cgagctttac	720
accagcggtt	accaatcttc	ttttagaact	ttggaaaatg	ttttgaactt	tagtcagtcc	780
gagacttgca	agaccaccga	gaagtctacc	aagtgcactt	tgcccagggc	tttgccctcc	840
gagcttaag	tcaactccga	taacgtctcc	ttgccaggag	catggctctc	ttcctccact	900
ttgaccgaga	ttttcttggt	gcaggaggca	caaggaatgc	cacaggtcgc	atggggtaga	960
attaccgggtg	aaaaggaatg	gagagacttg	ctgtctcttc	acaacgcccc	gttcgatctc	1020
ttgcagagaa	ccccagaggt	tgccagatcc	agagctactc	cacttttggg	tatgatcgac	1080
accgctttgc	tgaccaatgg	taccaccgag	aacagatacg	gtattaagtt	gccagtctcc	1140
ttgctgttca	ttgcagggtc	cgacaccaat	ttggccaact	tgtctggagc	cttggacctg	1200
aactgggtctt	tgccaggaca	gcccgcacaa	accccaccag	gaggcgaatt	ggttttcgaa	1260
aagtggaaaa	gaacctccga	taacaccgat	tgggtccaag	tctccttcgt	ctaccaaacc	1320
ttgagagata	tgcgtgacat	tcagccactg	tctttggaga	agcccgtggg	taaggttgac	1380
ttgaaattga	tcgcttgcca	agaaaagaac	tcccaggga	tgtgctcttt	gaagtccttt	1440
tcagatttga	tcaaggagat	tagagtcccc	gagtgcgctg	tcaccgagta	a	1491

<210> SEQ ID NO 5
<211> LENGTH: 1485
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: DAS promoter fused to phytase

<400> SEQUENCE: 5

gacgtcatag	ggagaaaaac	cgagacaaag	atggaactcc	catgtagatt	ccaccgcccc	60
aattactggt	ttgggcaatc	ctgttgataa	gacgcattct	agagttgttt	catgaaaggg	120
ttacgggtgt	tgattgggtt	gagatatgcc	agaggacaga	tcaatctgtg	gtttgtctaaa	180
ctggaagtct	ggtaaggact	ctagcaagtc	cgttactcaa	aaagtcatac	caagtaagat	240
tacgtaacac	ctgggcatga	ctttctaagt	tagcaagtca	ccaagagggg	cctatttaac	300
gtttggcgg	atctgaaaca	caagacttgc	ctatcccata	gtacatcata	ttacctgtca	360
agctatgcta	ccccacagaa	atcccccaaa	agttgaagtg	aaaaaatgaa	aattactggg	420
aacttcaccc	cataacaaac	ttaataattt	ctgtagccaa	tgaaagtaaa	ccccattcaa	480
tgttccgaga	tttagtatac	ttgcccctat	aagaaacgaa	ggatttcagc	ttccttacct	540
catgaacaga	aatcttccat	ttacccccca	ctggagagat	ccgcccacaa	gaacagataa	600
tagaaaaaag	aaattcggac	aaatagaaca	ctttctcagc	caattaaagt	cattccatgc	660
actcccttta	gctgccgttc	catccctttg	ttgagcaaca	ccatcgtagg	ccagtacgaa	720
agaggaaact	taaccgatac	cttgagagaaa	tctaaggcgc	gaatgagttt	agcctagata	780
tccttagtga	agggttggtc	cgatacttct	ccacattcag	tcatatagtg	gcagctttgt	840
tatcatgaag	agacggaaac	gggcattaag	ggtaaccgc	caattatat	aaagacaaca	900
tgtccccagt	ttaaagtttt	tctttcctat	tcttgatcc	tgagtgaccg	ttgtgtttta	960
tataacaagt	tcgttttaac	ttaagaccaa	aaccagttac	aacaaattat	aacctctcta	1020
aacactaaag	ttcactctta	tcaaaactatc	aaacatcaaa	acatatgaga	ttcccatcca	1080
tcttcactgc	tgttttggtc	gctgcttctc	ccgctttggc	tgtccaggtt	aacactacta	1140

US 2011/0129874 A1

Jun. 2, 2011

14

-continued

ctgaagatga aactgctcaa atcccagctg aagctgttat cggttactcc gacttgggaag	1200
gtgatttcga cgttgctgtt ttgccattct ccaactccac caacaacgga ttgttggtca	1260
ttaacaccac cattgtctcc atcgctgcta aggaagaagg tgtttccttg gaaaagagag	1320
aggaacagaa tggaatgaag ttggagagag ttgtcatcgt ttctagacac ggtgttagag	1380
ctcccaccaa attcactcca atcatgaaga acgtcaccce agatcagtgg ccacaatggg	1440
acgtccact gggctgggtg actccacgtg gtggagaact tgtct	1485

<210> SEQ ID NO 6
<211> LENGTH: 1493
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: DAS promoter fused to phytase

<400> SEQUENCE: 6

tcattgcata cgtcataggg agaaaaaccg agacaacgat ggaactccca ttagattccc	60
accgccccaa ttactgtttt gggcaatcct gttgataaga cgcattctag agttgtttca	120
tgaaaggggt acgggtgttg attggtttga gatatgccag aggacagatc aatctgtggt	180
ttgctaaact ggaagtcttg taaggactct agcaagtcog ttactcaaaa agtcatacca	240
agtaagatta cgtaaacact gggcatgact ttctaagtta gcaagtcacc aagagggtcc	300
tatttaacgt ttggcgggtat ctgaaacaca agacttgccct atcccatagt acatcatatt	360
acctgtcaag ctatgttacc ccacagaaat accccaaaag ttgaagtga aaaaatgaaa	420
ttactggtaa cttcacccca taacaaactt aataatttct gtagccaatg aaagtaaaacc	480
ccattcaatg ttccgagatt tagtatactt gccctataa gaaacgaagg atttcagctt	540
ccttacccca tgaacagaaa tcttccattt acccccact ggagagatcc gcccacacga	600
acagataata gaaaaaagaa attcggacaa atagaacact ttctcagcca attaaagtca	660
ttccatgcac tccctttagc tgccgttcca tccctttgtt gagcaacacc atcgtagacc	720
agtacgaaag aggaacttta accgatacct tggagaaatc taaggcgcgga atgagtttag	780
cctagatata cttagtgaag ggttggttccg atacttctcc acattcagtc atagatgggc	840
agctttgtta tcatgaagag acggaacggy gcattaagggy ttaacggcca aattatataa	900
agacaacatg tccccagttt aaagtttttc ttccctatcc ttgtatcctg agtgaccgtt	960
gtgtttaata taacaagttc gttttaactt aagacaaaa ccagttacaa caaattataa	1020
cccctctaaa cactaaagtt cactcttata aaactatcaa acatcaaaac atatgagatt	1080
cccatccatc ttcaetgctg ttttgttcgc tgettectcc gctttggctg ctccagttaa	1140
cactactact gaagatgaaa ctgctcaaat ccagctgaa gctgttatcg gttactccga	1200
cttgaaggt gatttcgacg ttgctgtttt gccattctcc aactccacca acaacggatt	1260
gttggtcatt aacaccacca ttgcttccat cgtgctaag gaagaaggty ttccttgga	1320
aaagagagag gaacagaatg gaatgaagtt ggagagagtt gtcacgttt ctgacacggy	1380
tgtagagct cccaccaa atcaactccat catgaagaac gtcacccag atcagtggcc	1440
acaatgggac gtcccactgy gctggttgac tccacgtggt ggagaacttg tct	1493

<210> SEQ ID NO 7
<211> LENGTH: 1421
<212> TYPE: DNA

US 2011/0129874 A1

Jun. 2, 2011

15

-continued

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Promoter variant fused to phytase reporter gene

<400> SEQUENCE: 7

```
ttccccgaaa agtgccacct gacgtcgttg gcgcgccaag gactagtccc ggaagatctt    60
ccagctttgt ttaaaactta ccagtagcaa agaggaaact taaccgatac cttggagaaa    120
tctaaggcgc gaatgagttt agcctagata tccttagtga aggggtgttc cgatactacg    180
taaacacctgg gcatgacttt ctaagttagc aagtcaccaa gaggggtcta tttaacgttt    240
ggcgggtatct gaaacacaag acttgccctat cccatagtac atcatattac ctgtcaagct    300
atgctacccc acagaaatac cccaaaagt gaagtgaata aatgaaaatt actggtaact    360
tcaccccata acaaaactta taatttctgt agccaatgaa agtaaaccac attcaatgtt    420
ccgagattta gtatacttgc ccctataaga aacgaaggat ttcagettcc ttaccccatg    480
aacagaaatc ttccatttac cccccactgg agagatccgc ccaacgaac agataataga    540
aaaaagaaat tcggacaaat agaacacttt ctcagccaat taaagtcatt ccatgcactc    600
cctttagctg ccgttccatc cctttgttga gcaacacccat cgttagccag tacgaaagag    660
gaaaacttaac cgataccttg gagaaatcta aggcgcgaat gagttagacc tagatatcct    720
tagtgaagggt ttgttccgat acttctccac attcagtcac agatgggcag ctttgttatc    780
atgaagagac ggaacagggtt attaagggtt aaccgccaac ttatataaag acaacatgtc    840
cccagtttaa agtttttctt tcctattctt gtatcctgag tgaccgttgt gtttaataata    900
acaagttcgt tttaacttaa gacccaaacc agttacaaca aattataacc cctctaaca    960
ctaaagtcca ctcttatcaa actatcaaac atcaaaacat atgagattcc catccatctt   1020
cactgctgtt ttgttcgttg cttctctcgc tttggtgtgt ccagttaaca ctactactga   1080
agatgaaact gctcaaatcc cagctgaagc tggtatcgtt tactccgact tggaagggtga   1140
tttcgacgtt gctgttttgc cattctccaa ctccaccaac aacggattgt tgttcattaa   1200
caccaccatt gcttccatcg ctgctaagga agaaggtgtt tccttggaac agagagagga   1260
acagaatgga atgaagttgg agagagttgt catcgtttct agacacgggtg ttagagctcc   1320
caccaaatcc actccaatca tgaagaacgt cccccagat cagtggccac aatgggacgt   1380
cccactgggc tggttgactc cactgggtgg agaacttgct t                                     1421
```

<210> SEQ ID NO 8
<211> LENGTH: 1513
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Promoter variant fused to phytase reporter gene

<400> SEQUENCE: 8

```
ttccccgaaa agtgccacct gacgtcgttg gcgcgccaag gactagtccc ggaagatctt    60
tagccagtag gaaagaggaa acttaaccga taccttgag aaatctaagg cgcgaatgag    120
tttagcctag atatccttag tgaagggttg ttcgatacgt gtttaaaact agccagtagc    180
aaagaggaaa cttaaccgat accttgagga aatctaaggc gcgaatgagt ttagcctaga    240
tatccttagt gaagggttgt tccgatacta cgtaaacacct gggcatgact ttctaagtta    300
gcaagtcacc aagagggtcc tatttaacgt ttggcgggtat ctgaaacaca agacttgctt    360
```

US 2011/0129874 A1

Jun. 2, 2011

16

-continued

atcccatagt	acatcatatt	acctgtcaa	ctatgtacc	ccacagaaat	accccaaaag	420
ttgaagtga	aaaatgaaa	ttactggta	cttcaccca	taacaaactt	aataatttct	480
gtagccaatg	aaagtaaacc	ccattcaatg	ttccgagatt	tagtatactt	gccctataa	540
gaaacgaag	atttcagctt	ccttacccca	tgaacagaaa	tcttcattt	acccccact	600
ggagagatcc	gccccaaacg	acagataata	gaaaaagaa	attcgacaa	atagaacact	660
ttctcagcca	attaaagtca	ttccatgcac	tccttttagc	tgccgttcca	tcctttgtt	720
gagcaacacc	atcgtttagc	agtacgaaag	aggaaactta	accgatacct	tggagaaatc	780
taaggcgcg	atgagtttag	cctagatata	cttagtgaag	ggtgttccg	atacttctcc	840
acattcagtc	atagatgggc	agctttgtta	tcatgaagag	acggaaacgg	gcattaaggg	900
ttaaccgcca	aattatataa	agacaacatg	tccccagttt	aaagtttttc	tttcctattc	960
ttgtatcctg	agtgaccgtt	gtgtttaata	taacaagttc	gttttaactt	aagacaaaa	1020
ccagttacaa	caaattataa	ccctctaaa	cactaaagtt	cactcttata	aaactatcaa	1080
acatcaaac	atatgagatt	cccatccatc	ttcactgctg	tttgttccg	tgcttctccc	1140
gctttggctg	ctccagttaa	cactactact	gaagatgaaa	ctgctcaaat	cccagctgaa	1200
gctgttatcg	gttactccga	cttgggaaggt	gatttcgacg	ttgctgtttt	gccattctcc	1260
aactccacca	acaacggatt	gttgttcatt	aacaccacca	ttgcttccat	cgctgctaag	1320
gaagaagggt	tttccttggg	aaagagagag	gaacagaatg	gaatgaagtt	ggagagagtt	1380
gtcatcgttt	ctagacacgg	tgtagagct	cccaccaa	tcactccaat	catgaagaac	1440
gtcaccacag	atcagtgccc	acaatgggac	gtccactgg	gctggttgac	tcacgtgggt	1500
ggagaacttg	tct					1513

<210> SEQ ID NO 9

<211> LENGTH: 1608

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Promoter variant fused to phytase reporter gene

<400> SEQUENCE: 9

ttcccga	agtgccac	gacgtcgtt	gcgcgcca	gactagtta	gccagtac	60
aagaggaa	ttaaccgata	ccttgagaa	atctaaggcg	cgaatgagtt	tagcctagat	120
atccttagt	aagggttgtt	ccgatacgag	atcttttagc	agtacgaaag	aggaaactta	180
accgatacct	tggagaaatc	taaggcgcg	atgagtttag	cctagatata	cttagtgaag	240
ggtgttccg	atacggttta	aacttagcca	gtacgaaa	ggaaacttaa	ccgatacctt	300
ggagaaatct	aaggcgcgaa	tgagtttagc	ctagatatcc	ttagtgaagg	gttgttccga	360
tactacgtaa	cacctgggca	tgactttcta	agtagcaag	tcaccaagag	ggctctattt	420
aacgtttggc	ggtatctgaa	acacaagact	tgcttatccc	atagtacatc	atattacctg	480
tcaagctatg	ctacccca	gaaatacccc	aaaagttgaa	gtgaaaaaat	gaaaattact	540
ggtaacttca	cccataaca	aacttaataa	tttctgtagc	caatgaaagt	aaacccatt	600
caatgttccg	agatttagta	tacttgcccc	tataagaaac	gaaggatttc	agcttcctta	660
ccccatgaac	agaaatcttc	catttacc	ccactggaga	gatccgcca	aacgaacaga	720
taatagaaaa	aagaaattcg	gacaaataga	acactttctc	agccaattaa	agtcattcca	780

US 2011/0129874 A1

Jun. 2, 2011

17

-continued

tgcactccct ttagctgccc ttccatccct ttgttgagca acaccatcgt tagccagtag	840
gaaagaggaa acttaaccga taccttgagg aaatctaagg cgcgaatgag ttagccctag	900
atatccttag tgaagggttg ttccgatact tctccacatt cagtcataga tgggcagctt	960
tggtatcatg aagagacgga aacgggcatt aagggttaac cgccaaatta tataaagaca	1020
acatgtcccc agtttaaagt ttttctttcc tattcttgta tctgagtga cgttggtgtt	1080
taatataaca agttcgtttt aacttaagac caaaaccagt tacaacaaat tataaccct	1140
ctaaacacta aagttcactc ttatcaaact atcaaacatc aaaacatatg agattcccat	1200
ccatcttcac tgctgttttg ttgctgctt cctccgcttt ggctgctcca gttaacacta	1260
ctactgaaga tgaactgct caaatcccag ctgaagctgt tatcggttac tccgacttgg	1320
aagggtgatt cgacgttgct gttttgccat tctccaactc caccaacaac ggattgttgt	1380
tcattaacac caccattgct tccatcgctg ctaaggaaga aggtgtttcc ttggaaaaga	1440
gagaggaaca gaatggaatg aagttggaga gagttgtcat cgtttctaga caggtgttga	1500
gagctccac caaattcact ccaatcatga agaacgtcac ccagatcag tggccacaat	1560
gggagctccc actgggctgg ttgactccac gtggtggaga acttgtct	1608

<210> SEQ ID NO 10
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 10

ctgctctagc cagtttgctg	20
-----------------------	----

<210> SEQ ID NO 11
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 11

gccgcccagt cctgctcgt	20
----------------------	----

<210> SEQ ID NO 12
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 12

gaaaagtgcc acctgagtc gttactcaaa aagtc	35
---------------------------------------	----

<210> SEQ ID NO 13
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 13

gaaaagtgcc acctgcggtg tctgaaacac aagac	35
--	----

US 2011/0129874 A1

Jun. 2, 2011

18

-continued

```

<210> SEQ ID NO 14
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 14

gaaaagtgcc acctgtgaaa attactggta acttc          35

<210> SEQ ID NO 15
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 15

gaaaagtgcc acctgtctata agaaacgaag gatatt          35

<210> SEQ ID NO 16
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 16

gaaaagtgcc acctgaaaga aattcggaca aatag          35

<210> SEQ ID NO 17
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 17

gaaaagtgcc acctgttctc cacattcagt catag          35

<210> SEQ ID NO 18
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 18

gaaaagtgcc acctgcagtt taaagttttt ctttc          35

<210> SEQ ID NO 19
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 19

gaaaagtgcc acctgttata acccctctaa acact          35

<210> SEQ ID NO 20
<211> LENGTH: 43
<212> TYPE: DNA

```

US 2011/0129874 A1

Jun. 2, 2011

19

-continued

```

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 20

aaccagccca gtgggacgtc ccattgtggc cactgatctg ggg          43

<210> SEQ ID NO 21
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 21

ccccgaaaag tgccacctga cgtc          24

<210> SEQ ID NO 22
<211> LENGTH: 41
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 22

actgaatgtg gagaacgatg gtgttgctca acaaagggat g          41

<210> SEQ ID NO 23
<211> LENGTH: 41
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 23

ttgagcaaca ccacgtttct ccacattcag tcatagatgg g          41

<210> SEQ ID NO 24
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 24

aacagataat agaaattctc cacattcagt catagatggg          40

<210> SEQ ID NO 25
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 25

actgaatgtg gagaatttct attatctgtt cgtttgggag          40

<210> SEQ ID NO 26
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

```

US 2011/0129874 A1

Jun. 2, 2011

20

-continued

<400> SEQUENCE: 26

ttagtataact tgcccttctc cacattcagt catagatggg 40

<210> SEQ ID NO 27
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 27

actgaatgtg gagaaggga agtatactaa atctcggaac 40

<210> SEQ ID NO 28
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 28

aacagataat agaaattagc cagtacgaaa gaggaactt 40

<210> SEQ ID NO 29
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 29

tttcgtactg gctaatttct attatctgtt cgtttggcg 40

<210> SEQ ID NO 30
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 30

ttagtataact tgcccttagc cagtacgaaa gaggaactt 40

<210> SEQ ID NO 31
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 31

tttcgtactg gctaaggga agtatactaa atctcggaac 40

<210> SEQ ID NO 32
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 32

ttagtataact tgcccaaaga aattcggaca aatagaacac 40

US 2011/0129874 A1

Jun. 2, 2011

21

-continued

<210> SEQ ID NO 33
<211> LENGTH: 41
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 33

tgtccgaatt tctttgggca agtatactaa atctcggaac a 41

<210> SEQ ID NO 34
<211> LENGTH: 41
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 34

taggctaaac tcattcgcg ctcggtaagt ttctcttttc g 41

<210> SEQ ID NO 35
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 35

gcgcgaatga gtttagccta gatatacctt 29

<210> SEQ ID NO 36
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 36

aagtatcgga acaactaaac tcattcgcg cttagatttc 40

<210> SEQ ID NO 37
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 37

gtttagttgt tccgatactt ctccacattc 30

<210> SEQ ID NO 38
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 38

aagatgcttt tctgtgactg gtgagtactc aaccaagtca 40

<210> SEQ ID NO 39
<211> LENGTH: 47
<212> TYPE: DNA

US 2011/0129874 A1

22

Jun. 2, 2011

-continued

```

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 39

ttccccgaaa agtgccacct gacgtcggtg gcgcgccaag gactagt      47

<210> SEQ ID NO 40
<211> LENGTH: 46
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer

<400> SEQUENCE: 40

agacaagtgc tccaccacgt ggagtcaacc agcccagtg gacgtc      46

<210> SEQ ID NO 41
<211> LENGTH: 9667
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: complete plasmid sequence

<400> SEQUENCE: 41

ccttagacat gactgttctt cagttcaagt tgggcactta cgagaagacc ggtcttgcta      60
gattctaate aagaggatgt cagaatgcc a ttgctctgag agatgcaggc ttcatttttg      120
atactttttt atttgaacc tatatagtat aggatttttt ttgtcatttt gtttcttctc      180
gtacgagctt gctcctgatc agcctatctc gcagctgatg aatatcttgt ggtagggggt      240
tgggaaaatc attcgagttt gatgtttttc ttggtatttc ccactcctct tcagagtaca      300
gaagattaag tgagacgttc gtttgtgcaa gcttcaacga tgccaaaagg gtataataag      360
cgtcatttgc agcattgtga agaaaactat gtggcaagcc aagcctgcga agaattgatt      420
ttaagtgtga ctttgatgta ttcaattgat taagccataa ttctcgagta tctatgattg      480
gaagtatggg aatgggtgata ccgcattctt tcagtgctct gaggtctctc atcagattat      540
gcccactaa agcaaccgga ggaggagatt tcatggtaaa tttctctgac ttttggtcat      600
cagtagactc gaactgtgag actatctcgg ttatgacagc agaaatgtcc ttcttgagga      660
cagtaaatga agtcccacca ataaagaaat ccttgttatc aggaacaaac ttcttgtttc      720
gaactttttc ggtgccttga actataaaat gtagagtgga tatgtcgggt aggaatggag      780
cgggcaaatg cttaccttct ggaccttcaa gaggtatgta gggttttag atactgatgc      840
caacttcagt gacaacgttg ctatttcgtt caaaccatc cgaatccaga gaaatcaaag      900
ttgtttgtct actattgatc caagccagtg cggctctgaa actgacaata gtgtgctcgt      960
gttttgaggt catctttgta tgaataaate tagtctttga tctaaataat cttgacgagc      1020
cagacgataa taccaatcta aactctttaa acgttaaagg acaagtatgt ctgcctgtat      1080
taaaccccaa atcagctcgt agtctgatcc tcatcaactt gaggggcact atcttgtttt      1140
agagaaattt gcggagatgc agatcttaag ggcgaattcg cgcccgctaa attcaattcg      1200
ccctatagtg agtcgtatta caattcactg gccgtcggtt tacaacgtcg tgactgggaa      1260
aaccctggcg ttaccacact taatcgctt gcagcacatc cccctttcgc cagctggcgt      1320
aatagcgaag agggccgcac cgatcgccct tcccaacagt tgcgcagcct atacgtacgg      1380

```

US 2011/0129874 A1

23

Jun. 2, 2011

-continued

cagtttaagg	tttacaccta	taaaagagag	agccgttatc	gtctgtttgt	ggatgtacag	1440
agtgatatta	ttgacacgcc	ggggcgacgg	atgggtgatcc	ccctggccag	tgcacgtctg	1500
ctgtcagata	aagtctcccg	tgaactttac	ccggtggtgc	atctcgggga	tgaaagctgg	1560
cgcagtatga	ccaccgatat	ggccagtgtg	ccggtctccg	ttatcgggga	agaagtggct	1620
gattctcagcc	accgcgaaaa	tgacatcaaa	aacgccatta	acctgatgtt	ctggggaata	1680
taaatgtcag	gcatgagatt	atcaaaaagg	atcttcacct	agatcctttt	cacgtagaaa	1740
gccagtccgc	agaaacgggtg	ctgaccccg	atgaatgtca	gctactgggc	tatctggaca	1800
agggaaaacg	caagcgcaaa	gagaaagcag	gtagcttgca	gtgggcttac	atggcgatag	1860
ctagactggg	cggttttatg	gacagcaagc	gaaccggaat	tgccagctgg	ggcgccctct	1920
ggtaagggtg	ggaagccctg	caaagtaaac	tggatggctt	tcttgccgcc	aaggatctga	1980
tggcgcaggg	gatcaagctc	tgatcaagag	acaggatgag	gacgttttcg	catgattgaa	2040
caagatggat	tgcacgcagg	ttctccggcc	gcttgggtgg	agaggctatt	cggctatgac	2100
tgggcacaac	agacaatcgg	ctgctctgat	gccgcgtgt	tccggctgtc	agcgcagggg	2160
cgcgccgttc	ttttgtcaa	gaccgacctg	tccggtgccc	tgaatgaact	gcaagacgag	2220
gcagcgcggc	tatcgtggct	ggccaacgag	ggcgttcctt	gcgcagctgt	gctcgacgtt	2280
gtcactgaag	cgggaaggga	ctggctgcta	ttgggcgaag	tgccggggca	ggatctcctg	2340
tcattctcacc	ttgctcctgc	cagagaaagta	tccatcatgg	ctgatgcaat	gcggcggctg	2400
catacgtctg	atccggtctc	ctgcccattc	gaccaccaag	cgaacatcgc	catcgagcga	2460
gcacgtactc	ggatggaagc	cggctctgtc	gatcaggatg	atctggacga	agagcatcag	2520
gggctcgcgc	cagccgaact	gttcgccagg	ctcaaggcga	gcatgcccca	cggcgaggat	2580
ctcgtcgtga	cccattggcg	tgctctgttg	ccgaatatca	tggtggaaaa	tggccgcttt	2640
tctggattca	tgaactgtgg	ccggctgggt	gtggcggaac	gctatcagga	catagcgttg	2700
gtacaccgtg	atattgctga	agagcttggc	ggcgaatggg	ctgaccgctt	cctcgtgctt	2760
tacggtatcg	ccgctcccca	ttcgcagcgc	atcgccttct	atcgccttct	tgacgagttc	2820
ttctgaatta	ttaacgctta	caatttctcg	atgcggtatt	ttctccttac	gcatctgtgc	2880
ggatatttcac	accgcatcag	gtggcacttt	tccgggaaat	gtgcgcggaa	cccctatttg	2940
tttatttttc	taaatacatt	caaatatgta	tccgctcatg	agattatcaa	aaaggatctt	3000
cacctagatc	cttttaaat	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	3060
aacttggtct	gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	3120
atttcgttca	tccatagtgt	cctgactccc	cgtcgtgtag	ataactacga	tacgggaggg	3180
cttaccatct	ggccccagtg	ctgcaatgat	accgcgagac	ccacgctcac	cggctccaga	3240
tttatcagca	ataaaccagc	cagccggaag	ggccgagcgc	agaagtggtc	ctgcaacttt	3300
atccgcctcc	atccagtcta	ttaattgttg	ccgggaagct	agagtaagta	gttcgccagt	3360
taatagtttg	cgcacagttg	tggccattgc	tacaggcatc	gtgggtgtcac	gctcgtcgtt	3420
tggtatggct	tcattcagct	ccggttccca	acgatcaagg	caggttacat	gatcccccac	3480
gttgtgcaaa	aaagcgggta	gtcctctcgg	tcctccgata	gttgtcagaa	gtaagttggc	3540
cgcagtgtta	tcactcatgg	ttatggcagc	actgcataat	tctcttactg	tcattgccatc	3600
cgtaaagatgc	ttttctgtga	ctggtgagta	ctcaaccaag	tcattctgag	aatagtgtat	3660

US 2011/0129874 A1

24

Jun. 2, 2011

-continued

gcggcgaccg agttgctctt gcccgcgctc aatacgggat aataccgcgc cacatagcag	3720
aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt	3780
accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc	3840
ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa	3900
gggaataagg gcgacacgga aatgttgaat actcatactc ttcttttttc aatattattg	3960
aagcatttat caggggttatt gtctcatgac caaaatccct taacgtgagt ttctgttcca	4020
ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg	4080
cgtaatctgc tgttgcaaaa caaaaaaacc accgctacca gcggtgggtt gtttgccgga	4140
tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa	4200
tactgttctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc	4260
tacatacctc gctctgctaa tctgtttacc agtggtctgt gccagtggcg ataagtcgtg	4320
tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggg cgggctgaac	4380
gggggggttc tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct	4440
acagcgtgag ctatgagaaa gccccacgct tccgaaggg agaaaggcgg acaggtatcc	4500
ggtaacggcg agggctcgaa caggagagcg cagcagggag cttccagggg gaaacgcctg	4560
gtatctttat agtctgtctg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg	4620
ctcgtcaggg gggcgagacc tatggaaaaa cgcagcaaac gcggcctttt tacggttcct	4680
ggccttttgc tggccttttg ctcacatgtt ctttctctcg ttatccccctg attctgtgga	4740
taaccgtatt accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg	4800
cagcagtgca gtgagcgagg aagcgggaaga gcgcccaata cgcaaaccgc ctctccccgc	4860
gcgttgcccg attcattaat gcagctggca cgacaggttt cccgactgga aagcgggcag	4920
tgagcgcaac gcaattaatg tgagttagct cactcattag gcaccccgagg ctttacactt	4980
tatgcttcgg gctcgtatgt tgtgtggaat tgtgagcggg taacaatttc acacaggaaa	5040
cagctatgac catgattacg ccaagctcag aattaaccct cactaaaggg actagtcctg	5100
cagggtttaa cgaattcgcc cttagatcta acatccaaag acgaaagggt gaatgaaacc	5160
tttttgccat ccgacatcca caggctcatt ctcacacata agtgccaaac gcaacaggag	5220
gggatacact agcagcagc cgttgcaaac gcaggacctc cactcctctt ctcctcaaca	5280
ccacttttg ccacgaaaa accagcccag ttattgggct tgattggagc tcgctcatc	5340
caattccttc tattaggcta ctaacacat gactttatta gcctgtctat cctggcccc	5400
ctggcgaggt tcattgtttg ttatttccga atgcaacaag ctccgcatta caccgaaca	5460
tcactccaga tgagggtctt ctgagtgtgg ggtcaaatag tttcatgttc cccaaatggc	5520
ccaaaactga cagtttaaac gctgtcttg aacctaatat gacaaaagcg tgatctcatc	5580
caagatgaac taagtttggt tcgttgaaat gctaacggcc agttggtcaa aaagaaactt	5640
ccaaaagtgc gcataccgtt tgtcttggtt ggtattgatt gacgaatgct caaaaataat	5700
ctcattaatg cttagcgcag tctctctatc gcttctgaac cccgggtgcac ctgtgccgaa	5760
acgcaaatgg ggaacacccc gctttttgga tgattatgca ttgtctccac attgtatgct	5820
tccaagatcc tgggtgggaat actgctgata gcctaacgtt catgatcaaa atttaactgt	5880
tctaaccctt acttgacagc aatatataaa cagaagggaag ctgccctgtc ttaaaccttt	5940

US 2011/0129874 A1

Jun. 2, 2011

25

-continued

ttttttatca	tcattattag	cttactttca	taattgcgac	tggttccaat	tgacaagctt	6000
ttgatttttaa	cgacttttaa	cgacaacttg	agaagatcaa	aaaacaacta	attattcgaa	6060
acgatggcta	tccccgaaga	gtttgatata	cccctgtagt	aaacagggag	ataccgtacg	6120
gaggtctgaa	ttccctacag	aagtagctgt	aaaaattcag	aattcgcaac	aacctataat	6180
tgagttaagt	gcctttccaa	gctaaaaagt	ttgaggttat	aggggcttag	catccacacg	6240
tcacaatctc	gggtatcgag	tatagtatgt	agaattacgg	caggagggtt	ccaatgaac	6300
aaaggacagg	ggcacgggtga	gctgtcgaag	gtatccattt	tatcatgttt	cgtttgtaca	6360
agcacgacat	actaagacat	ttaccgtatg	ggagttgttg	tcctagcgta	gttctcgctc	6420
ccccagcaaa	gctcaaaaaa	gtacgtcatt	tagaatagtt	tgtgagcaaa	ttaccagtcg	6480
gtatgctacg	ttagaaaggc	ccacagtatt	cttctaccaa	aggcgtgcct	ttgttgaact	6540
cgatccatta	tgagggtctc	cattattccc	cgatttttta	ttactctgaa	caggaataaaa	6600
aagaaaaaac	ccagtttagg	aaattatccg	ggggcgaaga	aatacgcgta	gcgttaatcg	6660
acccacgctc	cagggttttt	ccatggaggt	ttctggaaaa	actgacgagg	aatgtgatta	6720
taaatccctt	tatgtgatgt	ctaagacttt	taaggtagcg	ccgatgtttg	cctattacca	6780
tcatagagac	gtttcttttc	gaggaatgct	taaacgactt	tgtttgacaa	aaatgttgcc	6840
taagggtctc	atagtaaacc	atttggaaga	aagatttgac	gacttttttt	ttttggattt	6900
cgatcctata	atcctctctc	ctgaaaagaa	acataataat	agatatgtat	tattcttcaa	6960
aacattctct	tgttcttgtg	cttttttttt	accatatata	ttactttttt	ttttctctca	7020
gagaaacaag	caaaacaaaa	agctttttct	ttcactaacg	tatatgatgc	ttttgcaagc	7080
tttccttttc	cttttggtcg	gttttgacgc	caaaatatct	gcatcaatga	caaacgaaac	7140
tagcgataga	cctttggtcc	acttcacacc	caacaagggc	tggatgaatg	acccaaatgg	7200
gttgtgtgac	gatgaaaaag	atgccaaatg	gcatactgtac	tttcaataca	acccaaatga	7260
caccgtatgg	ggtagcccat	tgttttgggg	ccatgctact	tccgatgatt	tgactaattg	7320
ggaagatcaa	ccattgcta	tcgctcccaa	gcgtaacgat	tcagggtgctt	tctctggctc	7380
catgggtggt	gattacaaca	acacgagtg	gtttttcaat	gatactattg	atccaagaca	7440
aagatgcggt	gcgatttgga	cttataacac	tcctgaaagt	gaagagcaat	acattagcta	7500
ttctcttgat	ggtaggttaca	cttttactga	ataccaaaag	aacctgtgtt	tagctgccaa	7560
ctccactcaa	ttcagagatc	caaagggtgt	ctggtagtaa	ccttctcaaa	aatggattat	7620
gacggctgcc	aatcacaaag	actacaaaat	tgaattttac	tcctctgatg	acttgaagtc	7680
ctggaagcta	gaatctgcat	ttgccaatga	agggtttctta	ggctaccaat	acgaatgtcc	7740
aggtttgatt	gaagtcccaa	ctgagcaaga	tccttccaaa	tcttattggg	tcagttttat	7800
ttctatcaac	ccagggtcac	ctgctggcgg	ttccttcaac	caatattttg	ttggatcctt	7860
caatggtact	cattttgaag	cgtttgacaa	tcaatctaga	gtggtagatt	ttggaagga	7920
ctactatgcc	ttgcaaaact	tcttcaacac	tgaccaaac	tacggttcag	cattaggtat	7980
tgctctgggt	tcaaaactgg	agtacagtgc	ctttgtccca	actaaccat	ggagatcatc	8040
catgtctttg	gtccgcaagt	ttcttttgaa	caactgaat	caagctaata	cagagactga	8100
attgatcaat	ttgaaagccg	aaccaatatt	gaacattagt	aatgctggtc	cctggtctcg	8160
ttttgtact	aacacaaact	taactaaggc	caattcttac	aatgtcgatt	tgagcaactc	8220

US 2011/0129874 A1

26

Jun. 2, 2011

-continued

gactggtacc	ctagagtttg	8280
atccgtcttt	gccgacttat	8340
gagaatgggt	tttgaagtca	8400
caagtttgtc	aaggagaacc	8460
caagctctgag	aacgacctaa	8520
ggaaatgtac	ttcaacgatg	8580
taacgctcta	ggatctgtga	8640
gttccaagta	aggggaagtaa	8700
aaaagccatt	ctaaagggct	8760
agaatacctc	caaaccattg	8820
cctggaattt	gttctttatg	8880
tttgcaatag	acatcatttt	8940
actactagga	tggacctttt	9000
aaagtattag	ttatgtgaaa	9060
aaaaaagcgt	tgtcttctac	9120
accaagctat	ctgtaacagg	9180
gaaggaaact	caatcgggaa	9240
cgggaataaaa	atcgaaacag	9300
cctgtgactg	gaggggaaga	9360
aaaaacaata	atggagctaa	9420
gaagagcgaa	aacaaatcgc	9480
tccgatttta	aactgtttta	9540
cccatgacag	actccggtaa	9600
ccaaaggccc	caataaaact	9660
accgccc		9667

<210> SEQ ID NO 42
<211> LENGTH: 5741
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: complete plasmid sequence

<400> SEQUENCE: 42

agatcttttt	tgtagaaatg	60
atctggctcc	gttgcaactc	120
cttaaatgtg	gagtaatgga	180
gttacgctcc	ctaggaaatt	240
aatgctcttc	ccagcattac	300
ccagggatgg	aaaagtcgcc	360
gattattgga	aaccaccaga	420
ctcctgaccc	aaagacttta	480

US 2011/0129874 A1

Jun. 2, 2011

27

-continued

tatttcgaaa	tgccctcctaa	acatcggctg	gagcagagta	tacagcccat	ggcttctcaa	540
caaatagtag	ccggtaataa	ggttattctg	ccgaatccaa	aagtagatgc	aaaatctacc	600
ccaaacattt	cagttcagaa	gagaagaaga	gtcaccagag	cttgtgatga	atgtcggaaa	660
aagaaggtag	aatgtgatgg	tcaacaacca	tgcatctatt	gtaccgttta	ttcctatgag	720
tgcaacttaca	gccaaccttc	cagtaagaag	agacagggac	aatctctgag	tctgagtgtc	780
ccgtcaaca	ttaatgcaac	aagttccgta	caaaaatctg	taaaacctcc	tgaatcgat	840
ttocaaagga	tgagagacgc	actcaaatat	tacgaagatc	ttttaaacca	gttgatatac	900
cccaacagtg	ctccaactgt	tcgagttaat	ccgattcgtc	tagcatcgat	cttaaaacaa	960
ttgagagccg	ataaatcaag	tgatgaatta	atttcagtca	aggctctttc	tgacaattac	1020
attgagatgc	ttcacaaaa	gatgcaacaa	cctgtacagc	agccagctcc	tccttcattg	1080
gggcaaggag	ggctcttctc	taatcacagt	cccaatcata	ataatgcttc	tattgatggg	1140
tcctatagaat	ctaacttagg	gagggaaata	cgtatcatat	tacctccgag	agatattgag	1200
ctgaagctta	tctacaagac	ttgggacaac	gcgtgtgtac	ttttccgctt	ttatcacaga	1260
cccgcattta	ttgaggacct	gaatgagtta	tatgaaacag	atttggcaaa	ctacaccaat	1320
aaacaacaaa	gggtttttacc	tcttgatat	tcgggtgatgg	cttgtgtgtc	tcttttttgc	1380
aagactgatg	ggattaatca	cggccaaaag	agctccaagc	caaagactc	ttctgatgaa	1440
agtcctatag	acgatgaggg	ttacaagtat	tttattgccg	caagaaaact	aatagatatac	1500
acggatacca	gggataccta	cgggaattcag	actattgtta	tgctgatcat	ttttttacaa	1560
tgttcggtc	gtctttcaac	atgctattct	tattattggca	ttgctctaag	agctgcattg	1620
agagaagggt	tgcatcgtca	gttgaactat	cctttcaatc	caattgagtt	agaaacaaga	1680
aagcgtcttt	tttgactat	ctataaaatg	gacatctatg	tcaatacaat	gctggggctt	1740
ccaagaacca	ttctgaaaga	ggatttcgac	caggaaatgc	ctatcgaaact	tgatgatgag	1800
aacattagtg	aaaccggata	taggttcgat	ttacaaggta	caaagtatac	cagttcagga	1860
atagccaatg	ctcacactag	attgatattc	ataatgaaga	aaattgtgaa	aaaattatat	1920
cctgtcaaac	tacagaaacc	aacctcaaac	agtggcgata	cccacttga	gaacaatgat	1980
ttattggctc	atgaaatcgt	tcatagaact	gagatggatc	tccaaaattg	ggccaatagt	2040
ctacctgcag	aactaaaacc	ggggatagaa	ccaccgaccg	agtattttaa	agctaacaga	2100
ttgcttcatt	tggcatacct	gcattgtcaag	attattctct	acaggccatt	tattcattac	2160
atctcagaaa	aggataaggt	tggaatatgt	tctatccctc	cgtcgcccga	agagatcact	2220
tctatcgaga	aagccaagaa	ttgtgtcaat	gttgccagaa	ttgttgtaa	actagccgaa	2280
gacatgatta	ataggaaaat	gttaagtggg	tcatattggg	tttccattta	taccattttt	2340
ttttccgtgg	catgtctggg	gtactatgtt	catttcgctc	caccgaagaa	agacaatgga	2400
gaactggatc	cccaatacat	ggaaatcaag	aaagatacag	agagtggag	agaggtctta	2460
aatatcctca	aagatagtag	tatggcgga	agaagaacgt	ataatattct	caactctttg	2520
tttgagcagt	taaacagaag	aactgcaaag	gtcaacctag	caaaggcaca	gcaaccacca	2580
tcagggttga	ataaccagc	tgctaccag	tatcagaaac	aggggtgaaca	caggcagtta	2640
caaccaagta	actattctgg	aactgtgaaa	tctgtggacc	cagagaatat	cgattactct	2700
tcctttgggt	ctcagtttga	aaacactaac	atcgaagatg	gttcctcaaa	tacaaagatt	2760

US 2011/0129874 A1

Jun. 2, 2011

28

-continued

gatcagaaaag tgaatggggg gaactacatc gatgggtgtg ttacagggat caacctaaat	2820
atgcctaata tctcagaaac ttctaact caaggtatcg ataaccagc atttcaaagt	2880
ataaacaatt ctaatttgaa caataatttt gtacaaacaa agtacattcc cgcatgatg	2940
gaccagctag atatgaaaaa ttccggaaga ttcttccac cttacatgct gaactccaac	3000
aagggtgaac agggacaaaa tgaagggaac ctatcaggcc aaccatcctc gtcgaatact	3060
cctgatggat cacaacctgt gacagttctg gatggattat acccgttgca gaatgataat	3120
aataataacc acgaccagc aaattcaaag tctgttgtaa ataacagtaa ctcggtagaa	3180
aacttactac agaactttac aatgggtccc tcggggttgt catcaacagt gcaaaatcct	3240
gaagcgcccc aaaaattcaa taatcatatg tcaacatat cgaatatgaa tgatccaaga	3300
agagctagcg tagctacatc agatggatcc aatgacatgg atcatcatag ccaaggcccg	3360
ataacaaaag atttgaaacc gttgagcaac tacgagtttg acgatctctt ctttaatgat	3420
tggaacctg cgccagatac aataaatttt gacagttaag tcgacctca tcatcatcat	3480
cattgagttt tagccttaga catgactgtt cctcagttca agttgggac ttacgagaag	3540
accggtcttg ctgatttcta atcaagagga tgcagaatg ccatttgcct gagagatgca	3600
ggcttcattt ttgatacttt tttatttgta acctatatag tataggattt tttttgcat	3660
tttgtttctt ctgtacgag cttgtcctg atcagcctat ctgcagctg atgaatatct	3720
tggtgtagg gtttgggaaa atcatcagc tttgatgttt ttcttggtat tccccctcc	3780
tcttcagagt acagaagatt aagtgaagcc ttctgttgtg cggatcccc acacaccata	3840
gcttcaaat gtttctact cttttttact cttccagatt ttctcggact ccgcatcg	3900
ccgtaccact tcaaacacc caagcacagc atactaaatt ttccctcttt ctccctctag	3960
gggtgcgtta attaccgta ctaaagggtt ggaagaaa aaagagaccg cctcgtttct	4020
ttttctctg cgaagagc aataaaaatt tttatcacgt ttctttttct tgaattttt	4080
ttttttagt tttttctct tcatgacct ccattgatat ttaagttaat aaacggtctt	4140
caatttctca agtttcagt tcatttttct tgttctatta caacttttt tacttcttgt	4200
tcattagaaa gaaagcatag caatctaac taaggcggt gttgacaatt aatcatcgcc	4260
atagtatac ggcatagtat aatcagacaa ggtgaggaa taaacctgg ccaagttgac	4320
cagtgcctt ccggtgctc ccgcgcgca cgtcgcgga gcggtcgagt tctggaccga	4380
ccggtcggg ttctccggg acttcgtgga ggacgactt gccggtgtg tccgggacga	4440
cgtgacctg ttcacacg cggtccagga ccaggtgtg ccggacaaca cctggcctg	4500
gggtgtggg cgccgctg acgagctgta ccgcgagtg tcggaggtcg tgtccacgaa	4560
ctccgggac gcctccggc cgccatgac cgagatcgcc gagcagccgt gggggcgga	4620
gttcgacctg cgcgacccg ccggcaactg cgtgcactt gtggcgagg agcaggactg	4680
acacgtccga cggcgccca cgggtcccag gcctcgaga tccgtcccc ttttccttg	4740
tcgatatcat gtaattagt atgtcacgt tacattcac cctcccccc acatccgctc	4800
taaccgaaaa ggaaggagt agacaacctg aagtctagt ccctatttat tttttatag	4860
ttatgttagt attaagaacg ttatttatat ttcaatttt tcttttttt ctgtacagac	4920
gcgtgtacg atgtaacatt atactgaaa ccttgcttga gaaggttttg ggacgctcga	4980
aggctttaat ttgcaagctg gagaccaaca tgtgagcaaa aggcagcaa aaggccagga	5040

US 2011/0129874 A1

Jun. 2, 2011

29

-continued

accgtaaaaa ggccgcgttg ctggcggtttt tccataggct ccgccccct gacgagcatc	5100
acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg	5160
cgtttccccc tggaagctcc ctcgtagcgt ctctgttcc gaccctgcgc cttaccggat	5220
acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcaatgctca cgctgtagg	5280
atctcagttc ggtgtaggtc gtctgctcca agctgggctg tgtgcacgaa cccccgttc	5340
agcccgaccg ctgcgctta tccggttaact atcgtcttga gtccaaccg gtaagacacg	5400
acttatcgcc actggcagca gccactggtta acaggattag cagagcgagg tatgtaggcg	5460
gtgtacaga gttcttgaag tggtagccta actacggcta cactagaagg acagtatttg	5520
gtatctgcgc tctgtgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg	5580
gcaaacaaac caccgtggt agcgggtggt tttttgttg caagcagcag attacgcgca	5640
gaaaaaaagg atctcaagaa gatcctttga tctttctac ggggtctgac gctcagtgg	5700
acgaaaactc acgttaaggg attttggtca tgcagtgat c	5741

<210> SEQ ID NO 43
<211> LENGTH: 6282
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: complete plasmid sequence

<400> SEQUENCE: 43

agatcttttt tgtagaaatg tcttggtgtc ctctccaat caggtagcca tctctgaaat	60
atctggctcc gttgcaactc cgaacgacct gctggcaacg taaaattctc cggggtaaaa	120
cttaaatgtg gagtaatgga accagaaacg tctcttccct tctctctcct tccaccgcc	180
gttaccgtcc ctaggaaatt ttactctgct ggagagcttc ttctacggcc cccttgagc	240
aatgtctctc ccagcattac gttgcgggta aaacggagggt cgtgtaccgc acctagcagc	300
ccagggatgg aaaagtcccg gccgtcgctg gcaataatag cgggcggacg catgtcatga	360
gattattgga aaccaccaga atcgaatata aaaggcgaac acctttccca attttggttt	420
ctctgaccc aaagacttta aatttaattt atttgtccct atttcaatca attgaacaac	480
tatttcgaaa ccattgagcaa tctaccccca acttttggtt ccactagaca atctccagaa	540
gaccaatcac ctcccgctgc caaggagctg tcattcaatg ggaccacacc ctccagaaag	600
ctacgcttat ttgtctgtca gacatgtact cgagcatttg ctctgcagga acacttgaaa	660
cgacacgaaa ggtctcacac caaggagaaa cctttcagct gcggcatttg ttctcgtaaa	720
ttcagcgcgc gagatctgtt attgagacat gcccaaaaac tgcacagcaa ctgctctgat	780
gcggccataa caagactaag gcgcaaggca actcgtcggt cttctaattgc cgcgggttcc	840
atatctggtt ctactccggt gacaacgcca aatactatgg gtacgcccga agatggcgag	900
aaacgaaaaa ttcagaaact ggccggccgc cgggactcaa atgaacagaa actgcaactg	960
caacaacaac atctacagca acaaccacag ttgcaatacc aacaatctct taagcagcat	1020
gaaaatcaag tccagcagcc tgatcaagat ccattgatat ccccgagaat gcaattattc	1080
aatgattcca accatcacgt aaacaatttg tttgatcttg gactaagaag agcttccttc	1140
tccgccgcta gtggaaaata ttatgcccat tatgtgaata attttcaaca agatgcctct	1200
tctaccaatc caaatcaaga ttcaataat gccgaatttg agaattattga attttctacc	1260

US 2011/0129874 A1

Jun. 2, 2011

30

-continued

ccacaaatga	tgcccgttga	agatgctgaa	acttgatga	acaacatggg	tccaattccg	1320
aacttctctc	tcgatgtgaa	caggaacatt	ggtgatagct	ttacagatat	acaacacaag	1380
aattcagagc	ctattatata	cgaaccgccc	aaggacaccg	ctccaaacga	caagaagttg	1440
aatggctact	ctttttacga	agccccatc	aagccattag	aatccctatt	ttctgtcagg	1500
aatacaaaga	gaaacaagta	taaaacaaat	gacgactctc	cagacaccgt	ggataataac	1560
tccgcaccgg	ctgctaatac	cattcaagaa	cttgagtctt	ctttgaatgc	atccaagaat	1620
ttttgcttgc	caactgggta	ttccttctat	ggtaatttgg	accaacagac	tttctctaac	1680
acgttatcat	gcacttcttc	taatgccaca	atttcgcccc	ttctactcga	taactccatt	1740
aataataact	ccactagtga	cgtgagacca	gaatttagaa	cacaaagtg	cacctctgaa	1800
atgagtcaag	ccccctcccc	tctcaaaaa	aacaactcga	aatattccac	cgaagtctct	1860
tttaccagca	acatgcggtc	gtttattcac	tacgctcttt	ccaagtatcc	ttttattggg	1920
gtgcccactc	caactcttcc	ggagaacgaa	agactaaatg	aatatgctga	ttcattcacc	1980
aaccgtttct	taaatcatta	tcttttcata	catgtcacga	ttctcaaaga	atactccctt	2040
ttcaaggcaa	tttttagatga	gaatgagtcg	actaagaact	gggaaaaata	tcagttttac	2100
ttagagaacc	aacgaatata	aattgtttgt	cttctctctt	tggtggctac	gataggtgca	2160
gtactatcaa	acaacaaaaa	ggatgcttcg	aatttatacg	aagcttcaag	gcgttgattt	2220
catgtttact	tagattccag	gaaaaagata	cccacttcct	tgccgcgaaa	taacaatgac	2280
tctccacttt	ggctaattca	atccctgacg	ttatctgtta	tgtatgggtt	atttgcggac	2340
aatgacatta	gtttgaatgt	cgtgatcaga	caagttaacg	cacttaattc	tctgggtcaag	2400
acttcggggc	tgaataggac	ctcaattata	gatcttttca	acatcaacaa	acctttggat	2460
aatgaactct	ggaatcaatt	cgtgaaaata	gagtcaccgc	taaggacaat	ccacacgatt	2520
tttcaaatca	gttccaactt	aagcgecttg	tacaatatta	ttccatcggt	gaaaattgat	2580
gacctaatga	ttactctacc	agttcccaca	acactttggc	aagctgattc	ttttgtgaaa	2640
ttcaaaagtc	taagttacgg	aatcagatc	ccttttcaat	atacaagagt	actacagaat	2700
ttgattgatt	acaatcagcc	attgagcgat	ggaaaatttt	tgtatgaaaa	ccatgtaagt	2760
gagtttggac	tcatatgcct	acagaatgg	ctacaccaat	acagctattt	ccaaaaattg	2820
actgctgtca	ataacagaga	agatgcgcta	ttcacaaagg	ttgttaattc	acttcacagt	2880
tgggatagga	tgatttcgaa	ttctgatttg	tttccaaaga	agatatatca	gcagagttgc	2940
ttgattttgg	actcaaagtt	gcttaataat	ttcctgattg	tcaagagctc	attgaaagtt	3000
tcgaccggag	acgttagttc	ttgaataaag	ttaaagaaaa	acgtgtggct	taaaaactgg	3060
aatcaagtgt	gtgctatcta	ttataacagc	ttcatgaaca	ttcctgctcc	cagtattcaa	3120
aagaagtaca	atgacataga	gtttgtggat	gacatgatta	atttgagtct	aatcatcatc	3180
aagattatga	aactcatttt	ctataacaat	gtcaaagaca	attatgagga	tgaatatgac	3240
ttcaaatgtc	aagagttaaa	tttaacattt	gacaattttg	atgagaaaaa	atccttgaat	3300
ttgacaatat	tattcgatat	atttttgatg	atctacaaga	taattaccaa	ttacgagaag	3360
tttatgaaga	tcaaacacaa	gtttaattac	tacaattcta	attcgaatat	aagcttcttg	3420
catcatctcg	aactctcctc	ggttatcaat	aacacccaaa	tgaaccagaa	tgattatatg	3480
aaaacagata	ttgatgaaaa	gcttgatcag	cttttccaca	tctatcaaac	atttttccgg	3540

US 2011/0129874 A1

Jun. 2, 2011

31

-continued

ctgtatctgg atttagaaaa gtttatgaag ttcaaattca actatcatga ctttgagaca	3600
gagttttcaa gtctctcaat atccaatata ctgaacactc atgctgcttc taacaatgac	3660
acaaatgctg ctgatgctat gaatgccaaag gatgaaaaaa tatctccac aactttgaat	3720
agcgattac ttgctgatga aggaaatgaa aattccggtc gtaataacga ttcagaccgc	3780
ctgttcatgc tgaacgagct aattaatttt gaagtaggtt tgaaatttct caagataggt	3840
gagtcatttt ttgatttctt gtatgagaat aactacaagt tcatccactt caaaaactta	3900
aatgacggaa tgttccacat caggatatac ctagaaaacc gactagatgg tgggtgtctct	3960
agaacaaaa ctcatctcag aagaggatct gaatagcgcc gtcgaccatc atcatcatca	4020
tcattgagtt ttagccttag acatgactgt tcctcagttc aagttgggca cttacgagaa	4080
gaccggtctt gctagattct aatcaagagg atgtcagaat gccatttgcc tgagagatgc	4140
aggcttcatt tttgatactt ttttatttgt aacctatata gtataggatt tttttgtca	4200
ttttgtttct tctcgtacga gcttgctcct gatcagccta tctcgcagct gatgaatgc	4260
ttgtggtagg ggtttgggaa aatcattcga gtttgatgtt tttcttggtt tttccactc	4320
ctcttcagag tacagaagat taagtgcagc ctctgtttgt gcggatcccc cacacacat	4380
agcttcaaaa tgtttctact ccttttttac tcttcagat tttctcggac tccgcgcac	4440
gccgtaccac ttcaaacac ccaagcacag catactaaat ttccctctt tcttctctta	4500
gggtgtcgtt aattaccctg actaaagggt tggaaaagaa aaaagagacc gcctcgtttc	4560
tttttcttcg tcgaaaaagg caataaaaaat ttttatcacg tttcttttct ttgaaatttt	4620
tttttttagt tttttctct ttcagtgacc tccattgata ttaagttaa taaacggtct	4680
tcaatttctc aagtttcagt ttcatttttc ttgttctatt acaacttttt ttacttcttg	4740
ttcattagaa agaagacata gcaatctaata ctaagggcgg tgttgacaat taatcatcgg	4800
catagtatat cggcatagta taatacgaca aggtgaggaa ctaaaccatg gccaaagtga	4860
ccagtgcctg tccggtgctc accgcgcgcg acgtcgcgcg agcggtcgag ttctggaccg	4920
accggtcgcg gttctcccg gacttcgtgg aggacgactt cgcgggtgtg gtcgggacg	4980
acgtgacctt gttcatcagc gcgggtccagg accaggtggt gccggacaac accctggcct	5040
gggtgtgggt gcgcggcctg gacgagctgt acgccgagtg gtcggaggtc gtgtccacga	5100
acttccggga cgcctccggg ccggccatga ccgagatcgg cagcagccg tgggggcggg	5160
agttcgccct gcgcgaccgc gccggcaact gcgtgcactt cgtggccgag gagcaggact	5220
gacacgtccg acggcgcccc acgggtccca ggccctcggag atccgtcccc cttttctttt	5280
gtcgatatca tgtaattagt tatgtcacgc ttacattcac gccctcccc caccatccgt	5340
ctaaccgaaa aggaaggagt tagacaacct gaagtctagg tccctattta tttttttata	5400
gttatgttag tattaagaac gttatttata tttcaattt ttcttttttt tctgtacaga	5460
cgcgtgtacg catgtaacat tatactgaaa accttgcttg agaaggtttt gggacgctcg	5520
aaggctttaa tttgcaagct ggagaccaac atgtgagcaa aaggccagca aaaggccagg	5580
aaccgtaaaa aggcgcgctt gctggcgctt ttcoatagc tccgcccccc tgacgagcat	5640
cacaaaaatc gacgctcaag tcagaggtgg cgaacccga caggactata aagataccag	5700
gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgacctgcc gcttacggga	5760
tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcaatgctc acgctgtagg	5820

US 2011/0129874 A1

Jun. 2, 2011

32

-continued

tatctcagtt	cggtgtaggt	cgttcgctcc	aagctgggct	gtgtgcacga	accccccggt	5880
cagccccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	agtcacaccc	ggtaagacac	5940
gacttatcgc	cactggcagc	agccactggg	aacaggatta	gcagagcgag	gtatgtaggc	6000
ggtgtacag	agttcttgaa	gtggtggcct	aactacggct	acactagaag	gacagtattt	6060
ggtatctgcg	ctctgttgaa	gccagttacc	ttcgaaaaa	gagttggtag	ctcttgatcc	6120
ggcaaaaaa	ccaccgctgg	tagcggtggt	ttttttgttt	gcaagcagca	gattacgcgc	6180
agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	cggggtctga	cgtcagtggt	6240
aacgaaaact	cacgttaagg	gattttgtgc	atgcattgaga	tc		6282

<210> SEQ ID NO 44

<211> LENGTH: 3827

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Prm1 CDS controlled by GAP

<400> SEQUENCE: 44

agatcttttt	tgtagaaatg	tcttggtgtc	ctcgtccaat	caggtagcca	tctctgaaat	60
atctggctcc	gttgcaactc	cgaacgacct	gctggcaacg	taaaattctc	cggggtaaaa	120
cttaaatgtg	gagtaatgga	accagaaacg	tctcttcctt	tctctctcct	tcacacgccc	180
gttaccgtcc	ctaggaaatt	ttactctgct	ggagagcttc	ttctacggcc	cccttgacgc	240
aatgctcttc	ccagcattac	gttgccggta	aaacggaggt	cgtgtaccgc	acctagcagc	300
ccagggatgg	aaaagtcctg	gccgtcgtgt	gcaataatag	cgggcggacg	catgtcatga	360
gattattgga	aaccaccaga	atcgaatata	aaaggcgaac	acctttccca	attttggttt	420
ctctgacccc	aaagacttta	aatttaattt	attgtctcct	atttcaatca	attgaacaac	480
tatttcgaaa	tgcctectaa	acatcggctg	gagcagagta	tacagcccat	ggctttctca	540
caaatagtac	ccggtaataa	ggttattctg	ccgaatccaa	aagtagatgc	aaaatctacc	600
ccaaacattt	cagttcagaa	gagaagaaga	gtcaccagag	cttgtgatga	atgtcggaaa	660
aagaaggta	aatgtgatgg	tcaacaacca	tgcattcatt	gtaccgttta	ttcctatgag	720
tgcacttaca	gccaaccttc	cagtaagaag	agacagggac	aatctctgag	tctgagtgtc	780
ccgtcaaaca	ttaatgcaac	aagttccgta	caaaaatctg	taaaacctcc	tgaaatcgat	840
ttccaaagga	tgagagacgc	actcaaatat	tacgaagatc	ttttaaacca	gttgatatac	900
cccaacagtg	ctccaactgt	tcgagttaat	ccgattcgtc	tagcatcgat	cttaaaacaa	960
ttgagagccg	ataaatcaag	tgatgaatta	atttcagtca	aggctctttc	tgacaattac	1020
attgagatgc	ttcacaaaa	gatgcaacaa	cctgtacagc	agccagctcc	tccttcattg	1080
gggcaaggag	ggctctcttc	taatcacagt	ccaatcata	ataatgcttc	tattgatggg	1140
tcctataga	ctaacttagg	gagggaaata	cgtatcatat	tacctccgag	agatattgag	1200
ctgaagctta	tctacaagac	ttgggacaac	gcgtgtgtac	ttttccgctt	ttatcacaga	1260
cccgatttta	ttgaggacct	gaatgagtta	tatgaacag	atttgcaaaa	ctacaccaat	1320
aaacaacaaa	gggtttttacc	tcttgatat	tcggtgatgg	cttgtggtgc	tcttttttgc	1380
aagactgatg	ggattaatca	cggccaaaag	agctccaagc	ccaaagactc	ttctgatgaa	1440
agtcctcatg	acgatgaggg	ttacaagtat	tttattgccc	caagaaaact	aatagatata	1500

US 2011/0129874 A1

Jun. 2, 2011

33

-continued

acggatacca	gggataccta	cggaattcag	actattgtta	tgctgatcat	ttttttacaa	1560
tggtcggtctc	gtctttcaac	atgctattct	tatattggca	ttgctctaag	agctgcattg	1620
agagaaggtt	tgcatcgtea	gttgaactat	cctttcaatc	caattgagtt	agaacaaga	1680
aagcgtcttt	tttggactat	ctataaaatg	gacatctatg	tcaatacaat	gctggggcctt	1740
ccaagaacca	tttctgaaga	ggatttcgac	caggaaatgc	ctatcgaact	tgatgatgag	1800
aacattagt	aaaccggata	taggttcgat	ttacaaggta	caaagttatc	cagttcagga	1860
atagccaatg	ctcacactag	attgatattc	ataatgaaga	aaattgtgaa	aaaattatat	1920
cctgtcaaac	tacagaaacc	aacctcaaac	agtggcgata	cccacttga	gaacaatgat	1980
ttattggctc	atgaaatcgt	tcatgaactt	gagatggatc	tccaaaattg	ggccaatagt	2040
ctacctgcag	aactaaaacc	ggggatagaa	ccaccgaccg	agtattttaa	agctaacaga	2100
ttgcttcatt	tggcatacct	gcatgtcaag	attattctct	acaggccatt	tattcattac	2160
atctcagaaa	aggataaggt	tggaaatagt	tctatccctc	cgtcgcccg	agagatcact	2220
tctatcgaga	aagccaagaa	ttgtgtcaat	gttgccagaa	ttgttgtaa	actagccgaa	2280
gacatgatta	ataggaaaat	gttaagtggg	tcatattggg	ttccatttta	taccattttt	2340
ttttcgtgg	catgtctggt	gtactatgtt	catttcgctc	caccgaagaa	agacaatgga	2400
gaactggatc	ccaatacat	ggaaatcaag	aaagatacag	agagtggaag	agaggtctta	2460
aatatcctca	aagatagtag	tatggcggca	agaagaacgt	ataatattct	caactctttg	2520
tttgagcagt	taaacagaag	aactgcaaag	gtcaacctag	caaaggcaca	gcaaccacca	2580
tcagggttga	ataaccagc	tgctaccag	tatcagaaac	aggggtgaaca	caggcagtta	2640
caaccaagta	actattctgg	aactgtgaaa	tctgtggacc	cagagaatat	cgattactct	2700
tcctttggtt	ctcagtttga	aaacactaac	atcgaagatg	gttcctcaaa	tacaaagatt	2760
gatcagaaa	tgatggggg	gaactacatc	gatgggtgtg	ttacagggat	caacctaaat	2820
atgcctaate	tctcagaaac	ttctaacact	caaggtatcg	ataatccagc	atttcaaagt	2880
ataaacaatt	ctaatttgaa	caataatttt	gtacaaacaa	agtacattcc	cggcatgatg	2940
gaccagctag	atatgaaaat	tttcggaaga	ttccttcacc	cttacatgct	gaactccaac	3000
aagggtgaac	agggacaaaa	tgaaagggaac	ctatcaggcc	aaccatcctc	gtcgaataact	3060
cctgatggat	cacaacctgt	gacagttctg	gatggattat	accggttgca	gaatgataat	3120
aataataacc	acgaccaggg	aaattcaaag	tctgttgtaa	ataacagtaa	ctcggtagaa	3180
aacttactac	agaactttac	aatggtgccc	tcgggggttg	catcaacagt	gcaaaatcct	3240
gaagcggccc	aaaaattcaa	taatcatatg	tcaaacatat	cgaatatgaa	tgatccaaga	3300
agagctagcg	tagctacatc	agatggatcc	aatgacatgg	atcatcatag	ccaaggcccg	3360
ataaacaag	atttgaaacc	gttgagcaac	tacgagtttg	acgatctctt	ctttaatgat	3420
tggaccactg	cgccagatac	aataaatttt	gacagttaag	tcgaccatca	tcatcatcat	3480
cattgagttt	tagccttaga	catgactgtt	cctcagttca	agttgggcac	ttacgagaag	3540
accggtcttg	ctagattcta	atcaagagga	tgtoagaatg	ccatttgcct	gagagatgca	3600
ggcttcat	ttgatacttt	tttatttgta	acctatatag	tataggattt	ttttgtcat	3660
tttgtttctt	ctcgtacgag	cttgcctctg	atcagcctat	ctcgagctg	atgaatatct	3720
tgtggtaggg	gtttgggaaa	atcattcgag	tttgatgttt	ttcttggtat	ttccactcc	3780

US 2011/0129874 A1

Jun. 2, 2011

34

-continued

tcttcagagt acagaagatt aagtgagacc ttcgtttgtg cggtatcc 3827

<210> SEQ ID NO 45
<211> LENGTH: 4368
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Mxr CDS controlled by GAP

<400> SEQUENCE: 45

agatcttttt tgtagaaatg tcttggtgtc ctgtccaat caggtagcca tctctgaaat 60
atctggctcc gttgcaactc cgaacgacct gctggcaacg taaaattctc cggggtaaaa 120
cttaaatgtg gagtaatgga accagaaacg tctcttccct tctctctcct tccaccgccc 180
gttaccgtcc ctaggaaatt ttactctgct ggagagcttc ttctacggcc cccttgacgc 240
aatgtctctc ccagcattac gttgcgggta aaacggaggt cgtgtaccgc acctagcagc 300
ccagggatgg aaaagtcctg gccgtcgtg gcaataatag cggggcggac catgtcatga 360
gattattgga aaccaccaga atcgaatata aaaggcgaac acctttccca attttggttt 420
ctctgaccc aaagacttta aatttaattt atttgtccct atttcaatca attgaacaac 480
tatttcgaaa ccattgagcaa tctaccccca acttttggtt ccactagaca atctccagaa 540
gaccaatcac ctcccggtgc caaggagctg tcattcaatg ggaccacacc ctccaggaaag 600
ctacgcttat ttgtctgtca gacatgtact cgagcatttg ctgtcagga acacttgaaa 660
cgacacgaaa ggtctcacac caaggagaaa cctttcagct gcggcatttg ttctcgtaaa 720
ttcagccgtc gagatctgtt attgagacat gcccacaaac tgcacagcaa ctgctctgat 780
gcggccataa caagactaag gcgcaaggca actcgtcggg ctctcaatgc cgggggttcc 840
atatctggtt ctactcggg gacaacgcca aatactatgg gtacgccga agatggcgag 900
aaacgaaaag ttcagaaact ggccggccgc cgggaactca atgaacagaa actgcaactg 960
caacaacaac atctacagca acaaccacag ttgcaatacc aacaatctct taagcagcat 1020
gaaaatcaag tccagcagcc tgatcaagat ccattgatat ccccgagaat gcaattatcc 1080
aatgattcca accatcacgt aaacaatttg ttgatcttg gactaagaag agcttccttc 1140
tcgcgcgtta gtggaaaata ttatgcccat tatgtgaata attttcaaca agatgcctct 1200
tctaccaatc caaatcaaga ttcaataat gccgaatttg agaataattga attttctacc 1260
ccacaaatga tgcccggtga agatgtgaa acttgatga acaacatggg tccaattccg 1320
aacttctctc tcgatgtgaa caggacattt ggtgatagct ttacagatat acaacacaag 1380
aattcagagc ctattatata cgaaccgccc aaggacaccg ctccaaacga caagaagttg 1440
aatggctact ctttttacga agccccatc aagccattag aatccctatt ttctgtcagg 1500
aatacaaaag gaaacaagta taaaacaaat gacgactctc cagacaccgt ggataataac 1560
tcgcaccggg ctgctaatac cattcaagaa cttgagtctt ctttgaatgc atccaagaat 1620
ttttgcttgc caactgggta ttctctctat ggtaatttgg accaacagac ttctctaac 1680
acgttatcat gcacttcttc taatgccaca atttcgcca ttctactcga taactccatt 1740
aataataact ccactagtga cgtgagacca gaatttagaa cacaaagtgt cacctctgaa 1800
atgagtcag cccctcccc tctcaaaaa aacaactcga aatattccac cgaagtctct 1860
tttaccagca acatgcgggc gtttattcac tacgtctttt ccaagtatcc ttttattggg 1920

US 2011/0129874 A1

Jun. 2, 2011

35

-continued

gtgcccactc	caactcttcc	ggagaacgaa	agactaaatg	aatatgctga	ttcattcacc	1980
aaccgtttct	taaatcatta	tcctttcata	catgtcacga	ttctcaaaga	atactccctt	2040
ttcaaggcaa	ttttagatga	gaatgagtcg	actaagaact	gggaaaaata	tcagttttac	2100
ttagagaacc	aacgaatata	aattgtttgt	cttcctcttt	tggtggctac	gatagtgca	2160
gtactatcaa	acaacaaaaa	ggatgcttcg	aatttatacg	aagcttcaag	gcgttgattt	2220
catgtttact	tagattccag	gaaaaagata	cccacttcct	tgtccgcaaa	taacaatgac	2280
tctccacttt	ggctaattca	atccctgacg	ttatctgtta	tgtatgggtt	atttgccgac	2340
aatgacatta	gtttgaatgt	cgtgatcaga	caagttaacg	cacttaattc	tctgggtcaag	2400
acttcgggcc	tgaataggac	ctcaattata	gatcttttca	acatcaacaa	acctttggat	2460
aatgaactct	ggaatcaatt	cgtgaaaata	gagtcaccgc	taaggacaat	ccacacgatt	2520
tttcaaatca	gttccaactt	aagcgcttg	tacaatatta	ttccatcggt	gaaaattgat	2580
gacctaatga	ttactctacc	agttcccaca	acactttggc	aagctgattc	ttttgtgaaa	2640
ttcaaaaagtc	taagttacgg	aaatcagatc	ccttttcaat	atacaagagt	actacagaat	2700
ttgattgatt	acaatcagcc	attgagcgat	ggaaaatttt	tgtatgaaaa	ccatgtaagt	2760
gagtttggac	tcatatgcct	acagaatggt	ctacaccaat	acagctattt	ccaaaaattg	2820
actgctgtca	ataacagaga	agatgcgcta	ttcacaaagg	ttgttaattc	acttcacagt	2880
tgggatagga	tgatttcgaa	ttctgatttg	tttccaaaga	agatatatca	gcagagttgc	2940
ttgattttgg	actcaaagtt	gcttaataat	ttctgatttg	tcaagagctc	attgaaagtt	3000
tcgaccggag	acgttagttc	tttgaataag	ttaaaagaaa	acgtgtggct	taaaaactgg	3060
aatcaagtgt	gtgctatcta	ttataacagc	ttcatgaaca	ttcctgctcc	cagtattcaa	3120
aagaagtaca	atgacataga	gtttgtggat	gacatgatta	atttgagtct	aatcatcatc	3180
aagattatga	aactcatttt	ctataacaat	gtcaaagaca	attatgagga	tgaaaatgac	3240
ttcaaattgc	aagagttaaa	tttaacattt	gacaattttg	atgagaaaat	atccttgaat	3300
ttgacaatat	tattcgatat	atttttgatg	atctacaaga	taattaccaa	ttacgagaag	3360
tttatgaaga	tcaaacacaa	gtttaattac	tacaattcta	attcgaatat	aagcttcttg	3420
catcattttc	aactctctct	ggttatcaat	aacacccaaa	tgaaccagaa	tgattatatg	3480
aaaacagata	ttgatgaaaa	gcttgatcag	cttttccaca	tctatcaaac	atttttccgg	3540
ctgtatctgg	atttagaaaa	gtttatgaag	ttcaaattca	actatcatga	ctttgagaca	3600
gagttttcaa	gtctctcaat	atccaatata	ctgaacactc	atgctgcttc	taacaatgac	3660
acaaatgctg	ctgatgctat	gaatgccaa	gatgaaaaaa	tatctccac	aactttgaat	3720
agcgtattac	ttgctgatga	aggaaatgaa	aattccggtc	gtaataacga	ttcagaccgc	3780
ctgttcacgc	tgaacgagct	aattaatttt	gaagtagggt	tgaaatttct	caagataggt	3840
gagtcatttt	ttgatttctt	gtatgagaat	aactacaagt	tcattccactt	caaaaactta	3900
aatgacggaa	tgttccacat	caggatatac	ctagaaaaac	gactagatgg	tggtgtctct	3960
agaacaaaaa	ctcatctcag	aagaggatct	gaatagcgcc	gtcgaccatc	atcatcatca	4020
tcattgagtt	ttagccttag	acatgactgt	tcctcagttc	aagttgggca	cttacgagaa	4080
gaccggtctt	gctagattct	aatcaagagg	atgtcagaat	gccatttgcc	tgagagatgc	4140
aggcttcatt	tttgatactt	ttttatttgt	aacctatata	gtataggatt	ttttttgtca	4200

US 2011/0129874 A1

Jun. 2, 2011

36

-continued

ttttgtttct tctcgtaaga gcttgctect gatcagccta tctcgagct gatgaatgc	4260
ttgtggtagg ggtttgggaa aatcattcga gtttgatgtt tttcttgga tttcccactc	4320
ctcttcagag tacagaagat taagtgcagac ctctgtttgt gcggatcc	4368

1. An isolated polynucleotide comprising:
 - i) a nucleotide sequence consisting of the DAS promoter sequence from *Pichia* or a functional part thereof, wherein the said DAS promoter is comprised in SEQ ID NO: 1; and
 - ii) at least one additional UAS, wherein the said UAS is comprised in SEQ ID NO: 2.
2. The isolated polynucleotide according to claim 1, wherein the UAS comprises at least the sequence shown in SEQ ID NO: 3.
3. The isolated polynucleotide according to claim 1, wherein the DAS promoter is comprised in position 501 to 1055 in SEQ ID NO: 1.
4. The isolated polynucleotide according to claim 1, wherein the DAS promoter is comprised in position 601 to 1055 in SEQ ID NO: 1.
5. The isolated polynucleotide according to claim 1, comprising at least two additional UASs.
6. The isolated polynucleotide according to claim 1, comprising at least three additional UASs.
7. The isolated polynucleotide according to claim 1, wherein the promoter is chosen from the group consisting of:
 - a) a polynucleotide comprising or consisting of position 77 to 828, particularly 77 to 901, more particularly 77 to 948 of SEQ ID NO: 7;
 - b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with position 77 to 828, particularly 77 to 901, more particularly 77 to 948 of SEQ ID NO: 7;
 - c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with position 77 to 828, particularly 77 to 901, more particularly 77 to 948 of SEQ ID NO: 7 or a full-length complementary strand thereof.
8. The isolated polynucleotide according to claim 1, wherein the promoter is chosen from the group consisting of:
 - a) a polynucleotide comprising or consisting of position 60 to 920, particularly 60 to 993, more particularly 60 to 1040 of SEQ ID NO: 8;
 - b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with position 60 to 920, particularly 60 to 993, more particularly 60 to 1040 of SEQ ID NO: 8;
 - c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with position 60 to 920, particularly 60 to 993, more particularly 60 to 1040 of SEQ ID NO: 8 or a full-length complementary strand thereof.
9. The isolated polynucleotide according to claim 1, wherein the promoter is chosen from the group consisting of:
 - a) a polynucleotide comprising or consisting of position 48 to 1015, particularly 48 to 1088, more particularly 48 to 1135 of SEQ ID NO: 9;
 - b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with position 48 to 1015, particularly 48 to 1088, more particularly 48 to 1135 of SEQ ID NO: 9;
 - c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with position 48 to 1015, particularly 48 to 1088, more particularly 48 to 1135 of SEQ ID NO: 9 or a full-length complementary strand thereof.
10. A DNA construct comprising a polynucleotide sequence of claim 1 operably linked to a structural gene encoding a polypeptide of interest and a terminator.
11. An expression vector comprising a DNA construct of claim 10, further comprising a signal peptide coding region.
12. A *Pichia* host cell comprising an expression vector of claim 10.
13. A method of producing a polypeptide of interest comprising:
 - (a) cultivating the host cell of claim 12, under conditions conducive for the production of the polypeptide of interest; and
 - (b) recovering the polypeptide.
14. The method according to claim 13, wherein expression of the positive regulator Prm1 is increased by controlling the expression of Prm1 or by increasing the copy number of the gene encoding Prm1.
15. The method according to claim 13, wherein expression of the positive regulator Mxr1 is increased by controlling the expression of Mxr1 or by increasing the copy number of the gene encoding Mxr1.
16. The method according to claim 13, wherein expression of the positive regulators Prm1 and Mxr1 are increased at the same by controlling their expression or by increasing the copy number of the respective genes.
17. A promoter comprising an UAS selected from the group consisting of:
 - i) (a) a polynucleotide comprising or consisting of SEQ ID NO: 2; or
 - (b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 2; or
 - c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 2 or a full-length complementary strand thereof; or

US 2011/0129874 A1

Jun. 2, 2011

37

ii) (a) a polynucleotide comprising or consisting of SEQ ID NO: 3; or

(b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 3; or

(c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 3 or a full-length complementary strand thereof; and

wherein the UAS according to (i) or (ii) is either foreign to the promoter or present in more than one copy.

18. A use of an UAS for increasing transcription from a promoter, wherein the UAS is selected from the group consisting of:

i) (a) a polynucleotide comprising or consisting of SEQ ID NO: 2; or

(b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 2; or

(c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 2 or a full-length complementary strand thereof; or

ii) (a) a polynucleotide comprising or consisting of SEQ ID NO: 3; or

(b) a polynucleotide comprising or consisting of a polynucleotide having at least 90% identity, preferably at least 95%, more preferably at least 97%, even more preferably at least 99% identity with SEQ ID NO: 3; or

(c) a polynucleotide comprising or consisting of polynucleotide that hybridizes under at least high stringency conditions with SEQ ID NO: 3 or a full-length complementary strand thereof; and

wherein the UAS according to (i) or (ii) is either foreign to the promoter or present in more than one copy.

19. The use according to claim **18**, wherein the promoter is induced by methanol.

* * * * *