



US 20230357705A1

(19) **United States**(12) **Patent Application Publication**
MACH et al.(10) **Pub. No.: US 2023/0357705 A1**(43) **Pub. Date: Nov. 9, 2023**(54) **ERYTHRITOL PRODUCING SAPROTROPH****Publication Classification**(71) Applicant: **CONZIL ESTATE GMBH**,
Bodman-Ludwigshafen (DE)(72) Inventors: **Robert MACH**, Vienna (AT); **Astrid MACH-AIGNER**, Vienna (AT);
Katharina REGNAT, Vienna (AT)(51) **Int. Cl.****C12N 1/14** (2006.01)**C12P 7/18** (2006.01)**C12N 9/10** (2006.01)**C12N 9/04** (2006.01)(52) **U.S. Cl.**CPC **C12N 1/145** (2021.05); **C12P 7/18**(2013.01); **C12N 9/1022** (2013.01); **C12Y****202/01001** (2013.01); **C12Y 202/01002**(2013.01); **C12N 9/0006** (2013.01); **C12Y****101/01017** (2013.01)(21) Appl. No.: **18/012,440**(22) PCT Filed: **Jun. 25, 2021**(86) PCT No.: **PCT/EP2021/067473**

§ 371 (c)(1),

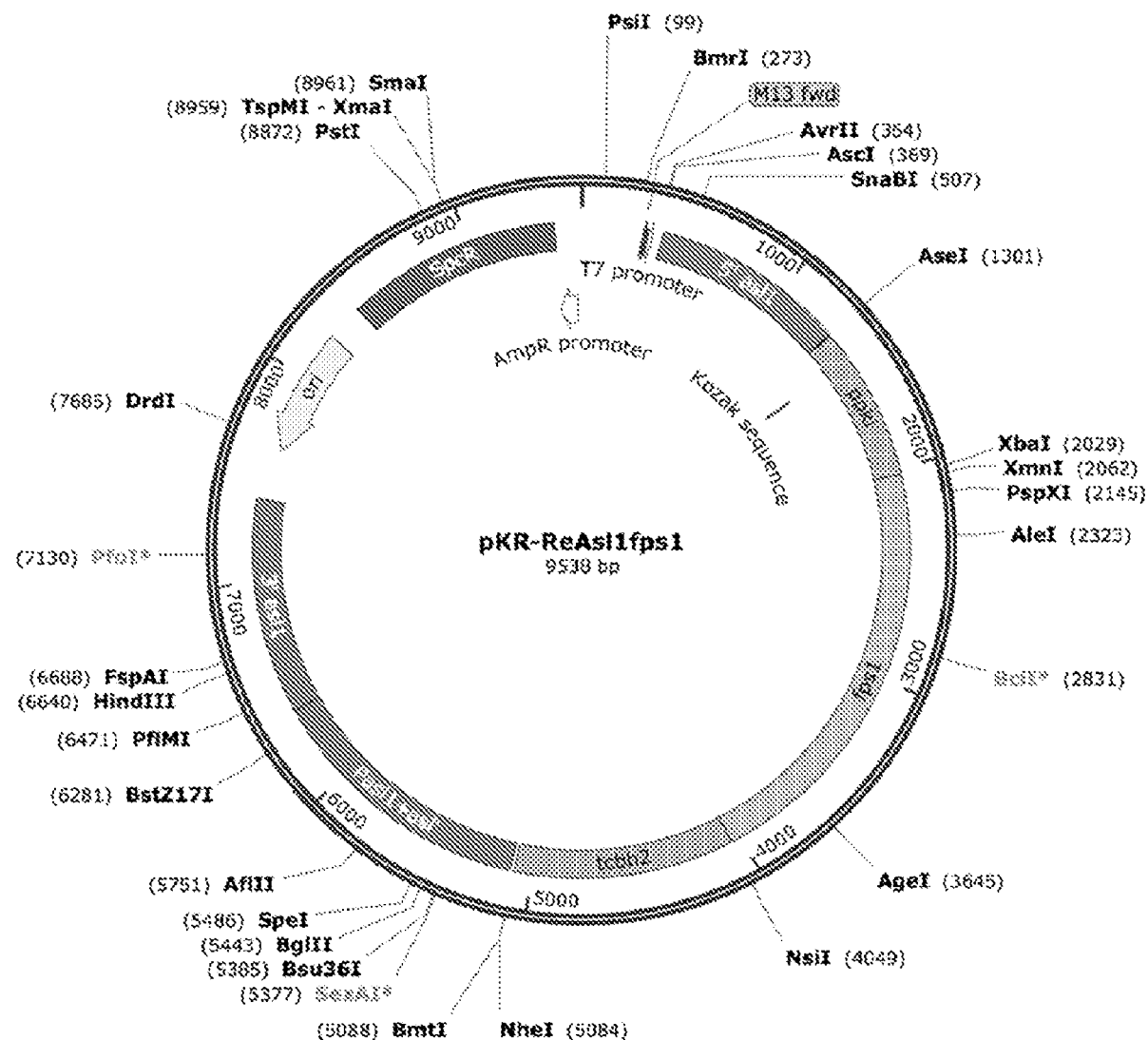
(2) Date: **Dec. 22, 2022**(30) **Foreign Application Priority Data**

Jun. 25, 2020 (EP) 20182382.0

(57)

ABSTRACT

The present invention pertains to a genetically modified saprotroph for the biotechnological production of erythritol and a method for the production of erythritol using said genetically modified saprotroph.

Specification includes a Sequence Listing.

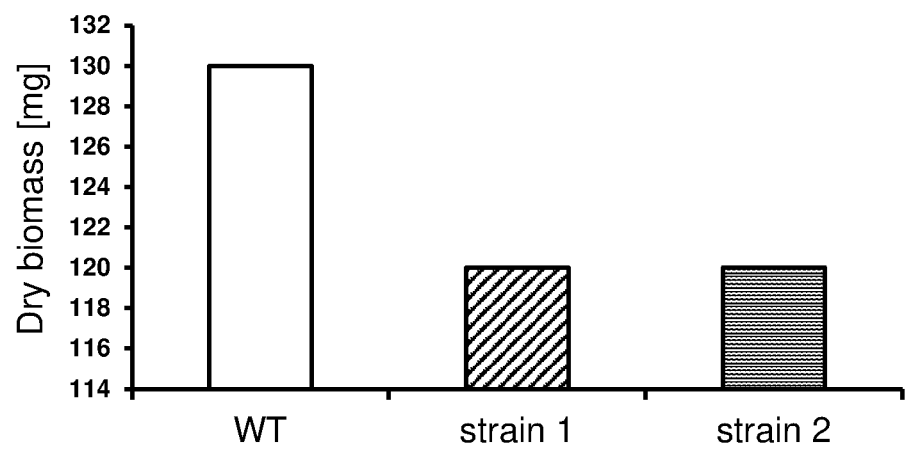


Fig. 1

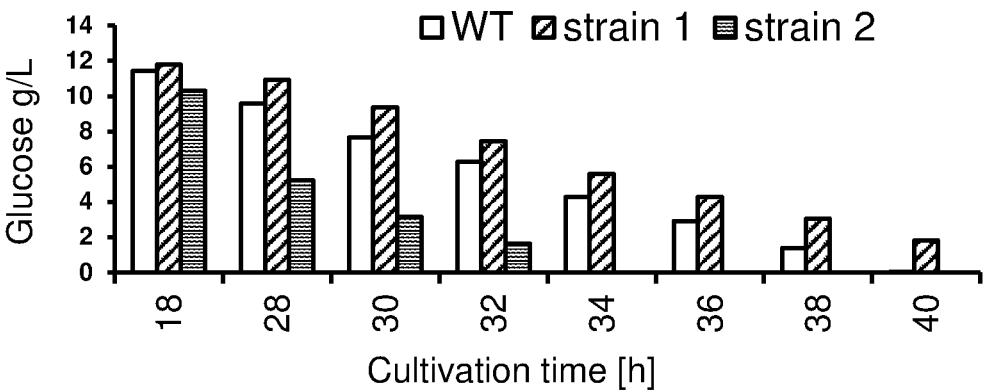


Fig. 2

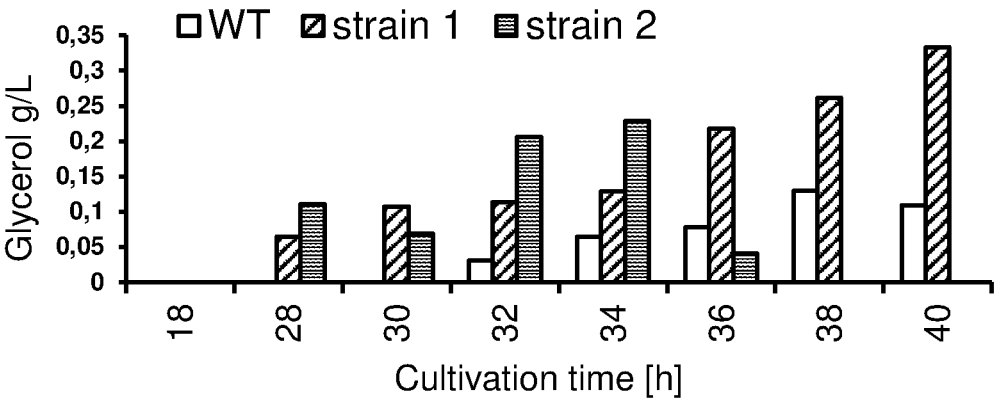


Fig. 3

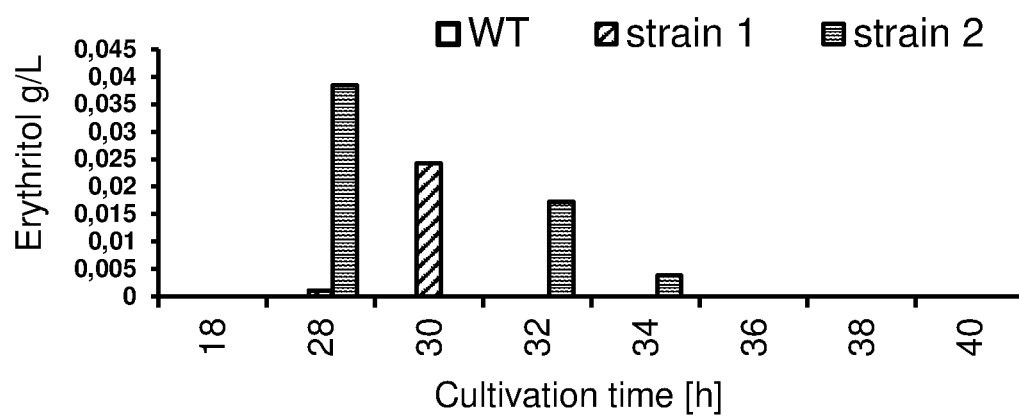


Fig. 4

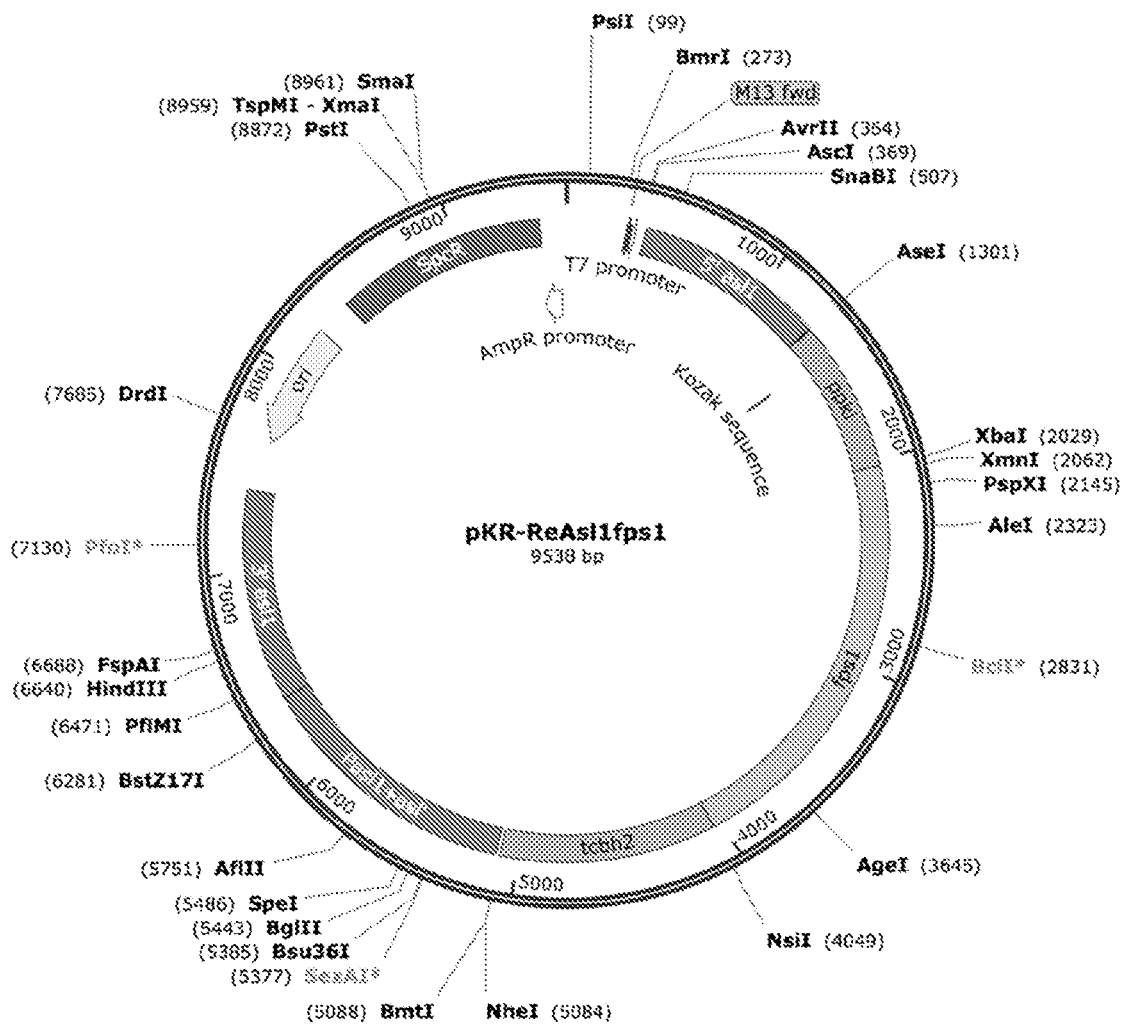


Fig. 5

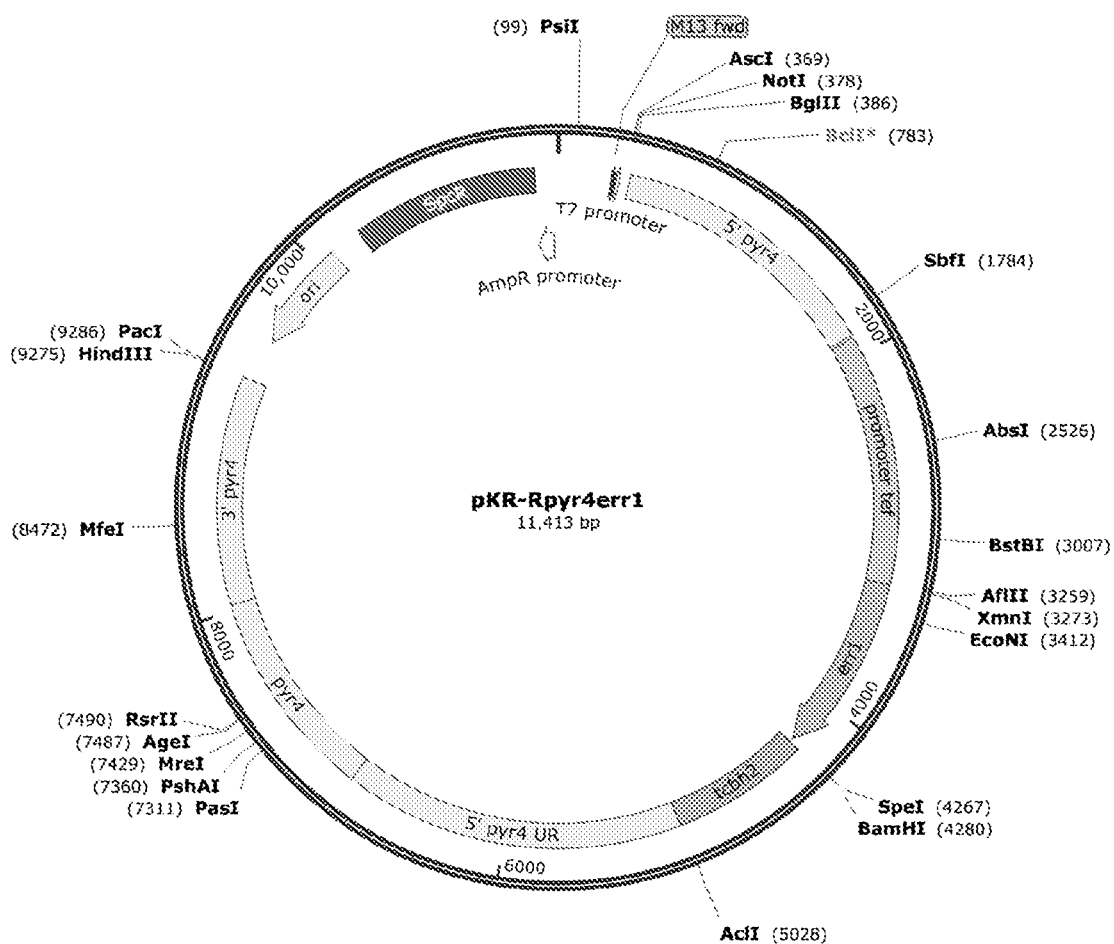


Fig. 6

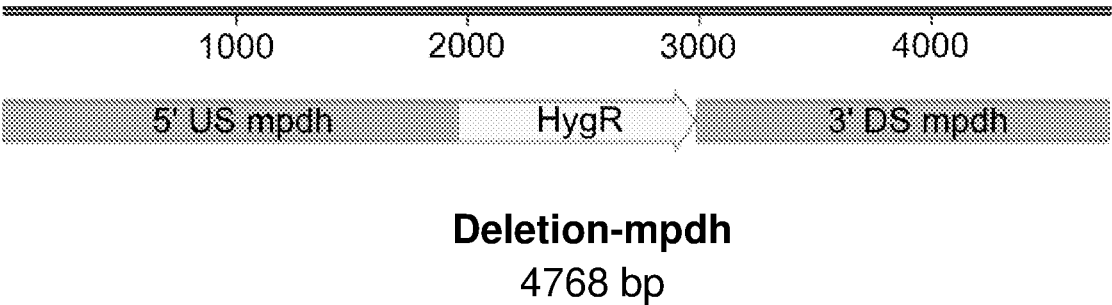


Fig. 7

ERYTHRITOL PRODUCING SAPROTROPH

[0001] The present invention pertains to a genetically modified saprotroph for the biotechnological production of erythritol and a method for the production of erythritol using said genetically modified saprotroph.

[0002] In the recent years people's lifestyle and the growing consumption of food products with high sugar content has resulted in a tremendous rise of blood glucose related diseases and disorders, such as diabetes mellitus type 2 (DMT2). Nowadays, a low-glycaemic nutrition and the avoidance of excessive peaks in blood glucose level is considered to reduce the risk for developing certain chronic diseases and to be beneficial for maintenance and improvement of health and for the treatment and/or prevention of a large number of blood glucose related diseases and disorders.

[0003] Erythritol is a naturally occurring four-carbon sugar alcohol gaining increasing importance in the food industry due to its specific properties and its manifold fields of application. It can be found in several fruits, such as pears, grapes and melons, mushrooms, alcoholic drinks (beer, wine, sake) and fermented food products, such as soy sauce and miso bean paste, but naturally also occurs in biofluids of humans and animals such as eye lens tissue, serum, plasma, fetal fluid and urine. Due to its small molecular weight, erythritol is easily absorbed already in the upper intestine and therefore causes less digestive distress than other sugar alcohols used in the food industry. The majority of ingested erythritol is not metabolized in the human body and is excreted unmodified into the urine without changing blood glucose and insulin levels (Regnat et al., Erythritol as sweetener wherefrom and whereto?, 2018, Applied Microbiology and Biotechnology, Vol. 102). Furthermore, erythritol is non-cariogenic, thermally stable, crystalizes well and is less hygroscopic than sucrose. Due to the negative enthalpy of dissolution, the consumption of erythritol causes a cooling sensation in the oral cavity. A 10% (w/v) solution of erythritol has 60-80% of the sweetness of sucrose.

[0004] However, in contrast to other sugar alcohols, such as sorbitol, xylitol, mannitol, lactitol, and maltitol, which are well-established as sugar alternatives for many years, so far erythritol cannot be chemically produced in a commercially worthwhile way. The production of erythritol from dialdehyde starch using a nickel catalyst at high temperatures results in unsatisfying low yields (Moon et al. Biotechnological production of erythritol and its applications, 2010, Appl. Microbiol. Biotechnol., Vol. 86).

[0005] In yeast and fungus species, erythritol is produced via the so-called pentose phosphate pathway. It is synthesized from D-erythrose-4-phosphate through dephosphorylation and subsequent reduction of erythrose. Based thereon, the suitability of osmophilic yeast, such as *Aureobasidium* sp., *Trichosporonoides* sp. and *Candida magnoliae*, for the biotechnological production of erythritol has been investigated in several studies (Ishizuka et al., Breeding of a mutant of *Aureobasidium* sp. with high erythritol production, 1989, J. Ferm. Bioeng., Vol. 68(5); U.S. Pat. Nos. 4,939,091A; 5,962,287 A; Oh et al., Increased erythritol production in fedbatch cultures of *Torula* sp. by controlling glucose concentration, 2001, JIM&B, Vol. 26; Koh et al., Scale-up of erythritol production by an osmophilic mutant of *Candida magnoliae*, 2003, Biotechnol. Lett., Vol. 25; Ryu et al., Optimization of erythritol production by *Candida magnoliae* in fed-batch culture, 2000, JIM&B, Vol. 25). More

recent studies examined the potential of filamentous fungi to produce erythritol (Jovanovic et al., 2013). However, the yields of erythritol obtained from the different strains was unsatisfactory for industrial scale production.

[0006] Accordingly, there is still a need for a biotechnological method for the production of erythritol with increased yield. The present invention overcomes the disadvantages of the methods in the prior art by the subject-matter of the independent claims, in particular by the genetically modified saprotroph according to the present invention and the method for the production of erythritol according to the present invention.

[0007] The present invention in particular pertains to a genetically modified saprotroph comprising at least one gene encoding at least one membrane-bound alditol transporter, at least one gene encoding at least one erythrose reductase and at least one inactivated gene encoding mannitol 1-phosphate 5-dehydrogenase.

[0008] In a preferred embodiment of the present invention, the genetically modified saprotroph is a filamentous fungus. Preferably, the saprotroph is a filamentous fungus selected from the genera *Hypocrea*, *Gibberella*, *Aspergillus* and *Penicillium*. Preferably, the saprotroph is a filamentous fungus selected from *Hypocrea jecorina* (*Trichoderma reesei*), *Gibberella zeae*, *Aspergillus niger*, *Aspergillus nidulans*, *Aspergillus oryzae* and *Penicillium chrysogenum*.

[0009] In a particularly preferred embodiment of the present invention, the genetically modified saprotroph is *Hypocrea jecorina* (*Trichoderma reesei*). Preferably, the genetically modified saprotroph is based on the *Trichoderma reesei* strain QM6aΔtmus53.

[0010] In a further preferred embodiment of the present invention, the at least one gene encoding at least one membrane-bound alditol transporter is *fps1*, in particular codon-optimized *fps1*. Preferably, the at least one gene encoding at least one membrane-bound alditol transporter is *fps1* from *Saccharomyces cerevisiae*, in particular codon-optimized *fps1* from *Saccharomyces cerevisiae*.

[0011] In a further preferred embodiment, the at least one gene encoding at least one membrane-bound alditol transporter comprises the nucleotide sequence of SEQ ID No. 1, in particular consists of the nucleotide sequence of SEQ ID No. 1.

[0012] In a preferred embodiment of the present invention, the at least one gene encoding at least one membrane-bound alditol transporter comprises a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 1. Preferably, the at least one gene encoding at least one membrane-bound alditol transporter consists of a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 1.

[0013] Preferably, the membrane-bound alditol transporter comprises the amino acid sequence of SEQ ID No. 2, in particular consists of the amino acid sequence of SEQ ID No. 2.

[0014] Preferably, the membrane-bound alditol transporter comprises an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least

97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 2. In a further preferred embodiment, the membrane-bound alditol transporter consists of an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 2.

[0015] In another preferred embodiment of the present invention, the membrane-bound alditol transporter comprises an amino acid sequence as defined in SEQ ID No. 2, in which one, preferably two, preferably three, preferably five, preferably six, preferably seven, preferably eight, preferably nine, preferably ten, amino acids are exchanged by other naturally occurring amino acids. In a particularly preferred embodiment, the membrane-bound alditol transporter comprises an amino acid sequence as defined in SEQ ID No. 2, in which at most ten, preferably at most nine, preferably at most eight, preferably at most seven, preferably at most six, preferably at most five, preferably at most four, preferably at most three, preferably at most two, preferably at most one, amino acids are exchanged by other naturally occurring amino acids.

[0016] In a preferred embodiment of the present invention, the at least one gene encoding at least one erythrose reductase is *err1*, in particular codon-optimized *err1*. Preferably, the at least one gene encoding at least one erythrose reductase is *err1*, in particular codon-optimized *err1*, from *Trichoderma reesei*, *Aspergillus niger* or *Fusarium graminearum*.

[0017] In a preferred embodiment, the at least one gene encoding at least one erythrose reductase comprises the nucleotide sequence of SEQ ID No. 3, in particular consist of the nucleotide sequence of SEQ ID No. 3.

[0018] Preferably, the at least one gene encoding at least one erythrose reductase comprises a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 3. In another preferred embodiment of the present invention, the at least one gene encoding at least one gene encoding at least one erythrose reductase consists of a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 3.

[0019] Preferably, the erythrose reductase comprises the amino acid sequence of SEQ ID No. 4, in particular consists of the amino acid sequence of SEQ ID No. 4.

[0020] In a preferred embodiment, the erythrose reductase comprises an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 4. Particularly preferred, the erythrose reductase consists of an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 4.

[0021] Preferably, the erythrose reductase comprises an amino acid sequence as defined in SEQ ID No. 4, in which one, preferably two, preferably three, preferably five, pref-

erably six, preferably seven, preferably eight, preferably nine, preferably ten, amino acids are exchanged by other naturally occurring amino acids. In a preferred embodiment of the present invention, the erythrose reductase comprises an amino acid sequence as defined in SEQ ID No. 4, in which at most ten, preferably at most nine, preferably at most eight, preferably at most seven, preferably at most six, preferably at most five, preferably at most four, preferably at most three, preferably at most two, preferably at most one, amino acids are exchanged by other naturally occurring amino acids.

[0022] In a preferred embodiment of the present invention, the at least one inactivated gene encoding mannitol 1-phosphate 5-dehydrogenase is *mpdh*.

[0023] Preferably, the at least one gene encoding mannitol 1-phosphate 5-dehydrogenase of the genetically modified saprotroph comprised the nucleotide sequence of SEQ ID No. 12 before inactivation, in particular consisted of the nucleotide sequence of SEQ ID No. 12 before inactivation.

[0024] Preferably, the at least one gene encoding mannitol 1-phosphate 5-dehydrogenase of the genetically modified saprotroph comprised a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 12 before inactivation. In another preferred embodiment, the at least one gene encoding mannitol 1-phosphate 5-dehydrogenase of the genetically modified saprotroph consisted of a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 12 before inactivation.

[0025] Preferably, the mannitol 1-phosphate 5-dehydrogenase encoded by the nucleotide sequence of SEQ ID No. 12 before inactivation comprises the amino acid sequence of SEQ ID No. 13, in particular consists of the amino acid sequence of SEQ ID No. 13.

[0026] In a preferred embodiment, the mannitol 1-phosphate 5-dehydrogenase encoded by the nucleotide sequence of SEQ ID No. 12 before inactivation comprises an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 13. Particularly preferred, the mannitol 1-phosphate 5-dehydrogenase encoded by the nucleotide sequence of SEQ ID No. 12 before inactivation consists of an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 13.

[0027] Preferably, the mannitol 1-phosphate 5-dehydrogenase encoded by the nucleotide sequence of SEQ ID No. 12 before inactivation comprises an amino acid sequence as defined in SEQ ID No. 13, in which one, preferably two, preferably three, preferably five, preferably six, preferably seven, preferably eight, preferably nine, preferably ten, amino acids are exchanged by other naturally occurring amino acids. In a preferred embodiment of the present invention, the mannitol 1-phosphate 5-dehydrogenase encoded by the nucleotide sequence of SEQ ID No. 12

before inactivation comprises an amino acid sequence as defined in SEQ ID No. 13, in which at most ten, preferably at most nine, preferably at most eight, preferably at most seven, preferably at most six, preferably at most five, preferably at most four, preferably at most three, preferably at most two, preferably at most one, amino acids are exchanged by other naturally occurring amino acids.

[0028] In a further preferred embodiment of the present invention, the at least one inactivated gene encoding mannitol 1-phosphate 5-dehydrogenase of the genetically modified saprotroph is deleted.

[0029] In a preferred embodiment of the present invention, the at least one inactivated gene encoding mannitol 1-phosphate 5-dehydrogenase of the genetically modified saprotroph is non-functional.

[0030] In a particularly preferred embodiment, the at least one inactivated gene encoding mannitol 1-phosphate 5-dehydrogenase of the genetically modified saprotroph is inactivated by gene knock-out, in particular gene replacement. Preferably, the at least one inactivated gene encoding mannitol 1-phosphate 5-dehydrogenase of the genetically modified saprotroph is inactivated by gene replacement using a deletion cassette.

[0031] In a preferred embodiment of the present invention, the at least one gene encoding mannitol 1-phosphate 5-dehydrogenase of the genetically modified saprotroph is inactivated, in particular made non-functional, by genome editing, in particular by meganucleases, zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) or by the clustered regularly interspaced short palindromic repeats (CRISPR) system.

[0032] In a particularly preferred embodiment, the genetically modified saprotroph comprises at least one, preferably at least two, preferably at least three, preferably at least four, preferably at least five, further inactivated genes. Preferably, the at least one further inactivated gene is selected from the group consisting of a gene encoding phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhps-1), a gene encoding erythritol utilization factor (EUF), a gene encoding erythrulose kinase (EYK1), a gene encoding erythritol dehydrogenase (EYD1), a gene encoding erythritol isomerase 1 (EYI1) and a gene encoding erythritol isomerase 2 (EYI2).

[0033] In a preferred embodiment of the present invention, the at least one further inactivated gene is a gene encoding phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhps-1). Preferably, the at least one gene encoding phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhps1) of the genetically modified saprotroph comprised the nucleotide sequence of SEQ ID No. 14 before inactivation, in particular consisted of the nucleotide sequence of SEQ ID No. 14 before inactivation.

[0034] Preferably, the at least one gene encoding phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhps1) of the genetically modified saprotroph comprised a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 14 before inactivation. In further preferred embodiment, the at least one gene encoding phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhps1) of the genetically modified saprotroph consisted of a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably

at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 14 before inactivation.

[0035] Preferably, the phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhps1) encoded by the nucleotide sequence of SEQ ID No. 14 before inactivation comprises the amino acid sequence of SEQ ID No. 15, in particular consists of the amino acid sequence of SEQ ID No. 15.

[0036] In a preferred embodiment, the phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhps1) encoded by the nucleotide sequence of SEQ ID No. 14 before inactivation comprises an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 15. Particularly preferred, the phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhps1) encoded by the nucleotide sequence of SEQ ID No. 14 before inactivation consists of an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 15.

[0037] Preferably, the phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhps1) encoded by the nucleotide sequence of SEQ ID No. 14 before inactivation comprises an amino acid sequence as defined in SEQ ID No. 15, in which one, preferably two, preferably three, preferably five, preferably six, preferably seven, preferably eight, preferably nine, preferably ten, amino acids are exchanged by other naturally occurring amino acids. In a preferred embodiment of the present invention, the mannitol 1-phosphate 5-dehydrogenase encoded by the nucleotide sequence of SEQ ID No. 12 before inactivation comprises an amino acid sequence as defined in SEQ ID No. 13, in which at most ten, preferably at most nine, preferably at most eight, preferably at most seven, preferably at most six, preferably at most five, preferably at most four, preferably at most three, preferably at most two, preferably at most one, amino acids are exchanged by other naturally occurring amino acids.

[0038] In another embodiment of the present invention, the at least one further inactivated gene is a gene encoding erythritol utilization factor (EUF). Preferably, the at least one gene encoding erythritol utilization factor (Euf1) of the genetically modified saprotroph comprised the nucleotide sequence of SEQ ID No. 16 before inactivation, in particular consisted of the nucleotide sequence of SEQ ID No. 16 before inactivation.

[0039] Preferably, the at least one gene encoding erythritol utilization factor (Euf1) of the genetically modified saprotroph comprised a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 16 before inactivation. Particularly preferred, the at least one gene encoding erythritol utilization factor (Euf1) of the genetically modified saprotroph consisted of a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 16 before inactivation.

[0040] Preferably, the erythritol utilization factor (Euf1) encoded by the nucleotide sequence of SEQ ID No. 16 before inactivation comprises the amino acid sequence of SEQ ID No. 17, in particular consists of the amino acid sequence of SEQ ID No. 17.

[0041] In a preferred embodiment, the erythritol utilization factor (Euf1) encoded by the nucleotide sequence of SEQ ID No. 16 before inactivation comprises an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 17. Particularly preferred, the erythritol utilization factor (Euf1) encoded by the nucleotide sequence of SEQ ID No. 16 before inactivation consists of an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 17.

[0042] Preferably, the erythritol utilization factor (Euf1) encoded by the nucleotide sequence of SEQ ID No. 16 before inactivation comprises an amino acid sequence as defined in SEQ ID No. 17, in which one, preferably two, preferably three, preferably five, preferably six, preferably seven, preferably eight, preferably nine, preferably ten, amino acids are exchanged by other naturally occurring amino acids. In a preferred embodiment of the present invention, the erythritol utilization factor (Euf1) encoded by the nucleotide sequence of SEQ ID No. 16 before inactivation comprises an amino acid sequence as defined in SEQ ID No. 17, in which at most ten, preferably at most nine, preferably at most eight, preferably at most seven, preferably at most six, preferably at most five, preferably at most four, preferably at most three, preferably at most two, preferably at most one, amino acids are exchanged by other naturally occurring amino acids.

[0043] In a particularly preferred embodiment of the present invention, the at least one further inactivated gene is a gene encoding erythrulose kinase (EYK1). Preferably, the at least one gene encoding erythrulose kinase (Eyk1) of the genetically modified saprotroph comprised the nucleotide sequence of SEQ ID No. 18 before inactivation, in particular consisted of the nucleotide sequence of SEQ ID No. 18 before inactivation.

[0044] According to another preferred embodiment of the present invention, the at least one gene encoding erythrulose kinase (Eyk1) of the genetically modified saprotroph comprised a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 18 before inactivation. Particularly preferred, the at least one gene encoding erythrulose kinase (Eyk1) of the genetically modified saprotroph consisted of a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 99%, sequence identity to SEQ ID No. 18 before inactivation.

[0045] Preferably, the erythrulose kinase (Eyk1) encoded by the nucleotide sequence of SEQ ID No. 18 before inactivation comprises the amino acid sequence of SEQ ID No. 19, in particular consists of the amino acid sequence of SEQ ID No. 19.

[0046] In a preferred embodiment, the erythrulose kinase (Eyk1) encoded by the nucleotide sequence of SEQ ID No. 18 before inactivation comprises an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 19. Particularly preferred, the erythrulose kinase (Eyk1) encoded by the nucleotide sequence of SEQ ID No. 18 before inactivation consists of an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 19.

[0047] Preferably, the erythrulose kinase (Eyk1) encoded by the nucleotide sequence of SEQ ID No. 18 before inactivation comprises an amino acid sequence as defined in SEQ ID No. 19, in which one, preferably two, preferably three, preferably five, preferably six, preferably seven, preferably eight, preferably nine, preferably ten, amino acids are exchanged by other naturally occurring amino acids. In a preferred embodiment of the present invention, the erythrulose kinase (Eyk1) encoded by the nucleotide sequence of SEQ ID No. 18 before inactivation comprises an amino acid sequence as defined in SEQ ID No. 19, in which at most ten, preferably at most nine, preferably at most eight, preferably at most seven, preferably at most six, preferably at most five, preferably at most four, preferably at most three, preferably at most two, preferably at most one, amino acids are exchanged by other naturally occurring amino acids.

[0048] In a preferred embodiment of the present invention, the at least one further inactivated gene is a gene encoding erythritol dehydrogenase (EYD1). Preferably, the at least one gene encoding erythritol dehydrogenase (Eyd1) of the genetically modified saprotroph comprised the nucleotide sequence of SEQ ID No. 20 before inactivation, in particular consisted of the nucleotide sequence of SEQ ID No. 20 before inactivation.

[0049] Preferably, the at least one gene encoding erythritol dehydrogenase (Eyd1) of the genetically modified saprotroph comprised a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 20 before inactivation. In another preferred embodiment of the present invention, the at least one gene encoding erythritol dehydrogenase (Eyd1) of the genetically modified saprotroph consisted of a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 20 before inactivation.

[0050] Preferably, the erythritol dehydrogenase (Eyd1) encoded by the nucleotide sequence of SEQ ID No. 20 before inactivation comprises the amino acid sequence of SEQ ID No. 21, in particular consists of the amino acid sequence of SEQ ID No. 21.

[0051] In a preferred embodiment, the erythritol dehydrogenase (Eyd1) encoded by the nucleotide sequence of SEQ ID No. 20 before inactivation comprises an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least

98%, preferably at least 99%, sequence identity to SEQ ID No. 21. Particularly preferred, the erythritol dehydrogenase (Eyd1) encoded by the nucleotide sequence of SEQ ID No. 20 before inactivation consists of an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 21.

[0052] Preferably, the erythritol dehydrogenase (Eyd1) encoded by the nucleotide sequence of SEQ ID No. 20 before inactivation comprises an amino acid sequence as defined in SEQ ID No. 21, in which one, preferably two, preferably three, preferably five, preferably six, preferably seven, preferably eight, preferably nine, preferably ten, amino acids are exchanged by other naturally occurring amino acids. In a preferred embodiment of the present invention, the erythritol dehydrogenase (Eyd1) encoded by the nucleotide sequence of SEQ ID No. 20 before inactivation comprises an amino acid sequence as defined in SEQ ID No. 21, in which at most ten, preferably at most nine, preferably at most eight, preferably at most seven, preferably at most six, preferably at most five, preferably at most four, preferably at most three, preferably at most two, preferably at most one, amino acids are exchanged by other naturally occurring amino acids.

[0053] In a preferred embodiment of the present invention, the at least one further inactivated gene is a gene encoding erythritol isomerase 1 (Eyi1). Preferably, the at least one gene encoding erythritol isomerase 1 (Eyi1) of the genetically modified saprotroph comprised the nucleotide sequence of SEQ ID No. 22 before inactivation, in particular consisted of the nucleotide sequence of SEQ ID No. 22 before inactivation.

[0054] Particularly preferred, the at least one gene encoding erythritol isomerase 1 (Eyi1) of the genetically modified saprotroph comprised a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 22 before inactivation. In a further preferred embodiment of the present invention, the at least one gene encoding erythritol isomerase 1 (Eyi1) of the genetically modified saprotroph consisted of a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 22 before inactivation.

[0055] Preferably, the erythritol isomerase 1 (Eyi1) encoded by the nucleotide sequence of SEQ ID No. 22 before inactivation comprises the amino acid sequence of SEQ ID No. 23, in particular consists of the amino acid sequence of SEQ ID No. 23.

[0056] In a preferred embodiment, the erythritol isomerase 1 (Eyi1) encoded by the nucleotide sequence of SEQ ID No. 22 before inactivation comprises an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 23. Particularly preferred, the erythritol isomerase 1 (Eyi1) encoded by the nucleotide sequence of SEQ ID No. 22 before inactivation consists of an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably

at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 23.

[0057] Preferably, the erythritol isomerase 1 (Eyi1) encoded by the nucleotide sequence of SEQ ID No. 22 before inactivation comprises an amino acid sequence as defined in SEQ ID No. 23, in which one, preferably two, preferably three, preferably five, preferably six, preferably seven, preferably eight, preferably nine, preferably ten, amino acids are exchanged by other naturally occurring amino acids. In a preferred embodiment of the present invention, the erythritol isomerase 1 (Eyi1) encoded by the nucleotide sequence of SEQ ID No. 22 before inactivation comprises an amino acid sequence as defined in SEQ ID No. 23, in which at most ten, preferably at most nine, preferably at most eight, preferably at most seven, preferably at most six, preferably at most five, preferably at most four, preferably at most three, preferably at most two, preferably at most one, amino acids are exchanged by other naturally occurring amino acids.

[0058] In another preferred embodiment of the present invention, the at least one further inactivated gene is a gene encoding erythritol isomerase 2 (Eyi2). Preferably, the at least one gene encoding erythritol isomerase 2 (Eyi2) of the genetically modified saprotroph comprised the nucleotide sequence of SEQ ID No. 24 before inactivation, in particular consisted of the nucleotide sequence of SEQ ID No. 24 before inactivation.

[0059] Preferably, the at least one gene encoding erythritol isomerase 2 (Eyi2) of the genetically modified saprotroph comprised a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 24 before inactivation. In a preferred embodiment of the present invention, the at least one gene encoding erythritol isomerase 2 (Eyi2) of the genetically modified saprotroph consisted of a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 24 before inactivation.

[0060] Preferably, the erythritol isomerase 2 (Eyi2) encoded by the nucleotide sequence of SEQ ID No. 24 before inactivation comprises the amino acid sequence of SEQ ID No. 25, in particular consists of the amino acid sequence of SEQ ID No. 25.

[0061] In a preferred embodiment, the erythritol isomerase 2 (Eyi2) encoded by the nucleotide sequence of SEQ ID No. 24 before inactivation comprises an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 25. Particularly preferred, the erythritol isomerase 2 (Eyi2) encoded by the nucleotide sequence of SEQ ID No. 24 before inactivation consists of an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 25.

[0062] Preferably, the erythritol isomerase 2 (Eyi2) encoded by the nucleotide sequence of SEQ ID No. 24

before inactivation comprises an amino acid sequence as defined in SEQ ID No. 25, in which one, preferably two, preferably three, preferably five, preferably six, preferably seven, preferably eight, preferably nine, preferably ten, amino acids are exchanged by other naturally occurring amino acids. In a preferred embodiment of the present invention, the erythritol isomerase 2 (Eyi2) encoded by the nucleotide sequence of SEQ ID No. 24 before inactivation comprises an amino acid sequence as defined in SEQ ID No. 25, in which at most ten, preferably at most nine, preferably at most eight, preferably at most seven, preferably at most six, preferably at most five, preferably at most four, preferably at most three, preferably at most two, preferably at most one, amino acids are exchanged by other naturally occurring amino acids.

[0063] In a further preferred embodiment of the present invention, the at least one further inactivated gene selected from the group consisting of a gene encoding phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhaps-1), a gene encoding erythritol utilization factor (EUF), a gene encoding erythrulose kinase (EYK1), a gene encoding erythritol dehydrogenase (EYD1), a gene encoding erythritol isomerase (EYI1) and a gene encoding erythritol isomerase (EYI2) of the genetically modified saprotroph is deleted.

[0064] Preferably, the at least one further inactivated gene selected from the group consisting of a gene encoding phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhaps-1), a gene encoding erythritol utilization factor (EUF), a gene encoding erythrulose kinase (EYK1), a gene encoding erythritol dehydrogenase (EYD1), a gene encoding erythritol isomerase (EYI1) and a gene encoding erythritol isomerase (EYI2) of the genetically modified saprotroph is non-functional.

[0065] In a particularly preferred embodiment, the at least one further inactivated gene selected from the group consisting of a gene encoding phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhaps-1), a gene encoding erythritol utilization factor (EUF), a gene encoding erythrulose kinase (EYK1), a gene encoding erythritol dehydrogenase (EYD1), a gene encoding erythritol isomerase (EYI1) and a gene encoding erythritol isomerase (EYI2) of the genetically modified saprotroph is inactivated by gene knock-out, in particular gene replacement. Preferably, the at least one further inactivated gene selected from the group consisting of a gene encoding phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhaps-1), a gene encoding erythritol utilization factor (EUF), a gene encoding erythrulose kinase (EYK1), a gene encoding erythritol dehydrogenase (EYD1), a gene encoding erythritol isomerase (EYI1) and a gene encoding erythritol isomerase (EYI2) of the genetically modified saprotroph is inactivated by gene replacement using a deletion cassette.

[0066] In a preferred embodiment of the present invention, the at least one further inactivated gene selected from the group consisting of a gene encoding phospho-2-dehydro-3-deoxyheptonate aldolase 1 (Dhaps-1), a gene encoding erythritol utilization factor (EUF), a gene encoding erythrulose kinase (EYK1), a gene encoding erythritol dehydrogenase (EYD1), a gene encoding erythritol isomerase (EYI1) and a gene encoding erythritol isomerase (EYI2) of the genetically modified saprotroph is inactivated, in particular made non-functional, by genome editing, in particular by meganucleases, zinc-finger nucleases (ZFNs), tran-

scription activator-like effector nucleases (TALENs) or by the clustered regularly interspaced short palindromic repeats (CRISPR) system.

[0067] In a further preferred embodiment of the present invention, the genetically modified saprotroph further comprises at least one gene encoding at least one transketolase. Preferably, the at least one gene encoding at least one transketolase is tk11, in particular codon-optimized tk11. Preferably, the genetically modified saprotroph further comprises at least one gene encoding tk11 from *T. reesei*, in particular codon-optimized tk11 from *T. reesei*. In a preferred embodiment, the at least one gene encoding at least one transketolase comprises the nucleotide sequence of SEQ ID No. 5, in particular consists of the nucleotide sequence of SEQ ID No. 5.

[0068] In a further preferred embodiment of the present invention, the at least one gene encoding at least one transketolase comprises a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 5. Preferably, the at least one gene encoding at least one transketolase consists of a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 5.

[0069] Preferably, the at least one transketolase comprises the amino acid sequence of SEQ ID No. 6, in particular consists of the amino acid sequence of SEQ ID No. 6.

[0070] In another preferred embodiment of the present invention, the at least one transketolase comprises an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 99%, sequence identity to SEQ ID No. 6. Particularly preferred, the at least one transketolase consists of an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 6.

[0071] In a further preferred embodiment of the present invention, the at least one transketolase comprises an amino acid sequence as defined in SEQ ID No. 6, in which one, preferably two, preferably three, preferably five, preferably six, preferably seven, preferably eight, preferably nine, preferably ten, amino acids are exchanged by other naturally occurring amino acids. In a particularly preferred embodiment, the at least one transketolase comprises an amino acid sequence as defined in SEQ

[0072] ID No. 6, in which at most ten, preferably at most nine, preferably at most eight, preferably at most seven, preferably at most six, preferably at most five, preferably at most four, preferably at most three, preferably at most two, preferably at most one, amino acids are exchanged by other naturally occurring amino acids. In a preferred embodiment of the present invention, the genetically modified saprotroph further comprises at least one gene encoding at least one transaldolase.

[0073] Preferably, the at least one gene encoding at least one transaldolase is tal1, in particular codon-optimized tal1. Preferably, the genetically modified saprotroph further comprises at least one gene encoding tal1 from *T. reesei*, in particular codon-optimized tal1 from *T. reesei*. In a preferred

embodiment, the at least one gene encoding at least one transaldolase comprises the nucleotide sequence of SEQ ID No. 7, in particular consist of the nucleotide sequence of SEQ ID No. 7.

[0074] In another preferred embodiment, the at least one gene encoding at least one transaldolase comprises a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 7. Preferably, the at least one gene encoding at least one transaldolase consists of a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 7.

[0075] Preferably, the at least one transaldolase comprises the amino acid sequence of SEQ ID No. 8, in particular consists of the amino acid sequence of SEQ ID No. 8.

[0076] Particularly preferred, the at least one transaldolase comprises an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 8. Preferably, the at least one transaldolase consists of an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 8.

[0077] In another preferred embodiment of the present invention, the at least one transaldolase comprises an amino acid sequence as defined in SEQ ID No. 8, in which one, preferably two, preferably three, preferably five, preferably six, preferably seven, preferably eight, preferably nine, preferably ten, amino acids are exchanged by other naturally occurring amino acids. In a particularly preferred embodiment, the at least one transaldolase comprises an amino acid sequence as defined in SEQ ID No. 8, in which at most ten, preferably at most nine, preferably at most eight, preferably at most seven, preferably at most six, preferably at most five, preferably at most four, preferably at most three, preferably at most two, preferably at most one, amino acids are exchanged by other naturally occurring amino acids.

[0078] In a further preferred embodiment of the present invention, the genetically modified saprotroph further comprises at least one gene encoding at least one erythritol utilization factor (EUF). Preferably, the at least one gene encoding at least one erythritol utilization factor, is *euf1*, in particular codon-optimized *euf1*. Preferably, the genetically modified saprotroph further comprises at least one gene encoding *euf1* from *T. reesei*, in particular codon-optimized *euf1* from *T. reesei*. In a preferred embodiment, the at least one gene encoding at least one erythritol utilization factor (EUF) comprises the nucleotide sequence of SEQ ID No. 16, in particular consist of the nucleotide sequence of SEQ ID No. 16.

[0079] In a preferred embodiment, the at least one gene encoding at least one erythritol utilization factor comprises a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 16. Preferably, the at least one gene encoding

at least one erythritol utilization factor consists of a nucleotide sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 16.

[0080] Preferably, the at least one erythritol utilization factor (EUF) comprises the amino acid sequence of SEQ ID No. 17, in particular consists of the amino acid sequence of SEQ ID No. 17.

[0081] In a further preferred embodiment, the at least one erythritol utilization factor comprises an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 17. Preferably, the at least one erythritol utilization factor consists of an amino acid sequence having at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 96%, preferably at least 97%, preferably at least 98%, preferably at least 99%, sequence identity to SEQ ID No. 17.

[0082] In a further preferred embodiment of the present invention, the at least one erythritol utilization factor comprises an amino acid sequence as defined in SEQ ID No. 17, in which one, preferably two, preferably three, preferably five, preferably six, preferably seven, preferably eight, preferably nine, preferably ten, amino acids are exchanged by other naturally occurring amino acids. In a particularly preferred embodiment, the at least one erythritol utilization factor comprises an amino acid sequence as defined in SEQ ID No. 17, in which at most ten, preferably at most nine, preferably at most eight, preferably at most seven, preferably at most six, preferably at most five, preferably at most four, preferably at most three, preferably at most two, preferably at most one, amino acids are exchanged by other naturally occurring amino acids.

[0083] In a preferred embodiment of the present invention, the at least one gene encoding at least one membrane-bound alditol transporter, the at least one gene encoding at least one erythrose reductase, the at least one gene encoding at least one transketolase, the at least one gene encoding at least one transaldolase and/or the at least one gene encoding at least one erythritol utilization factor (EUF) of the genetically modified saprotroph is an exogenous polynucleotide sequence.

[0084] Preferably, the at least one gene encoding at least one membrane-bound alditol transporter, the at least one gene encoding at least one erythrose reductase, the at least one gene encoding at least one transketolase, the at least one gene encoding at least one transaldolase and/or the at least one gene encoding at least one erythritol utilization factor (EUF) of the genetically modified saprotroph is an endogenous polynucleotide sequence.

[0085] In a further preferred embodiment, the genetically modified saprotroph according to the present invention comprises at least one exogenous gene encoding at least one membrane-bound alditol transporter. Preferably, genetically modified saprotroph according to the present invention comprises at least one endogenous gene encoding at least one membrane-bound alditol transporter.

[0086] Preferably, the genetically modified saprotroph according to the present invention comprises at least one exogenous gene encoding at least one erythrose reductase. In

another preferred embodiment, the genetically modified saprotroph according to the present invention comprises at least one endogenous gene encoding at least one erythrose reductase.

[0087] Further preferred, the genetically modified saprotroph according to the present invention comprises at least one exogenous gene encoding at least one transketolase. In a preferred embodiment of the present invention, the genetically modified saprotroph according to the present invention comprises at least one endogenous gene encoding at least one transketolase.

[0088] Preferably, the genetically modified saprotroph according to the present invention comprises at least one exogenous gene encoding at least one transaldolase. Preferably, the genetically modified saprotroph according to the present invention comprises at least one endogenous gene encoding at least one transaldolase.

[0089] Preferably, the genetically modified saprotroph according to the present invention comprises at least one exogenous gene encoding at least one erythritol utilization factor (EUF). Preferably, the genetically modified saprotroph according to the present invention comprises at least one endogenous gene encoding at least one erythritol utilization factor (EUF).

[0090] In a preferred embodiment of the present invention, the at least one gene encoding at least one membrane-bound alditol transporter, the at least one gene encoding at least one erythrose reductase, the at least one gene encoding at least one transketolase, the at least one gene encoding at least one transaldolase and/or the at least one gene encoding at least one erythritol utilization factor (EUF) is stably or transiently introduced into the genome of the genetically modified saprotroph.

[0091] In a particularly preferred embodiment of the present invention, the at least one gene encoding at least one membrane-bound alditol transporter, the at least one gene encoding at least one erythrose reductase, the at least one gene encoding at least one transketolase, the at least one gene encoding at least one transaldolase and/or the at least one gene encoding at least one erythritol utilization factor (EUF), is overexpressed.

[0092] Preferably, the at least one gene encoding at least one membrane-bound alditol transporter, in particular the at least one exogenous or endogenous gene encoding at least one membrane-bound alditol transporter, is overexpressed. In a further preferred embodiment, the at least one gene encoding at least one erythrose reductase, in particular the at least one exogenous or endogenous gene encoding at least one erythrose reductase, is overexpressed. Preferably, the at least one gene encoding at least one transketolase, in particular the at least one exogenous or endogenous gene encoding at least one transketolase, is overexpressed. Further preferred, the at least one gene encoding at least one transaldolase, in particular the at least one exogenous or endogenous gene encoding at least one transaldolase, is overexpressed. Preferably, the at least one gene encoding at least one erythritol utilization factor (EUF), in particular the at least one exogenous or endogenous gene encoding at least one erythritol utilization factor (EUF), is overexpressed.

[0093] In a further preferred embodiment of the present invention, the at least one gene encoding at least one membrane-bound alditol transporter, the at least one gene encoding at least one erythrose reductase, the at least one gene encoding at least one transketolase, the at least one

gene encoding at least one transaldolase and/or the at least one gene encoding at least one erythritol utilization factor (EUF) is under the control of a constitutive or inducible promoter. In a preferred embodiment, the constitutive or inducible promoter is a naturally occurring promoter. In a further preferred embodiment, the constitutive or inducible promoter is a synthetic promoter. Preferably, the constitutive promoter is selected from *pki*, *tef*, *gpd*. The inducible promoter is preferably selected from *bga1*, *bx11*, *cbh1*, *cbh2*, *xyn1*, *xyn2*. Preferably, the at least one gene encoding at least one membrane-bound alditol transporter is under the control of a promoter selected from *pki*, *tef*, *gpd*, preferably under control of a *pki* promoter. Preferably, the at least one gene encoding at least one erythrose reductase is under the control of a promoter selected from *pki*, *tef*, *gpd*, preferably under control of a *tef* promoter. Preferably, the at least one gene encoding at least one transketolase is under the control of a promoter selected from *pki*, *tef*, *gpd*, preferably under control of a *tef* promoter. Preferably, the at least one gene encoding at least one transaldolase is under the control of a promoter selected from *pki*, *tef*, *gpd*, preferably under control of a *tef* promoter. Preferably, the at least one gene encoding at least one erythritol utilization factor (EUF) is under the control of a promoter selected from *pki*, *tef*, *gpd*, preferably under control of a *tef* promoter.

[0094] In a further preferred embodiment of the present invention, the at least one gene encoding at least one membrane-bound alditol transporter, the at least one gene encoding at least one erythrose reductase, the at least one gene encoding at least one transketolase, the at least one gene encoding at least one transaldolase and/or the at least one gene encoding at least one erythritol utilization factor (EUF) is present on a plasmid. Preferably, each of the at least one gene encoding at least one membrane-bound alditol transporter, the at least one gene encoding at least one erythrose reductase, the at least one gene encoding at least one transketolase, the at least one gene encoding at least one transaldolase and/or the at least one gene encoding at least one erythritol utilization factor (EUF) is present on a separate plasmid.

[0095] In a particularly preferred embodiment of the present invention, the genetically modified saprotroph is the *Trichoderma reesei* strain deposited at the Westerdijk Fungal Biodiversity Institute under CBS number 146708.

[0096] The present invention further pertains to a method for the production of erythritol, comprising the steps:

[0097] a) providing at least one genetically modified saprotroph according to the present invention,

[0098] b) culturing the at least one genetically modified saprotroph provided in step a) in the presence of a culture medium, so as to obtain erythritol in the culture medium,

[0099] c) recovering erythritol from the culture medium.

[0100] In a preferred embodiment of the present invention, the culture medium comprises lignocellulosic biomass, in particular straw, and/or at least one residue of dairy product production, in particular whey. Particularly preferred, the culture medium comprises lignocellulosic biomass, in particular straw, preferably hydrolysed straw. According to this particular embodiment, the genetically modified saprotroph is able to grow on lignocellulosic biomass as carbon source, in particular on lignocellulosic biomass as the sole carbon source. In a further preferred embodiment of the present

invention, the culture medium comprises at least one residue of dairy product production. Particularly preferred, the culture medium comprises whey. In this way, the method according to the present invention advantageously utilizes inexpensive, abundant and renewable substrates as educts for the production of erythritol.

[0101] In a preferred embodiment of the present invention, the culture medium comprises ammonium as a nitrogen source. In a further preferred embodiment of the present invention, the culture medium comprises nitrate as a nitrogen source.

[0102] In a further preferred embodiment of the present invention, the culture medium comprises glycerol.

[0103] In a preferred embodiment of the present invention, the at least one genetically modified saprotroph provided in step a) is cultured in step b) at a temperature of 25 to 35° C., preferably 26 to 34° C., preferably 27 to 33° C., preferably 28 to 32° C., preferably 29 to 31° C., preferably 30° C.

[0104] Preferably, the pH of the culture medium in step b) is within the range of 3 to 8, preferably 3 to 7, preferably 3.5 to 6.5, preferably 3.5 to 6, preferably 4 to 5.5, preferably 4 to 5.

[0105] In a particularly preferred embodiment of the present invention, step b) is conducted until a concentration of erythritol in the culture medium of at least 1 g/L, preferably at least 1.5 g/L, preferably at least 2 g/L, preferably at least 2.5 g/L, preferably at least 3 g/L, is reached.

[0106] In a preferred embodiment of the present invention, in step b) at least one substrate that causes osmotic stress is added to the culture medium. Preferably, the at least one substrate that causes osmotic stress is added to the culture medium after 10 hours of cultivation, preferably 12 hours of cultivation, preferably 14 hours of cultivation, preferably 16 hours of cultivation.

[0107] In a further preferred embodiment of the present invention, the substrate that causes osmotic stress is selected from glycerol and sodium chloride or a combination thereof. Preferably, the substrate that causes osmotic stress is glycerol. Preferably, the substrate that causes osmotic stress is sodium chloride.

[0108] In a preferred embodiment of the present invention, in step c) erythritol is recovered by crystallisation. Particularly preferred, the recovery of erythritol from the culture medium in step c) comprises the steps:

[0109] i) removal of biomass, preferably removal of biomass by filtration,

[0110] ii) decolorisation of the culture medium, preferably decolorisation of the culture medium with active carbon,

[0111] iii) desalting the culture medium, preferably desalting and decolorizing the culture medium with at least one ion exchange resin, and

[0112] iv) concentrating and crystallising the erythritol.

[0113] The present invention also pertains to the use of a genetically modified saprotroph according to the present invention for the production of erythritol.

[0114] The term “genetically modified” and its grammatical equivalents as used herein refer to one or more alterations of a nucleic acid, e.g. the nucleic acid within the genome of an organism, in particular within the genome of a saprotroph. A “genetically modified” saprotroph can refer to a saprotroph with an added, deleted and/or altered, in particular inactivated, gene.

[0115] In the context of the present invention, the term “erythrose reductase” pertains to any enzyme that catalyses reversibly the reduction of an aldose to an alditol, wherein the enzyme has predominant specific activity for the reduction of erythrose to erythritol. Thus, an “erythrose reductase” according to the present invention exhibits a high specificity and affinity for the substrate erythrose. Particularly, the “erythrose reductase” has a higher specificity for erythrose than for any other substrate, such as glycerol.

[0116] The term “membrane-bound alditol transporter” as used in the context of the present invention designates a membrane protein suitable to reversibly transport alditols across the outer membrane, in particular to transport intracellularly produced alditols to the culture medium. According to the present invention, a “membrane-bound alditol transporter” is suitable to transport intracellularly produced erythritol across the membrane to the culture medium.

[0117] In the context of the present invention, the term “transketolase” refers to an enzyme of the pentose phosphate pathway catalysing the thiamine-dependent reversible transfer of a C-2 unit from a ketose donor to an acceptor aldopentose. In particular, the “transketolase” catalyses the transfer of a C-2 unit from xylulose-5-phosphate to ribose-5-phosphate forming seduheptulose-7-phosphate and glyceraldehyde-3-phosphate.

[0118] The term “transaldolase” designates an enzyme of the non-oxidative branch of the pentose phosphate pathway catalysing the reversible transfer of a C-3 unit from a ketose donor to an acceptor aldose. In particular, the “transketolase” catalyses the transfer of a C-3 unit from seduheptulose-7-phosphate to glyceraldehyde-3-phosphate forming erythrose-4-phosphate and fructose-6-phosphate.

[0119] In the context of the present invention an “inactivated” gene is a gene which can temporarily or permanently not be transcribed or can be transcribed but the transcript is not accessible for gene translation. According to the present invention, the inactivation of a gene includes but is not limited to inactivation by partial or complete gene deletion, partial or complete gene replacement, gene repression, gene insertion, and gene mutation. “Inactivation” in the context of the present invention further includes any post-transcriptional gene regulation, in particular by RNAi or siRNA, resulting in inhibition of translation of the transcript into the gene product.

[0120] In the context of the present invention a “non-functional” gene is a gene that is present in the genome of an organism or a plasmid in an organism which gene is either not transcribed or is transcribed but the transcript is not translated to the gene product.

[0121] In the context of the present invention a “deleted” gene is a gene that has previously been present in the genome of an organism or a plasmid in an organism but has been completely removed from the genome of the organism or the plasmid in the organism.

[0122] In the context of the present invention the verb “to overexpress” refers to the artificial expression of a gene in increased quantity. This includes the inducible overexpression of a gene which can be regulated using an inducible promoter and the constant overexpression of a gene using a constitutive promoter.

[0123] In the context of the present invention, the term “a” is meant to include the meaning of “one” or “one or more”.

[0124] In the context of the present invention, the term “comprising” preferably has the meaning of “containing” or

“including”, meaning including the specifically identified elements without excluding the presence of further elements. However, in a preferred embodiment, the term “comprising” is also understood to have the meaning of “consisting essentially of” and in a further preferred embodiment the meaning of “consisting of”.

[0125] The terms “and/or” is used herein to express the disclosure of any combination of individual elements connected within a listing. Solely for illustrative purposes, the expression “A, B, and/or C” can mean A individually, B individually, C individually, (A and B), (B and C), (A and C), and (A, B and C).

[0126] Further preferred embodiments of the invention are subject of the subclaims.

[0127] The invention is further described by way of the following example and the accompanying figures.

[0128] FIG. 1 shows a comparison dry biomass of a *T. reesei* wild-type strain (WT), of a *T. reesei* strain (strain 1) overexpressing the alditol transporter gene *fps1* and comprising an inactivated gene encoding mannitol 1-phosphate 5-dehydrogenase (*mpdh*) and of a *T. reesei* strain (strain 2) overexpressing the alditol transporter gene *fps1* and the gene *err1* encoding erythrose reductase as well as comprising an inactivated gene encoding mannitol 1-phosphate 5-dehydrogenase (*mpdh*) after 40 hours of cultivation.

[0129] FIG. 2 shows the concentration of glucose in the culture medium after 18, 28, 30, 32, 34, 36, 38 and 40 hours of cultivation of the wild-type strain (WT) and strains 1 and 2.

[0130] FIG. 3 shows the concentration of glycerol in the culture medium after 18, 28, 30, 32, 34, 36, 38 and 40 hours of cultivation of the wild-type strain (WT) and strains 1 and 2.

[0131] FIG. 4 shows the concentration of erythritol in the culture medium after 18, 28, 30, 32, 34, 36, 38 and 40 hours of cultivation of the wild-type strain (WT) and strains 1 and 2.

[0132] FIG. 5 shows a plasmid map of pKR_ReAsl1fps1 (see also SEQ ID No. 9).

[0133] FIG. 6 shows a plasmid map of pKR_Rpyr4err1 (see also SEQ ID No. 10).

[0134] FIG. 7 shows a deletion cassette for deletion of *mpdh* (see also SEQ ID No. 11).

EXAMPLE

[0135] 1.1 Comparison of *T. reesei* Strains

[0136] For the characterization of the phenotype the following strains were used:

[0137] the wild type strain QM6aAtmus53, short-termed “WT”,

[0138] QM6aAtmus53 *fps1*(Reas1)*Ampdh*(hygR), short-termed “strain 1”, and

[0139] QM6aAtmus53 *fps1*(Reas1)*OEerr1*(Repyr4)*Ampdh*(hygR), short-termed “strain 2”. This strain has been deposited on May 25, 2020 in the Restricted CBS collection of the Westerdijk Fungal Biodiversity Institute, Utrecht, The Netherlands by the depositor Technische Universität Wien (Gumpendorfer Straße 1a, 1060 Vienna, Austria) under the stipulations of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure. The corresponding viability certificate has been emitted and the strain has been given accession number CBS 146708.

[0140] 1.2 Materials and Methods

[0141] 1.2.1 Generation of *T. reesei* Strains

[0142] The parental *T. reesei* strain used in the experiments is auxotroph for uridine (lack of functional *pyr4*), arginine (lack of functional *asl1*), histidine (lack of functional 80820 HisG) and is further resistant to hygromycin (*hygR*). The first transformation with the plasmid pKR_Reas1fps1 (see FIG. 5) to restore the *asl1* locus not only introduces the *fps1* gene into the genome, but also makes the strain hygromycin sensible again. The *err1* gene encoding erythrose reductase was introduced using the plasmid pKR_Repyr4err1 (see FIG. 6). The deletion of *mpdh* was done by transformation of a deletion cassette (see FIG. 7) to regain hygromycin resistance by homologous recombination. All transformations are performed as protoplast transformations. Typically, 30 to 50 µg of digested plasmid DNA (in 15 µl sterile double-distilled water) were used for transformation of 10⁷ protoplasts (in 200 µl). For selection for resistance against hygromycin B, 100 µl to 2 ml of the transformation reaction mixture was added to 10 ml melted, 50° C. warm malt extract (MEX) agar containing 1.2 M sorbitol. This mixture was poured into sterile petri dishes and incubated at 30° C. for 5 h after solidification. Subsequently, 10 ml melted, 50° C. MEX agar containing 1.2 M sorbitol and a double concentration of hygromycin B was poured on top of the protoplast-containing layer. Plates were incubated at 30° C. for 2 to 5 days until colonies were visible. For selection for prototrophy, 100 µl to 2 ml of the transformation reaction mixture was added to 20 ml melted, 50° C. warm minimal medium agar containing 1.2 M sorbitol and, in the case of *asl1* back transformation, additionally 0.25 mM L-arginine. This mixture was poured into sterile petri dishes and after solidification was incubated at 30° C. for 3 to 7 days until colonies were visible.

[0143] 1.2.2 Cultivation of *T. reesei* and Induction of Erythritol Synthesis

[0144] In the experiment ammonium was used as nitrogen source and glucose as carbon source under osmotic stress. The total duration of cultivation was 40 hours. After 10 hours of cultivation for biomass production, sodium chloride was added to the medium in order to initiate osmotic stress. Culture broth samples have been taken at the time of this addition and next after 28 hours of total cultivation time and then every two hours until end of the cultivation. Samples were centrifuged to separate the mycelium from the supernatant.

[0145] Cultivation was performed with the wild-type strain (WT), QM6a_fm (strain 1) and QM6a_fem (strain 2) in 100 ml shake flasks containing 25 ml cultivation medium ((NH₄)₂SO₄ 2.80 g/l, KH₂PO₄ 4.00 g/l, MgSO₄·xH₂O 1.0 g/l, NaCl 0.5 g/l, peptone from Casein 0.1 g/l, Tween®80 0.5 g/l), supplemented with 1.5 ml/l trace element solution (FeSO₄·xH₂O 0.90 mM, MnSO₄·xH₂O 0.50 mM, ZnSO₄·xH₂O 0.24 mM, CaCl₂·xH₂O 0.68 mM). For inoculation 10⁹ conidiospores per liter were used. To all flasks 0.3 ml of 0.2 M histidine and 0.25 ml of 0.2 M uridine was added. After 10 hours 0.5 ml of 5 M NaCl was added. The agitation rate was 180 rpm, the temperature was 30° C.

[0146] After the cultivation the whole cultivation broth was filtered through miracloth, the mycelia were pressed dry with Whatman filter paper and dried at 80° C. overnight.

[0147] Glucose, glycerol, mannitol and erythritol concentrations were quantified by high-performance liquid chromatography (HPLC) (Shimadzu Prominence Series

equipped with a RID-20A refractive-index detector). Culture supernatants were filtered through 0.20 µm membranes and measured by HPLC using a Rezex RCM-Monosaccharide column (300 mm×7.8 mm; Phenomenex, Torrance, CA, USA). The mobile phase was 30% acetonitrile (HPLC grade). For sample analysis, the column was eluted at 60° C. with 30% acetonitrile (HPLC grade) at a flow rate of 0.6 ml/min.

[0148] 1.3 Results

[0149] The biomass production at the end of the cultivation is nearly the same for all strains. WT produces slightly more biomass (FIG. 1). The glucose consumption is faster in strain 2 compared to WT and strain 1, which is deduced from concentration measurements in the respective supernatants (FIG. 2).

[0150] The glycerol concentration in the supernatant of strain 1 is always higher than in the WT due to the deletion of mannitol 1-phosphate 5-dehydrogenase (FIG. 3) since this deletion results in elimination of the production of mannitol and the strains produce more of the osmolyte glycerol. Whereas strain 1 shows a constant increase in glycerol production, strain 2 bearing the overexpressed *err1* gene shows a fluctuating accumulation of glycerol over time (FIG. 3). Erythritol is only detectable in the supernatants of strains bearing the transporter FPS1 of *Saccharomyces cerevisiae* (FIG. 4). Investigations on the consumption of erythritol demonstrated that all strains, including WT, can grow on erythritol as sole carbon source (data not shown). The presence of erythritol in the supernatant seems to coincide with the secretion of glycerol and it is higher for strain 2 (FIG. 3 and FIG. 4).

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 25

<210> SEQ ID NO 1

<211> LENGTH: 2010

<212> TYPE: DNA

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 1

atgagcaatc cccaaaaggc actgaacgac ttcctcagct ccgagagcgt ccacaccac	60
gactcctcac gtaagcagtc caacaagcaa tcctcggacg agggccgtag ctcgagccag	120
ccctcccacc atcactccgg aggaactaat aacaataaca acaataacaa caacaacaac	180
aacagcaaca acaacaacaa cggtaacgat ggcggaacg acgacgacta tgactacgag	240
atgcaggact accgtccctc cccacagtct gctaggccta ctcccaccta cgtgccccaa	300
tacagtgtcg agtcgggcac cgcattcccc attcaagagg tcattccctc ggccctacatt	360
aacacccaag acatcaacca caaggacaac ggccccctt ccgcacgtgc aaaccgcgct	420
ttccgcccc gcggccagac gaccgtcagc gccaacgtcc tgaacataga ggacttctac	480
aagaatgcgg acgatgcaca cacaattcca gagtcgcatc tgagccggcg tcgctcccca	540
tcacgggcca cgtcaaatgc agggccactc gccaacaccg gcgccaccaa tggctcgtact	600
acgggcgccc agaccaacat ggagagcaac gaggccccca ggaacgtgcc gataatggtc	660
aagcctaaga ccctctacca gaaccccgag acccccacag ttctccctc aacttaccac	720
ccgattaaca agtggagctc cggttaagaat acctatttaa aagagttcct cgccgagttt	780
atgggtacga tggtcgatgat catcttcggc agcgccgctg tatgccaggt caacgtcgcc	840
ggcaagatcc agcaggacaa cttcaacgta gccctcgaca acctcaatgt caccggctct	900
tccgccgaga caatcgacgc catgaagagc ctgactagcc tagtctcctc tgcgcgggt	960
ggcaccttcg atgacgttgc cctcgggtgg gccgctgccg tggtaatggg ctacttttgt	1020
gcaggtggct cggcaatctc tggagccccc ctgaacccat caatcaccct cgctaacctc	1080
gtttatcgag gctttccctt gaagaaggtc ccatattact tcgcaggcca gctgattggt	1140
gctttcacgg gagccctcat tctgttcatt tggtaacaaga gggtccttca ggaggcatac	1200
tcggactggt ggatgaacga atccgtcgct ggcatgtttt gtgtcttccc aaagccctac	1260
ctgagctctg ggccgcagtt ttttagcgag ttctctcgcg gcgccatgct ccaggccggg	1320
acattcgccc tgacggatcc ctacacgtgc ttgagctcag atgtatttcc cctgatgatg	1380

-continued

```

ttcattctca ttttcatcat taacgcttcg atggcatatc agacaggtag tgcgatgaac 1440
ctggcacggg atctggggccc ccgcctcgcc ctctacgcag tgggcttcga ccacaagatg 1500
ctctgggtcc accaccacca ctctcttttg gtgccgatgg tagggccgtt tattggtgcc 1560
ctgatggggg gcctcgtota cgacgtctgt atctaccaag gccacgaatc accggttaac 1620
tgggtctctcc ccgtgtacaa ggaaatgatt atgcgcgcgc ggtttaggcg ccccggatgg 1680
aagaagcgga accgagctag gaggacttct gacctctccg acttctctta caataatgac 1740
gacgacgagg aattcggcga acgcatggcg cttcagaaga cgaagacgaa gtctagcatt 1800
tccgacaacg agaacgaggc cggcgagaag aaggttcagt tcaagtccgt tcagcgaggg 1860
aagcggaact tcggcgcat cccgacaatt cttgaggagg aggatagcat cgagactgcc 1920
agcctcggcg ccacgacaac ggattctatt ggtctctctg acactagctc tgaggacagc 1980
cactacggca acgcgaagaa agtcacctga 2010

```

<210> SEQ ID NO 2

<211> LENGTH: 669

<212> TYPE: PRT

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 2

```

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

```

-continued

Pro	Ile	Asn	Lys	Trp	Ser	Ser	Val	Lys	Asn	Thr	Tyr	Leu	Lys	Glu	Phe	245	250	255
Leu	Ala	Glu	Phe	Met	Gly	Thr	Met	Val	Met	Ile	Ile	Phe	Gly	Ser	Ala	260	265	270
Val	Val	Cys	Gln	Val	Asn	Val	Ala	Gly	Lys	Ile	Gln	Gln	Asp	Asn	Phe	275	280	285
Asn	Val	Ala	Leu	Asp	Asn	Leu	Asn	Val	Thr	Gly	Ser	Ser	Ala	Glu	Thr	290	295	300
Ile	Asp	Ala	Met	Lys	Ser	Leu	Thr	Ser	Leu	Val	Ser	Ser	Val	Ala	Gly	305	310	315
Gly	Thr	Phe	Asp	Asp	Val	Ala	Leu	Gly	Trp	Ala	Ala	Ala	Val	Val	Met	325	330	335
Gly	Tyr	Phe	Cys	Ala	Gly	Gly	Ser	Ala	Ile	Ser	Gly	Ala	His	Leu	Asn	340	345	350
Pro	Ser	Ile	Thr	Leu	Ala	Asn	Leu	Val	Tyr	Arg	Gly	Phe	Pro	Leu	Lys	355	360	365
Lys	Val	Pro	Tyr	Tyr	Phe	Ala	Gly	Gln	Leu	Ile	Gly	Ala	Phe	Thr	Gly	370	375	380
Ala	Leu	Ile	Leu	Phe	Ile	Trp	Tyr	Lys	Arg	Val	Leu	Gln	Glu	Ala	Tyr	385	390	395
Ser	Asp	Trp	Trp	Met	Asn	Glu	Ser	Val	Ala	Gly	Met	Phe	Cys	Val	Phe	405	410	415
Pro	Lys	Pro	Tyr	Leu	Ser	Ser	Gly	Arg	Gln	Phe	Phe	Ser	Glu	Phe	Leu	420	425	430
Cys	Gly	Ala	Met	Leu	Gln	Ala	Gly	Thr	Phe	Ala	Leu	Thr	Asp	Pro	Tyr	435	440	445
Thr	Cys	Leu	Ser	Ser	Asp	Val	Phe	Pro	Leu	Met	Met	Phe	Ile	Leu	Ile	450	455	460
Phe	Ile	Ile	Asn	Ala	Ser	Met	Ala	Tyr	Gln	Thr	Gly	Thr	Ala	Met	Asn	465	470	475
Leu	Ala	Arg	Asp	Leu	Gly	Pro	Arg	Leu	Ala	Leu	Tyr	Ala	Val	Gly	Phe	485	490	495
Asp	His	Lys	Met	Leu	Trp	Val	His	His	His	His	Phe	Phe	Trp	Val	Pro	500	505	510
Met	Val	Gly	Pro	Phe	Ile	Gly	Ala	Leu	Met	Gly	Gly	Leu	Val	Tyr	Asp	515	520	525
Val	Cys	Ile	Tyr	Gln	Gly	His	Glu	Ser	Pro	Val	Asn	Trp	Ser	Leu	Pro	530	535	540
Val	Tyr	Lys	Glu	Met	Ile	Met	Arg	Ala	Trp	Phe	Arg	Arg	Pro	Gly	Trp	545	550	555
Lys	Lys	Arg	Asn	Arg	Ala	Arg	Arg	Thr	Ser	Asp	Leu	Ser	Asp	Phe	Ser	565	570	575
Tyr	Asn	Asn	Asp	Asp	Asp	Glu	Glu	Phe	Gly	Glu	Arg	Met	Ala	Leu	Gln	580	585	590
Lys	Thr	Lys	Thr	Lys	Ser	Ser	Ile	Ser	Asp	Asn	Glu	Asn	Glu	Ala	Gly	595	600	605
Glu	Lys	Lys	Val	Gln	Phe	Lys	Ser	Val	Gln	Arg	Gly	Lys	Arg	Thr	Phe	610	615	620
Gly	Gly	Ile	Pro	Thr	Ile	Leu	Glu	Glu	Glu	Asp	Ser	Ile	Glu	Thr	Ala	625	630	635
Ser	Leu	Gly	Ala	Thr	Thr	Thr	Asp	Ser	Ile	Gly	Leu	Ser	Asp	Thr	Ser			

-continued

645	650	655	
Ser Glu Asp Ser His Tyr Gly Asn Ala Lys Lys Val Thr			
660	665		
<210> SEQ ID NO 3 <211> LENGTH: 996 <212> TYPE: DNA <213> ORGANISM: Trichoderma reesei <400> SEQUENCE: 3			
atgtcttccg gaaggaccgt caccctcaac accggtaca agatcccca gatcggtac		60	
ggcacctggc aggccgctcc cggcgaggtc ggcgctggtg tctttgaggc cctcaagggt		120	
ggctaccgcc acctcgacct ggccaagggtc tacggcaacc agaaggaggt tggtaggggc		180	
atcaagaagg ctcttgctga ggtcccggtc ctgaagcgcg aggatatttt catcacctcc		240	
aagctgtgga acaactccca caagcccgag gacgtcgagc ccgctctcga cgacacctg		300	
gccgagcttg gcctcgacta ccttgacctc tacctcatcc actggcccggt tgcctttgct		360	
cccgcgcccg acctcttccc caagtccgag gacggctccg aggtgcagct caaccagaat		420	
gtgtccattg tccagacctg gaaggccatg accgagctgc ccaagtccaa ggtccgctcc		480	
gtcgggtgtc ccaactttac catagagcac ctcgacgccc tcatcgaggc caccggcggtc		540	
gtcccccgcg tcaaccagat cgagcgccac ccccgctccc ccaaccagcc tctgatcgac		600	
tactgcgcca agaagggcat catcatcacc gcctactccg cctttggcaa caacacaaag		660	
ggcctgcccc tgctcgctcag ctccgacgag gtcaaggccg tcgcccacaa cctgtccaag		720	
aagcagggca agaccgtcac tcccggccag gtcactctcg cctgggtccc gattggtggc		780	
cacaccgtca ttcccaagtc cgtcaccaag gcgcgcattg cggagaactt ccaggaggtt		840	
gagctggatg acgaggccat tgctgcgctg aacaagttgg gcgagaagcc tcagcggttc		900	
aacattcctt acacctacaa gcctaggtgg aacattaacc tgttcaacac cgaggaggag		960	
aaggccgctg cccacactgc tgcatacaag ctgtaa		996	

<210> SEQ ID NO 4 <211> LENGTH: 331 <212> TYPE: PRT <213> ORGANISM: Trichoderma reesei <400> SEQUENCE: 4			
Met Ser Ser Gly Arg Thr Val Thr Leu Asn Thr Gly Tyr Lys Ile Pro			
1	5	10	15
Gln Ile Gly Tyr Gly Thr Trp Gln Ala Ala Pro Gly Glu Val Gly Ala			
20	25	30	
Gly Val Phe Glu Ala Leu Lys Val Gly Tyr Arg His Leu Asp Leu Ala			
35	40	45	
Lys Val Tyr Gly Asn Gln Lys Glu Val Gly Glu Gly Ile Lys Lys Ala			
50	55	60	
Leu Ala Glu Val Pro Gly Leu Lys Arg Glu Asp Ile Phe Ile Thr Ser			
65	70	75	80
Lys Leu Trp Asn Asn Ser His Lys Pro Glu Asp Val Glu Pro Ala Leu			
85	90	95	
Asp Asp Thr Leu Ala Glu Leu Gly Leu Asp Tyr Leu Asp Leu Tyr Leu			
100	105	110	

-continued

Ile	His	Trp	Pro	Val	Ala	Phe	Ala	Pro	Gly	Ala	Asp	Leu	Phe	Pro	Lys
	115						120					125			
Ser	Glu	Asp	Gly	Ser	Glu	Val	Gln	Leu	Asn	Gln	Asn	Val	Ser	Ile	Val
	130					135					140				
Gln	Thr	Trp	Lys	Ala	Met	Thr	Glu	Leu	Pro	Lys	Ser	Lys	Val	Arg	Ser
145					150					155					160
Val	Gly	Val	Ser	Asn	Phe	Thr	Ile	Glu	His	Leu	Asp	Ala	Val	Ile	Glu
				165					170					175	
Ala	Thr	Gly	Val	Val	Pro	Ala	Val	Asn	Gln	Ile	Glu	Arg	His	Pro	Arg
			180					185					190		
Leu	Pro	Asn	Gln	Pro	Leu	Ile	Asp	Tyr	Cys	Ala	Lys	Lys	Gly	Ile	Ile
		195					200					205			
Ile	Thr	Ala	Tyr	Ser	Ala	Phe	Gly	Asn	Asn	Thr	Lys	Gly	Leu	Pro	Leu
	210					215					220				
Leu	Val	Ser	Ser	Asp	Glu	Val	Lys	Ala	Val	Ala	Asp	Asn	Leu	Ser	Lys
225					230					235					240
Lys	Gln	Gly	Lys	Thr	Val	Thr	Pro	Ala	Gln	Val	Ile	Leu	Ala	Trp	Ser
				245					250					255	
Gln	Ile	Gly	Gly	His	Thr	Val	Ile	Pro	Lys	Ser	Val	Thr	Lys	Ala	Arg
			260					265					270		
Ile	Ala	Glu	Asn	Phe	Gln	Glu	Val	Glu	Leu	Asp	Asp	Glu	Ala	Ile	Ala
		275					280					285			
Ala	Leu	Asn	Lys	Leu	Gly	Glu	Lys	Pro	Gln	Arg	Phe	Asn	Ile	Pro	Tyr
	290					295					300				
Thr	Tyr	Lys	Pro	Arg	Trp	Asn	Ile	Asn	Leu	Phe	Asn	Thr	Glu	Glu	Glu
305					310					315					320
Lys	Ala	Ala	Ala	His	Thr	Ala	Val	Ile	Lys	Leu					
				325					330						

<210> SEQ ID NO 5

<211> LENGTH: 2055

<212> TYPE: DNA

<213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 5

atgggctaca cagagggtga ccagaaggcc atcaacacca tccgcctgct cgcggtcgat	60
gccaccttca agagcaactc gggccaccgc ggtgccccca tgggcatggc tcccgttgcc	120
cacgtccttt gggacaagtt catgagggtc aacccaaga accccaagtg gctgaaccgt	180
gaccgattcg ttctctccaa cggccacggc tgcattgtcc agtatgtctt gcttcacctg	240
ttcggctacg acctgacctc cgatgacctc aaggccttcc gaaccgtcga cagcaagact	300
cctggtcacc ccgaggccca cgatactccc ggtgtcgagg tcaccaccgg ccccctgggt	360
cagggatatc ccaatgcctg cggtctggcc atggcccagg cccacaccgc cgccaccttc	420
aacaagcccg gcttcacctc gtgtgacaac tacacctact gcttctcggc cgatggctgc	480
ttgatggagg gtgtctctgc cgaggcctcg tccctggccg gccacctcca gctcggaac	540
ctcatctgca tctacgatga caaccacatc tccattgacg gcgacaccaa ctgtgccttc	600
accgaggacg tggccaagcg ctacgaggcc tacggctggc acgtcgaggt cgtcgagAAC	660
ggcgaccacg acctcgaggc catggaggcc gccatcaaga agtgccagga ggtcaaggac	720
aagccctcaa tcatcaagct gcgcaccacc attggcttcg gctctctcca ggagggcacc	780

-continued

```

cacggagtc acggtctctc cctcaaggcc gacgacatca agcagctcaa gcagaagttt 840
ggcttcaacc ccgacgagag cttcgtcgtc cccaaggagg tctccgacct ctacgcccgc 900
cactcagccg agggcgccgc cgccgaggag gagtggaaca agctgttcgc caagtacgcc 960
gagaagtacc ccgccgaggc tgctgacctc cagcgccgcc tcaagggcga gctgcccag 1020
ggctgggaga agaacctgcc cgtctactcg ccacgcatg ctgccattgc cagccgaaag 1080
ctgtccgaga cggctctcac caagatccac gacgtctgc ccgagctcca ggggtggttc 1140
gccgatctga ctggttccaa cctgacccgc tggaaggacg ctgttgactt ccagccccc 1200
gccaccggcc tcggtgacta ctccggacga tacatccgct acggtgtccg cgagcacggc 1260
atgggcgcca tcataagcg cctggccgcc tacggcacca tcatccccta cggcggtacc 1320
ttctcaact tcgtctccta cgccgcgggt gccgtccgcc tgcctcgtct gtcccgcgtc 1380
cgcgctcatc gggtcgccac ccacgactcc attggcctgg gcgaggacgg tcctactcac 1440
cagcctatcg agactctggc tcacttccgc gccctgccca actgcatggt gtggcgcccc 1500
gccgacggca acgagaccag cgtgcctac tacattgccc tgacctcaa gcacacccc 1560
agcatcattg ccctgtctcg ccagaacctg cccagctcg agggctccgt catcgagaag 1620
gccatcaagg gaggatacgt cctgcaggag gtcgagggtg ccagatcac cctggtctcc 1680
accggttccg aggtttccat cgccgtggat gccgccaagt acctcctgga gaagcacaac 1740
gtcaaggccc gcgtcgtctc catgccctgc ttcgaggttt tcgatgctca gcccaggag 1800
tacagactct ccgtctgcc cgacggcatc cctctctgt ccattgaggt gatgagcacc 1860
atgggctggg agcgatacac tcacgagcag ttggtctca accgattcgg tgcctctggc 1920
gcctacaagg atgtctacaa gaagtccgag ttcacccccc agggcatcac caagcgcgt 1980
cttctcacca ttgactctg gaaggacgtc cccaacgtcc gctctcccat caaccgtgcc 2040
ttccagcagc tcatac 2055

```

<210> SEQ ID NO 6

<211> LENGTH: 685

<212> TYPE: PRT

<213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 6

```

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

```

Leu 130	Ala	Met	Ala	Gln	Ala	His 135	Thr	Ala	Ala	Thr	Phe 140	Asn	Lys	Pro	Gly
Phe 145	Thr	Leu	Cys	Asp	Asn 150	Tyr	Thr	Tyr	Cys	Phe 155	Leu	Gly	Asp	Gly	Cys 160
Leu	Met	Glu	Gly	Val 165	Ser	Ala	Glu	Ala	Ser	Ser	Leu	Ala	Gly	His 175	Leu
Gln	Leu	Gly	Asn	Leu 180	Ile	Cys	Ile	Tyr 185	Asp	Asp	Asn	His	Ile	Ser	Ile
Asp	Gly	Asp	Thr	Asn 195	Cys	Ala	Phe 200	Thr	Glu	Asp	Val	Ala 205	Lys	Arg	Tyr
Glu	Ala	Tyr	Gly	Trp 210	His	Val 215	Glu	Val	Val	Glu	Asn 220	Gly	Asp	His	Asp
Leu 225	Glu	Ala	Met	Glu	Ala 230	Ala	Ile	Lys	Lys	Cys 235	Gln	Glu	Val	Lys	Asp 240
Lys	Pro	Ser	Ile	Ile 245	Lys	Leu	Arg	Thr 250	Thr	Ile	Gly	Phe	Gly	Ser 255	Leu
Gln	Glu	Gly	Thr	His 260	Gly	Val	His	Gly 265	Ser	Pro	Leu	Lys	Ala	Asp 270	Asp
Ile	Lys	Gln	Leu	Lys 275	Gln	Lys	Phe 280	Gly	Phe	Asn	Pro	Asp 285	Glu	Ser	Phe
Val 290	Val	Pro	Lys	Glu	Val	Ser 295	Asp	Leu	Tyr	Ala	Arg 300	His	Ser	Ala	Glu
Gly 305	Ala	Ala	Ala	Glu	Glu 310	Glu	Trp	Asn	Lys	Leu 315	Phe	Ala	Lys	Tyr	Ala 320
Glu	Lys	Tyr	Pro	Ala 325	Glu	Ala	Ala	Asp	Leu 330	Gln	Arg	Arg	Leu	Lys 335	Gly
Glu	Leu	Pro	Glu	Gly 340	Trp	Glu	Lys	Asn 345	Leu	Pro	Val	Tyr	Ser	Pro 350	Ser
Asp	Ala	Ala	Ile	Ala 355	Ser	Arg	Lys 360	Leu	Ser	Glu	Thr	Val 365	Leu	Thr	Lys
Ile 370	His	Asp	Val	Leu	Pro 375	Glu	Leu	Gln	Gly	Gly	Ser 380	Ala	Asp	Leu	Thr
Gly 385	Ser	Asn	Leu	Thr	Arg 390	Trp	Lys	Asp	Ala	Val 395	Asp	Phe	Gln	Pro	Pro 400
Ala	Thr	Gly	Leu	Gly 405	Asp	Tyr	Ser	Gly	Arg 410	Tyr	Ile	Arg	Tyr	Gly 415	Val
Arg	Glu	His	Gly	Met 420	Gly	Ala	Ile	Met 425	Asn	Gly	Leu	Ala	Ala	Tyr 430	Gly
Thr	Ile	Ile	Pro	Tyr 435	Gly	Gly	Thr	Phe 440	Leu	Asn	Phe	Val 445	Ser	Tyr	Ala
Ala	Gly 450	Ala	Val	Arg	Leu 455	Ser	Ala	Leu	Ser	Arg	Val 460	Arg	Val	Ile	Trp
Val 465	Ala	Thr	His	Asp 470	Ser	Ile	Gly	Leu	Gly	Glu 475	Asp	Gly	Pro	Thr	His 480
Gln	Pro	Ile	Glu	Thr 485	Leu	Ala	His	Phe	Arg 490	Ala	Leu	Pro	Asn	Cys 495	Met
Val	Trp	Arg	Pro	Ala 500	Asp	Gly	Asn	Glu 505	Thr	Ser	Ala	Ala	Tyr 510	Tyr	Ile
Ala	Leu	Thr	Ser	Lys 515	His	Thr	Pro	Ser 520	Ile	Ile	Ala	Leu 525	Ser	Arg	Gln
Asn	Leu	Pro	Gln	Leu	Glu	Gly	Ser	Val	Ile	Glu	Lys	Ala	Ile	Lys	Glu

-continued

530				535				540							
Gly	Tyr	Val	Leu	Gln	Glu	Val	Glu	Gly	Ala	Gln	Ile	Thr	Leu	Val	Ser
545				550				555							560
Thr	Gly	Ser	Glu	Val	Ser	Ile	Ala	Val	Asp	Ala	Ala	Lys	Tyr	Leu	Leu
				565				570							575
Glu	Lys	His	Asn	Val	Lys	Ala	Arg	Val	Val	Ser	Met	Pro	Cys	Phe	Glu
			580					585							590
Val	Phe	Asp	Ala	Gln	Pro	Lys	Glu	Tyr	Arg	Leu	Ser	Val	Leu	Pro	Asp
		595					600					605			
Gly	Ile	Pro	Ser	Leu	Ser	Ile	Glu	Val	Met	Ser	Thr	Met	Gly	Trp	Glu
	610					615					620				
Arg	Tyr	Thr	His	Glu	Gln	Phe	Gly	Leu	Asn	Arg	Phe	Gly	Ala	Ser	Gly
	625				630					635					640
Ala	Tyr	Lys	Asp	Val	Tyr	Lys	Lys	Phe	Glu	Phe	Thr	Pro	Glu	Gly	Ile
			645					650							655
Thr	Lys	Arg	Ala	Leu	Leu	Thr	Ile	Asp	Phe	Trp	Lys	Asp	Val	Pro	Asn
			660					665							670
Val	Arg	Ser	Pro	Ile	Asn	Arg	Ala	Phe	Gln	Gln	Leu	Ile			
		675					680					685			

<210> SEQ ID NO 7
 <211> LENGTH: 975
 <212> TYPE: DNA
 <213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 7

atgagctcct ctcttgacca gctcaaggcc actggcaccc ttgtcgtctc tgacactggt	60
gactttgctg ccctcgccaa gtacaagccc caggatgcca ccaccaaccc ctccctcatc	120
ctggccgcct ccaagaagcc cgagtacgcc aagctgatag acgtcgccat cgactacgcc	180
aagcagaagg gcggcgacat tgagcagcag gtcgacgatg ccctcgaccg cctgctggtc	240
gagttcggca aggagattct caagattgtt cccggcaagg tctccaccga ggtcgacgcc	300
cggttctcct ttgacaccaa ggctctgtc gacaaggccc tccacattat cgagctctac	360
aaggagctcg gcctcccaa ggagcgctc ctcatcaaga tcgcctccac ctgggagggc	420
atcaaggccg ccgagatcct gcagcgcgac cacggcatca actgcaacct gacgctcatg	480
ttctccctgc cccaggccat cggcgccgcc gaggcggcg ccttctctcat ctcccccttc	540
gttgcccgca tctcctgactg gttcaaggcg agcaccaaga aggactactc caaggaggag	600
gaccccggtg tcgcctccgt caagagcatc ttcaactact acaagaagtt tggctacaag	660
accattgtca tgggtgcctc gttccgcaac acgggcgaga ttaccgagct cgctggctgc	720
gactacctga ccatttctcc caacctgtc gaggagctcc tcaactccac cgagcccgtc	780
cccaagaagc tcgacgcgcg cgcgctgcc tcccttaaca ttgagaagaa gtcgtacatc	840
aacgacgagt cctgtttccg cttcgacttc aacgaggacc agatggccgt cgagaagctg	900
cgcgagggca tcagcaagtt tgccgcgcgac gccgtcacc tcaaggacat tctgaggcag	960
aagctgtctg cttaa	975

<210> SEQ ID NO 8
 <211> LENGTH: 324
 <212> TYPE: PRT
 <213> ORGANISM: Trichoderma reesei

-continued

<400> SEQUENCE: 8

```

Met Ser Ser Ser Leu Asp Gln Leu Lys Ala Thr Gly Thr Val Val Val
 1           5           10           15

Ser Asp Thr Gly Asp Phe Ala Ala Ile Ala Lys Tyr Lys Pro Gln Asp
 20           25           30

Ala Thr Thr Asn Pro Ser Leu Ile Leu Ala Ala Ser Lys Lys Pro Glu
 35           40           45

Tyr Ala Lys Leu Ile Asp Val Ala Ile Asp Tyr Ala Lys Gln Lys Gly
 50           55           60

Gly Asp Ile Glu Gln Gln Val Asp Asp Ala Leu Asp Arg Leu Leu Val
 65           70           75           80

Glu Phe Gly Lys Glu Ile Leu Lys Ile Val Pro Gly Lys Val Ser Thr
 85           90           95

Glu Val Asp Ala Arg Phe Ser Phe Asp Thr Lys Ala Ser Val Asp Lys
100           105           110

Ala Leu His Ile Ile Glu Leu Tyr Lys Glu Leu Gly Ile Pro Lys Glu
115           120           125

Arg Val Leu Ile Lys Ile Ala Ser Thr Trp Glu Gly Ile Lys Ala Ala
130           135           140

Glu Ile Leu Gln Arg Asp His Gly Ile Asn Cys Asn Leu Thr Leu Met
145           150           155           160

Phe Ser Leu Pro Gln Ala Ile Gly Ala Ala Glu Ala Gly Ala Phe Leu
165           170           175

Ile Ser Pro Phe Val Gly Arg Ile Leu Asp Trp Phe Lys Ala Ser Thr
180           185           190

Lys Lys Asp Tyr Ser Lys Glu Glu Asp Pro Gly Val Ala Ser Val Lys
195           200           205

Ser Ile Phe Asn Tyr Tyr Lys Lys Phe Gly Tyr Lys Thr Ile Val Met
210           215           220

Gly Ala Ser Phe Arg Asn Thr Gly Glu Ile Thr Glu Leu Ala Gly Cys
225           230           235           240

Asp Tyr Leu Thr Ile Ser Pro Asn Leu Leu Glu Glu Leu Leu Asn Ser
245           250           255

Thr Glu Pro Val Pro Lys Lys Leu Asp Ala Ala Ala Ala Ser Leu
260           265           270

Asn Ile Glu Lys Lys Ser Tyr Ile Asn Asp Glu Ser Leu Phe Arg Phe
275           280           285

Asp Phe Asn Glu Asp Gln Met Ala Val Glu Lys Leu Arg Glu Gly Ile
290           295           300

Ser Lys Phe Ala Ala Asp Ala Val Thr Leu Lys Asp Ile Leu Arg Gln
305           310           315           320

Lys Leu Ser Ala

```

<210> SEQ ID NO 9

<211> LENGTH: 9538

<212> TYPE: DNA

<213> ORGANISM: artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: pKR_ReAs1lfp1

<400> SEQUENCE: 9

ctaaattgta agcgtaaata ttttgtaaa attcgcgtta aatttttggt aaatcagctc 60

-continued

atthtttaac	caataggcgc	aaatcgccaa	aatcccttat	aatcaaaaag	aatagaccga	120
gatagggttg	agtggccgct	acagggcgct	cccattcgcc	attcaggctg	cgcaactgtt	180
gggaaggcgc	tttcgggtgc	ggcctcttcg	ctattacgcc	agctggcgaa	agggggatgt	240
gctgcaaggc	gattaagtgt	ggtaacgcc	gggttttccc	agtcacgacg	ttgtaaaacg	300
acggccagtg	agcgcgacgt	aatacgaact	actatagggc	gaattgagtg	aaggccgtca	360
aggcctaggc	gcgcagcg	ccgcgaaata	gaagacgaag	acgagaggag	gagaggggtgc	420
gagaggggtgc	gaaactgcgc	ccttcagctc	gctaattccg	cccattacat	tcagcgacaca	480
cttggtggta	catacatacc	tacatacgta	cccgtagtgt	tgcgaagcga	gtcgagccag	540
cggctcggatg	tgccaatatg	taccttgtag	gtactcatga	tagtcagcag	gaagccgagg	600
gatgatagca	cgcacagcca	gtctggtgtt	tgtgctctgg	attcagcccc	tagattcact	660
gacaaaaaga	agcaagagaa	gggaaaggga	aaaggaagcc	ctaaaaaaga	gcagggttgt	720
ccttcggcac	caccccgaaa	gcaacagcag	tgatggacgg	gactacgaca	ctgtagtaca	780
tacgccaacc	caaagattcc	gagttgcaat	cattgccaca	aaattgctcg	actgtggcgg	840
gagaggggaa	agtctccgcg	aggcatctc	tgatgagtc	cgtgccacgc	cggttttctt	900
gccttgaatc	accacaaaat	gtctcagcgc	ggctcatcca	tgccgcctt	ttggggcggt	960
tcgtattact	cttgggtttg	actgtattat	cggccaagag	ccaaaggcca	ctcggtttt	1020
ccgccttgc	ctctcttcca	tttcccag	cgctgatagt	atgatagttt	ccctcattg	1080
cagctttgga	cgtcaatgcc	aataactgta	ttgtattgac	aatcatttg	tccaaagacg	1140
accgacacct	ctgatatac	atacagtag	tagtacatac	aatccccccc	aaactccgcc	1200
tttttgccgc	gtgtgtgccg	cgtagctcgt	gcacaatgtc	ggccgcgcgc	cgccagcggc	1260
cacttgctta	gcaggaaacct	agtcagccgt	ccttggtgta	ttaatataac	ggtgagacta	1320
gcccgcggtc	ccttatccag	ctgtccacgt	tgccctgccc	tcagttagcg	ctcaactcat	1380
gccctcactg	gcgaggcgag	ggcaaggatg	gaggggcagc	atcgccctgag	ttggagcaaa	1440
gcgcgcgcat	gggagcagcg	aaccaacgga	gggatgccgt	gctttgtcgt	ggctgctgtg	1500
gccaatccgg	gcccttggtt	ggctcacaga	gcgttgctgt	gagaccatga	gctattattg	1560
ctaggtacag	tatagagaga	ggagagagag	agagagagtg	tgagagggaa	aaggtgaggt	1620
tgaagtgagg	tagtaagtga	aaaaaaaaa	aaaatcaaca	ctgacggctg	ccgtctgcc	1680
ccctcctcca	cccagccacc	tgacacactc	gcgcgcagca	tcacctaatc	ttggctcgcc	1740
ttccgcagct	cagggtgttt	tttttttttc	tctctcctcg	tcgaagccgc	ccttggtccc	1800
ttatttattt	ccctctccta	ccttgctctg	ctttgggtcca	tctgccccct	tgtctgcatc	1860
tcttttgca	gcacgcctt	atcgctcgtc	ctttttttcac	tcacgggagc	ttgacgaaga	1920
cctgactcgt	gagcctcacc	tgctgatttc	tctccccccc	tcccgaccgg	cttgactttt	1980
gtttctctc	cagtacctta	tcgcgaagcc	ggaagaacct	tcttaacctc	tagaatgagc	2040
aatcccaaaa	aggcactgaa	cgacttcctc	agctccgaga	gcgtccacac	ccagactcc	2100
tcacgtaagc	agtcacaaca	gcaatcctcg	gacgagggcc	gtagctcgag	ccagccctcc	2160
caccatcact	ccggagggaac	taataacaat	aacaacaata	acaacaacaa	caacaacagc	2220
aacaacaaca	acaacggtaa	cgatggcggc	aacgacgacg	actatgacta	cgagatgcag	2280
gactaccgtc	cctccccaca	gtctgctagg	cctactccca	cctacgtgcc	ccaatacagt	2340

-continued

gtcagagtcgg	gcaccgcatt	ccccattcaa	gaggtcattc	cctcggccta	cattaacacc	2400
caagacatca	accacaagga	caacggcccc	ccttcgcgat	cgtaaaaccg	cgctttccgc	2460
ccccgcggcc	agacgacgct	cagcgccaac	gtcctgaaca	tagaggactt	ctacaagaat	2520
gggacgatg	cacacacaat	tccagagtcg	catctgagcc	ggcgtcgctc	ccgatcacgg	2580
gccacgtcaa	atgcaggcca	ctccgccaac	accggcgcca	ccaatggtcg	tactacgggc	2640
gccagacca	acatggagag	caacgagtc	cccaggaacg	tgccgataat	ggtaaacct	2700
aagaccctct	accagaaccc	gcagaccccc	acagttctcc	cctcaactta	ccaccgatt	2760
aacaagtgga	gctccgttaa	gaatacctat	ttaaaagagt	tcctcgccga	gtttatgggt	2820
acgatggtea	tgatcatctt	cggcagcgcc	gtcgtatgcc	aggtcaacgt	cgccggcaag	2880
atccagcagg	acaacttcaa	cgtagccctc	gacaacctca	atgtcacggg	ctcttcgcgc	2940
gagacaatcg	acgccatgaa	gagcctgact	agcctagtct	cctctgtcgc	cggtggcacc	3000
ttcgatgacg	ttgccctcgg	gtggggccgct	gccgtggtaa	tgggctactt	ttgtgcaggt	3060
ggctcggcaa	tctctggagc	ccacctgaac	ccatcaatca	ccctcgctaa	cctcgtttat	3120
cgaggctttc	ccctgaagaa	ggccccatat	tacttcgcag	gccagctgat	tggtgctttc	3180
acgggagccc	tcattctgtt	catttggtac	aagagggctc	ttcaggaggc	atactcggac	3240
tggtggatga	acgaatccgt	cgttggcgat	ttttgtgtct	tcccaaagcc	ctacctgagc	3300
tctgggcggc	agtttttttag	cagattctct	tgccggcgcca	tgctccaggc	cgggacattc	3360
gccctgacgg	atccctacac	gtgcttgagc	tcagatgtat	ttccctgat	gatgttcatt	3420
ctcattttca	tcattaacgc	ttcgatggca	tatcagacag	gtactgcgat	gaacctggca	3480
cgggatctgg	gcccccgct	cgccctctac	gcagtgggct	tcgaccacaa	gatgctctgg	3540
gtccaccacc	accactctct	ttgggtgcgc	atggtagggc	cgtttattgg	tgccctgatg	3600
gggggcctcg	tctacgacgt	ctgtatctac	caaggccacg	aatcacccgt	taactggtct	3660
ctccccgtgt	acaaggaaat	gattatgcgc	gcattggtta	ggcgccccgg	atggaagaag	3720
cggaaaccgag	ctaggaggac	ttctgacctc	tccgactctc	cttacaataa	tgacgacgac	3780
gaggaaattcg	gcgaacgcgt	ggcgcttcag	aagacgaaga	cgaagtctag	catttccgac	3840
aacgagaacg	aggccggcga	gaagaagggt	cagttcaagt	ccgttcagcg	agggaagcgg	3900
accttcggcg	gcattcccgac	aattcttgag	gaggaggata	gcacgagac	tgccagcctc	3960
ggcgccacga	caacggattc	tattggtctc	tctgacacta	gctctgagga	cagccactac	4020
ggcaacgcga	agaaagtcc	ctgaatgcgt	ggctttcgtg	accgggcttc	aaacaatgat	4080
gtgcgatggg	gtgggtcccg	gttgccggag	tctttgtcta	ctttggttgt	ctgtcgcagg	4140
tcggtagacc	gcaaatgagc	aactgatgga	ttggtgccag	cgatactata	attcacatgg	4200
atggtctttg	cgatcagtag	ctagtgcgag	agagagaaca	tctatccaca	atgtcgagtg	4260
tctattagac	atactccgag	aataaagtca	actgtgtctg	tgatctaaag	atcgattcgg	4320
cagtcgagta	gcgtataaca	actccgagta	ccagcaaaag	cacgtcgtga	caggagcagg	4380
ctttgccaac	tgccgaacct	tgcttgaatg	aggatacacg	gggtgcaaca	tggtgtact	4440
gatccatcgc	aacaaaaatt	tctgtttata	gatcaagctg	gtagattcca	attactccac	4500
ctcttcgcgt	tctccatgac	atgtaagtgc	acgtaggaaa	ccatacccaa	attgcctaca	4560
gctgcggagc	atgagcctat	ggcgatcagt	ctggctcatgt	taaccagcct	gtgctctgac	4620

-continued

gttaatgcag aatagaaagc cgcggttgca atgcaaatga tgatgccttt gcagaaatgg	4680
cttgctcgct gactgatacc agtaacacct ttgcttgccc gtctagcgct gttgattgta	4740
ttcatcacaa cctcgtctcc ctcccttggg ttgagctctt tggatggctt tccaaacgtt	4800
aatagcgcgt ttttctccac aaagtattcg tatggacgcg cttttggctg tattgcgtga	4860
gctaccagca gccaatttgg cgaagtcttg agccgcactc gcatagaata attgattgcg	4920
catttgatgc gatttttgag cggctgtttc aggcgacatt tcgccccccc ttatttgctc	4980
cattatatca tcgacggcat gtccaatagc cgggtgatag tcttgctgaa tatggctgtc	5040
gtgaataacc catcggcagc agatgataat gattccgcag cacgctagcg agctctccgg	5100
agcccaaaat agaagcacc gaattcaggg cctttttctc cacaccaatc accaaattca	5160
ctgatttcaa ggttggttagc cgcctcttat cactaggtag actagcacct aattttaccg	5220
atatgtgtag tccatacatg tatgagatat acatcatgat ctggaagtaa aaccgggtac	5280
gaagaggat aaaagccgcg gaacaacaac ttgcagatgc catcacgcaa gacgcagaca	5340
cggctctctt ctccggggca tcaccagagc ccaatcacct ggtcctcagg tggccacttg	5400
agcgcgctca gaatggcgc acagcagcag cagcatcacc acagatcttg gcctcaactc	5460
ttgatatttc ggcgagaata caattactag tctagtctat ggaaattgga tgtgtgctcg	5520
ctctttggag actggttagta cataagccac gaacatgata ttggcgggga aggttttagct	5580
ggaaattgga acaagtcaga gaagtcgccc cgtattagct tgggtgttagt agacacagca	5640
agtcaagagt gggagttaga tttctgaatt gagtgacgca gatggatgta gtattcatta	5700
ttgaagacag agcctttcca tcaacgtgca gagccatgaa atgcaatgac cttaaggagg	5760
ggcatctctc cgctttccat ctgactgaga tactaaccgg attccatctc cttttctag	5820
cacacacata cctggagtc aacctcaacc ggccttcagt agctaaacag ctatgtagcg	5880
catccgcgct aagccgctta ggcgataacg ttggccgggtg tcggccatac aattcttgcc	5940
tatgactctg tctcgactaa ctgattgacc aacgtcaatt ccttatcaa gttttcgacc	6000
ttgataactc tcgatcttaa tctcgagctt tttatttctt caacatcatt cacagtaacg	6060
actcaagccc ttggacttgg ccagcaattg acacaaccac gcgccatctg ctccattggc	6120
ctcatcgatc ccatcgactt cgtcagcctc gacaacgcat ccaacgcaat cacaaatact	6180
tcaacatggc tgagaaaaag gcggacgccc cggcagccgg catgctctgg ggtggcagg	6240
tcacaggttg cagcctcgca aatctagctg cggattcagt atactgacct ctgcgatgc	6300
gcaggtgcaa ttgatcctct catgcacaag tacaatgcct ccatcgggta cgacaaggcc	6360
ctgtacaagg aagatatcct tggatccatt gcctttcgca gggccaattc aaaggctggc	6420
atcatcagcg aagacgaatt ccaggcgatc gagcgtggct tgctccaagt gatggaagaa	6480
tggaaagcagg gcactttcgc catcatgccg aatgacgaag atgtgagagg gcacactcga	6540
ttctaactct aaacttcttt cgaggctgac acaatggacc agattcatac ggcaaacgaa	6600
cgtcgcctgg gagaggctcat cggcaaggac atcgcgggaa agcttcacac aggtcgcagc	6660
cgcaacgagc aagttgtctg tgatatggc atgtggcttc gcgatcggat ccgcgaaatc	6720
gacagccaac tcgttgcttt cctcaaggtc attatcacac gtgccgaate tgaatgtaa	6780
gcccagcgt acctttccag ctaaattgcc ccccccccc cccccctc tcttactgac	6840
ggacaacctc agcgactaca tcatgcccg atacacccat ctccagcgcg ctccagccagt	6900

-continued

tccgttggtct	caatggctca	tgtccctatgc	cgtgcggttc	aagcaggatc	tcgagcggtt	6960
gcgtcaagtc	ttcgagaggg	tcaacctcag	tccacttggc	tgtggcgctt	tgcccgcaa	7020
tccattcgga	atcgaccgag	actcaatagc	tgcggagctt	ggcttcagcg	gaatcacctt	7080
gaactccatg	aatacttcgg	ccgaccgga	tttctctgc	gatttctctg	cctggaactc	7140
catcttcacc	aaccacgtca	gcagggtggc	tgaggacctc	atcatctact	cgacctctga	7200
gttcagtttc	gtccgtctgg	ccgacgccta	ttccacgggc	agctctctca	tgcccaacaa	7260
attcaaccgg	gactcgttgg	agcttcttcg	cggcaagtca	gggcgtgctt	tcggcctaat	7320
ggctggtttg	atgatggtga	gtcctcatat	caaacatctc	agccagtttg	acttggtttt	7380
tgagggtgac	cgcggcgcc	gttccggtaa	ttaactggcc	tcattgggctt	tcacttcact	7440
gcccgccttc	cagtcgggaa	acctgtctg	ccagctgcat	taacatggtc	atagctgttt	7500
ccttgcgat	tgggcgctct	ccgcttctc	gtcactgac	tcgctgcgct	cggctcgttcg	7560
ggtaaagcct	gggttgcccta	atgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	7620
gcccgcgttc	tggcggtttt	ccataggttc	cgcaccttcg	acgagcatca	caaaaatcga	7680
cgtcaagtc	agaggtggcg	aaacccgaca	ggactataaa	gataccaggc	gtttccccct	7740
ggaagctccc	tcgtgcgctc	tctgttccg	accctgcgcg	ttaccggata	cctgtccgcc	7800
tttctccctt	cgggaagcgt	ggcgctttct	catagctcac	gctgtaggta	tctcagttcg	7860
gtgtaggtcg	ttcgtcccaa	gctgggctgt	gtgcacgaac	cccccgttca	gcccgaaccgc	7920
tgcgccctat	cggtaacta	tcgtcttgag	tccaacccgg	taagacacga	cttatcgcca	7980
ctggcagcag	ccactggtaa	caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	8040
ttcttgaagt	ggtggcctaa	ctacggctac	actagaagaa	cagtatttgg	tatctgcgct	8100
ctgctgaagc	cagttacctt	cggaaaaaga	gttggtagct	cttgatccgg	caaaacaaacc	8160
accgtggta	gcggtggttt	ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaaggga	8220
tctcaagaag	atcctttgat	cttttctacg	gggtctgacg	ctcagtggaa	cgaaaactca	8280
cgttaaggga	ttttggtcat	gagattatca	aaaaggatct	tcacctagat	cctttttaa	8340
taaaaatgaa	gttttaaatc	aatctaaagt	atatatgagt	aaacttggtc	tgacagttat	8400
tatttaccga	ccactttggt	aatttcgcct	ttcacgtaat	gcacaaattc	ttccagctga	8460
tccgcacggc	tggccagacg	atcttcttcc	tggcccgat	acgcctgacg	cgttccaga	8520
atcacgggct	gatactgcgc	cggcagacgt	tccatggccc	aatccgcgcg	cacatctttc	8580
ggcgcaattt	tgccgggtcac	cgcgctatac	caaatacggc	tcagggtcag	caccacgttg	8640
cgttcacac	ccgccaatc	cggcgggctg	ttccacaggg	tcagggttcc	gttcagcgct	8700
tcaaacaggt	cctgttccgg	caccggatca	aacagttctt	ccgctgccgg	accaaccagc	8760
gccacgctat	gttcacgcgc	tttggtcagc	agaatggcca	gatcgatatc	aatggtcgcc	8820
ggttcaaaaa	tgccggccag	aatatcggtt	cgtgccatt	cgccaaactg	cagctcacgt	8880
ttcgccggat	aacgccacgg	aataatgtca	tcattgcacca	cgatgggtcac	ttccaccgca	8940
cgcagaattt	cgttttcacc	cgggctcgcg	ctggtttcca	gcagatcggt	aatcagcgca	9000
cgcgggttg	tttcatccag	acgcacggtc	acggtcacca	gcagatcaat	atcgctatgc	9060
ggtttcaggc	cgccatccac	cgcgctgcca	tacagatgca	cggccagcag	ggtcggttcc	9120
agatgacgtt	caatcacgcc	caccacttca	gacagctggg	tgctcacttc	cgcaatcacc	9180

-continued

gcttcacgca taatgttcag tttggttttc ggacggctgc cctgctgggt cacatcgttg	9240
ctgctccaca gcatcaggca acggcccacc gcataacgcg cctgctgttt gctcgcacgc	9300
ggcatgctct gcacgccaaa aaagcaatca tagcacgccca taaacaccgc caccgcacca	9360
ttaccgccgc tacgttcggg cagggtacgg ctccaattac ggctacgcat actcttcctt	9420
tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa	9480
tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccac	9538

<210> SEQ ID NO 10
 <211> LENGTH: 11413
 <212> TYPE: DNA
 <213> ORGANISM: artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: pKR_Rpyr4err1

<400> SEQUENCE: 10

ctaaattgta agcgttaata ttttgttaaa attcgcgtta aatttttgtt aaatcagctc	60
attttttaac caataggcgc aaatcggcaa aatcccttat aaatcaaaag aatagaccga	120
gatagggttg agtggccgct acagggcgct cccattcgcc attcagcgtg cgcaactgtt	180
gggaaggcgct tttcgggtgc ggctcttcg ctattacgcc agctggcgaa agggggatgt	240
gctgcaaggc gattaagttg ggtaacgcc gggttttccc agtcacgacg ttgtaaaacg	300
acggccagtg agcgcgacgt aatacagctc actatagggc gaattgagtg aaggccgtca	360
aggcctaggc gcgccagcgc ccgccagatc tcgagatagt atctcattca gtaaaatact	420
cgtacgagtt tgagggggcg cgaggctaag ctgcaagcgc gcgcggcaac actacctgta	480
tgtattaata aagcagcaga gagagagaga gccctggact gccggcaatc ctgagaaggg	540
gggatcgacc ggaccatctg cagcgcagcag ggctgctgct gccaacatgg agagactaaa	600
gatgggattg gtggtacata gcaagacgct tggatgcttg ttgacgggct gctccaataa	660
gcaagctcac aagcagtgtt ttgggggtcc gcacggacgc tcgagagggc tcggacggct	720
gatattgcag acctggacga gtcagttgtg caggtaggca ctcgtagtg tgtctttgtg	780
cttgatcaac cttgagtaag tattattagg gtgatggctt aggtaggtag gttaggtagg	840
tagtttaggt acagggaggt gaggagggga gacgtccagc tgccatatgc gccacttgca	900
gcctcaactc tgcaaaagcc agccccgac acggcatgtc gtacatactg taaggtagta	960
cctagtagta gaacgaatgc taaaggtgca gagcaactaa cttgagcaaa accccccctc	1020
cccccatcga gattgattgc cctctacgct gaaaatagtc tcgccagccg gcgacgctta	1080
gccgtggcgc tacccgcaca tggcgcttgt tagcctgcgc tccaggttga tcggaagcat	1140
gtcgcagggc ggggtctctg cgtgccgaca gccagcagag agacagaaga gagcgagaga	1200
gcgagagagc gagagagaga gagagagaga aagacgcctc ggggggtttg ctggtggctt	1260
cgatgcttgt ccgtgtcaag ggttcacgaa agacgggacg gctgagttgg tgactggcgc	1320
cgggtggcatc gaggctggag tccaggcaca tgcattggcg aaggcatgtc aggccatgtc	1380
gtgtcatgcc ccggcacagc tgaagggaat actactattg agagctgttc tccaaggcgt	1440
caagcatgtg agcaaaaact cgagttgtct agactcgact cgacttgagg ctgcagtaca	1500
acaatacagt actactcgta cacgtatgta cagcaaggag caggtggatc agcaaaagttg	1560
catgataatg gactggaccg gcaacgctca aagttcgctc tgcagctccg gcatgaaggg	1620

-continued

cagacccccg	ccattattag	cgatgactag	ccgtggaggc	cacttttgct	ttagtgattt	1680
aaaccccgag	attcatgctc	gatgctctac	aggatgtat	gtactagctc	agctcagcct	1740
tagtaggtag	gtattgaagt	ctgggtggtc	gaggaggacc	tgcagggtgg	agttttgggg	1800
gaaaaagatg	atgctcacca	actaaaataa	atagtcaggc	cacacacaca	acctactgag	1860
cagaaccacc	acaaccagtg	aagagctact	ccggaattct	cgactccgga	gagttgggca	1920
aatcaggcgc	cggaccagcg	ccaaggtgtg	ggtggattgg	gtcgtaagaa	aaagaggcat	1980
ccctacatgc	agcaaccggc	tggagacgag	ccagcccgag	cactggcagc	ctcagaccag	2040
aggcatctcg	accacggctt	caagtatcag	actgggagct	gtgccgagtt	tgctggctac	2100
ttacctagtc	aaaaaggctg	ccgctgactg	tctgtgtcat	attaacgagc	cagggtctcaa	2160
ccttctgtcg	aatgaaatgg	aaagcggcga	atgactcttc	ggggttgccc	agaaacaacg	2220
acgaccactg	agggatcggg	acgatgtaca	tctactgcga	tatgccgggc	tcgaatgtga	2280
ctgcacacac	cacacacaca	cacacacacg	gacacatctg	agcagtacta	tcaaagcggg	2340
gggagtgcac	ctgcactggg	cacccttact	gtaggtagga	aggataccta	aggtagaggg	2400
ctcgctgctg	caagttgata	tgacggccac	caatcggaag	ggtgtcgtat	ctcacatcac	2460
cagcgtcagg	acacgcttca	caaaattccc	acaagttcca	gttgaaatga	cccgtaggag	2520
agcgctcga	gggacagaat	gtacagtact	atactaccta	ccttagatgc	catcacccct	2580
tcgcaggcca	tggaaattct	ggcaagtgcg	acccttcagc	catgcctgtc	cgccatcatc	2640
ataccagatc	gtgtcgtggc	gggcacttgg	gaaggacgag	ctccaaccgc	cgccgctgat	2700
tggcctgac	cgcacacagc	actggcgctc	gcggcgagga	aaaaaaaaaa	tcttgcgggg	2760
gtgcagcagc	attgagtgga	aagcgggagc	cgctgcacga	cgctgattcg	tccgggctgc	2820
ctcgttaggg	ctagaccagc	aggcggggcg	cgttattttt	gcccctcctc	aggctttttt	2880
tcacaagccc	cgtttttctg	ccccgcgctg	ggcagaaaaa	aaaaatacag	cacgagccac	2940
tggcacttgc	accctccac	ccctccgcg	accacctttt	aaatctcctc	ctcccaccat	3000
ctgctttcga	aattttctct	cttctgctct	tcgtctcgca	taccgggttc	aagcatccga	3060
tctgogaatt	tgtatgtgct	gcctctccct	ctgaccttct	ggtctggtga	taccatctct	3120
cctcagtttg	gatcatcgcc	ttattcttct	tccctcttct	gcattctgct	cctgctcggt	3180
tgaggaacat	cgcacagctg	ctctgcttgc	ctcgcagcga	tctagtcaag	aacaacacag	3240
ctctcacgct	acatcacact	taagtctaga	atgtcttccg	gaaggaccgt	cacctcaac	3300
accggctaca	agatccccc	gatcggtac	ggcacctggc	aggccgctcc	cggcgaggtc	3360
ggcgctgggt	tcttttaggc	cctcaagggt	ggctaccgcc	acctcgacct	ggccaagggt	3420
tacggcaacc	agaaggaggt	tggtgagggc	atcaagaagg	ctcttgctga	ggtccccggc	3480
ctgaagcgcg	aggatatttt	catcacctcc	aagctgtgga	acaactccca	caagcccgag	3540
gacgtcgagc	ccgctctcga	cgacaccctg	gccgagcttg	gcctcgacta	ccttgacctc	3600
tacctcatcc	actggcccg	tgctttgct	cccggcgccg	acctcttccc	caagtccgag	3660
gacggctccg	agggtcgagct	caaccagaat	gtgtccattg	tccagacctg	gaaggccatg	3720
accgagctgc	ccaagtccaa	ggtccgctcc	gtcggtgtct	ccaactttac	catagagcac	3780
ctcgacgccg	tcacgaggc	caccggcgct	gtcccccgcc	tcaaccagat	cgagcgccac	3840
ccccgcctcc	ccaaccagcc	tctgatcgac	tactcgccca	agaagggcac	catcatcacc	3900

-continued

gectactcgg	cctttggcaa	caacacaaag	ggcctgcccc	tgtctgtcag	ctcggacgag	3960
gtcaaggcgg	tcgccgacaa	cctgtccaag	aagcagggca	agaccgtcac	tcccgcccag	4020
gtcatcctcg	cctgggtcca	gattggtggc	cacacgtca	ttcccaagtc	cgtcaccaag	4080
gcgcgcatcg	cggagaactt	ccaggagggt	gagctggatg	acgaggccat	tgtcgcgctg	4140
aacaagttag	gcgagaagcc	tcagcgggtc	aacattcctt	acacctacaa	gcctaggtag	4200
aacattaacc	tgttcaacac	cgaggaggag	aaggccgctg	cccacactgc	tgtcatcaag	4260
ctgtaaaacta	gtgaattctg	gattcctttcg	tgaccggggt	tcaaacatg	atgtgcgatg	4320
gtgtgggttc	cgggtggcgg	agtctttgtc	tactttggtt	gtctgtcgca	ggcggtaga	4380
ccgcaaatga	gcaactgatg	gattgttgcc	agcgatacta	taattcacat	ggatgggtctt	4440
tgcgatcagt	agctagttag	agagagagaa	catctatcca	caatgtcgag	tgtctattag	4500
acatactcgg	agaataaagt	caactgtgtc	tgtgatctaa	agatcgattc	ggcagtcgag	4560
tagcgtataa	caactccgag	taccagcaaa	agcagctcgt	gacaggagca	ggctttgcc	4620
actgcgcaac	cttgcttgaa	tgaggataca	cgggggtgca	catggctgta	ctgatccatc	4680
gcaacaaaaa	tttctgttta	tagatcaagc	tggttagattc	caattactcc	acctcttgcg	4740
cttctccatg	acatgtaagt	gcacgttaga	aaccataccc	aaattgccta	cagctgcgga	4800
gcatgagcct	atggcgatca	gtctggatcat	gttaaccagc	ctgtgtctctg	acgttaaatgc	4860
agaatagaaa	gccgcggttg	caatgcaaat	gatgatgcct	ttgcagaaat	ggcttgctcg	4920
ctgactgata	ccagtaacac	ctttgcttg	ccgtctagcg	ctgttgattg	tattcatcac	4980
aacctcgtct	ccctcctttg	ggttgagctc	tttggtgggc	tttccaaacg	ttaatagcgc	5040
tagtagtaga	aatggagggt	aaactgaaca	tcagtcaaaa	gacgtcgaga	cgcagatact	5100
actagatacc	tatcccagcg	acggaaaaatg	gacaagggtg	cgccgcgcgc	aagctctcag	5160
atgcgctatc	gtgcgatgat	cctctcatca	acaaacggcg	ccagtgcagg	gctagccgtg	5220
ggattgactc	gagctcgctt	cttctttgtg	cttttctcct	gtcgacatcc	cggttgcgcc	5280
ttggacctcg	ccgggacgtg	caaacgaaca	catcactttc	aaagagagtg	tttgatgctc	5340
acgctcggat	gctccagccc	ccccttgctt	gcttggtccat	cattgagggt	cgaaggggca	5400
gttgctcccc	attccctctg	ctccaacagg	aactgtaacc	gaaaaaaaag	ccatcctcct	5460
tcttctctct	cctcttcgcg	ctcctcctct	tctcctcctc	actaatacct	actcctcctc	5520
tcctcatcga	ctagactgac	ccccccgggt	gggccccctg	tcccgctctc	aacagagcac	5580
caccagacaa	agacccttgc	ccgcgcgaat	ccagaccccc	ccagcaattc	cgggcctcgt	5640
tgatcctcct	ctactgtagt	tgtacataca	tacctaccga	ctgcattgca	ttggtacagc	5700
tgcaggcact	tccaggcacg	gccaccaa	tgcagcggcc	cttgcttgct	tgttggttc	5760
gagaactagg	atctgtgtct	tttgcttgct	cttgcttgct	ctgggttcct	gctcgtctgc	5820
ggcaatcggg	acgcccgcag	tgccgtgcca	agcaaacacc	ccaaggtagg	tacctaccac	5880
taggcttctt	ttctcgttgt	ctcactctct	cttttctct	ttgtctctct	ttatcccat	5940
cttttctctc	tctctctgct	cctttcctaa	ccacttcctc	acctttctct	ttttcctttt	6000
cttgctcatc	ccatcttggtc	tgacgaaaaa	ggtctgactg	ggtaggtatt	atctggcaga	6060
cttggtgtga	tcattcacc	tatttctgct	tcatagtaca	tgtactgtac	ctgaacggct	6120
caaccgctat	ttacgactct	tatttttttg	tggcgttggt	cacgtttgct	agctgttgct	6180

-continued

cgtcttttcta	gggctectca	aacttgacct	gaccgagctc	cctttctgga	ccgggtgggc	6240
ttcacttcca	gctgctgagc	gacctgagcc	gaacatcttc	agtccttgtc	cagcgcaatt	6300
cattttcttt	ccttttcttt	ttttttatct	ctttctttac	ttttattctc	tctttttctc	6360
ctcttctct	tcttctctct	tctctctctc	ctccatctcc	tcactctcgt	ctccctcatt	6420
actacctct	cggctctctca	ggctccaccaa	ccctcccgca	cccaaacctc	tgccgctgaa	6480
acccattcgg	tggtcgccgt	tttttttttt	ttttttttct	cacccccaaa	gtcgcaatat	6540
cgggtatcgc	cgcgggcatt	gaatcgccct	ctccgctagc	atcgactact	gctgctctgc	6600
tctcgttgcc	agcgtgctc	cctagaattt	tgaccagggg	acgagcccgca	cattaaagca	6660
actccctcgc	ctcgagacga	ctcggtatgc	acgaaattct	cccaatcgcc	gacagttcct	6720
actctctctc	ctcccgcaag	gctgtcgcgc	ttccaacgtc	attcgcacag	cagaattgtg	6780
ccatctctct	cttttttttc	ccccctctca	aaccgccaca	acggcacctc	aagggttaaa	6840
ctatccaacc	agccgcagcc	tcagctctct	tcagcctcat	cagccatggc	accacacccg	6900
acgctcaagg	ccacctctgc	ggccaggagc	gagacggcga	cgcacccgct	gacggcttac	6960
ctgttcaagc	tcattggacct	caaggcgctc	aacctgtgcc	tgagcgccga	cgtgccgaca	7020
gcgcgcgagc	tgctgtacct	ggccgacaag	attggcccg	cgattgtcgt	gctcaagacg	7080
cactacgaca	tggtctcggg	ctgggacttc	cacccgga	cgggcacggg	agcccagctg	7140
gcgtcgctgg	cgcgcaagca	cggcttctct	atcttcgagg	accgcaagtt	tgccgacatt	7200
ggccacaccc	tcgagctgca	gtacacgggc	gggtcggcgc	gcatcatcga	ctgggcgcac	7260
attgtcaaac	tcaacatggt	gcccggcaag	gcgtcggtgg	cctcgctggc	ccagggcgcc	7320
aagcgctggc	tcgagcgcta	cccctgcgag	gtcaagacgt	ccgtcacctg	cggcacgccc	7380
accatggact	cgtttgacga	cgacgcccgc	tccagggacg	ccgagcccgc	cggcgccgct	7440
aacggcatgg	gctccattgg	cgctctggac	aagcccatct	actcgaaccc	gtccggcgac	7500
ggccgcaagg	gcagcatcgt	ctccatcacc	accgtcacc	agcagtaaga	gtccgtctcc	7560
tcgccccggt	taacaaaggc	catcgccgag	ggcgacgagt	cgctcttccc	gggcatcgag	7620
gaggcgccgc	tgagcccgcg	cctcctgac	ctcgcccaaa	tgtccagcca	gggcaacttc	7680
atgaacaagg	agtacacgca	ggcctcgctc	gaggccgccc	gggagcacia	ggactttgtc	7740
atgggcttca	tctcgcagga	gacgctcaac	accgagcccg	acgatgcctt	tatccacatg	7800
acgcccggct	gccagctgcc	ccccgaagac	gaggaccagc	agaccaacgg	atcggtcgg	7860
ggagacggcc	agggccagca	gtacaacacg	ccgcacaagc	tgattggcat	cgcggcagc	7920
gacattgcca	ttgtgggccc	gggcatcctc	aaggcctcag	accccgtaga	ggaggcagag	7980
cggtagcgat	cagcagcggt	gaaagcctac	accgagaggg	tgctgcgata	ggggagggaa	8040
gggaagaaa	aagtaaagaa	aggcatttag	caagaagggg	gaaaaggagg	ggaggacaaa	8100
cggagctgag	aaagagctct	tgtccaaagc	ccggcatcat	agaatgcagc	tgtatttagg	8160
cgacctcttt	ttccatcttg	tcgatttttg	ttatgacgta	ccagttggga	tgatggatga	8220
ttgtacccca	gctgcgattg	atgtgtatct	ttgcatgcaa	caacacgcga	tgccggaggc	8280
gaactgcaca	ttggaagggt	catatatggt	cctgacatat	ctggtggatc	tggaagcatg	8340
gaattgtatt	tttgatttgg	catttgcttt	tgccgctgga	gggaacatat	caccctcggg	8400
catttttcat	ttggtaggat	ggtttggatg	cagttgtcga	cgatatcagc	ttccatatat	8460

-continued

acggaattgt tcaattgagt attacctgat ggctatttgt gtcgatcctt tgtgacattg	8520
acatgtgatg gcgcctgtac ctctaaatgt atgcagcgta acatcttcag agcacggcgg	8580
catgttctcc cctctacaat attcagataa gcagtacttg tgctgtgaga aaccattga	8640
tcgctgggtg caaccaaggc gatatgtctc caggagagga aattgcacga tcattgatcg	8700
tttgtgctac tcgtttcgtg tccaagcgaa cagagtatta atatatacca ttatatcata	8760
tgcacgtggt agtcgttact atgcagccta acctaaggta cctacaaatc ggcatcccgg	8820
agttaacgcg gaggtctctc tcgccgggcg tctccaccaa ttcgacatat caatgcta	8880
cgggcgcaaa ggctccaaac cggtatgctg tcgcattgag gttccgcatt ccttttctc	8940
cttccgaccc tctttcttac atctcactct cacatcccg	9000
gagacacgaa acggacggca acaatcaacc gtcccccg aacagcgacc tgacctccc	9060
tcccccata tacgattacg agctgaacga gacgagagag agagagagag agaagaaga	9120
gacagagaag aggcattggc gacattgagg tggcagtcac cgtacgggccc gagcggggcc	9180
cgactaagaa gcggagacgg atcgctcatc cgtgcaccga gtgccatcga aggaagcaaa	9240
aggtcaggct tctttcttac caagaaacaa agtgaagctt attaatcaac tggcctcatg	9300
ggccttcact tcaactgccg ctttccagtc gggaaacctg tcgtgccagc tgcattaaca	9360
tggctcatagc tgtttccttg cgtattgggc gctctccgct tctcgtctca ctgactcgt	9420
gcgctcggtc gttcgggtaa agcctggggg gcctaattgag caaaaggcca gcaaaggcc	9480
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc ccttgacgag	9540
catcacaaaa atcgacgctc aagtcagagg tggcgaaaac cgacaggact ataaagatac	9600
caggcgtttc cccctggaag ctccctcgtg cgctctcttg ttcgaccct gccgcttacc	9660
ggatacctgt ccgcctttct ccttcggga agcgtggcgc tttctcatag ctacacgtgt	9720
aggatatctc gttcgggtga ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc	9780
gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga	9840
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta	9900
ggcgtgtgta cagagtctct gaagtgggtg cctaactacg gctacactag aagaacagta	9960
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga	10020
tccggcaaac aaaccaccgc tggtagcggg ggtttttttg ttgcaagca gcagattacg	10080
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacgggggc tgacgctcag	10140
tggaacgaaa actcacgtta agggattttg gtcattgagat tatcaaaaag gatcttcacc	10200
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact	10260
tggctctgaca gttattatct acccaccact ttggtaattt cgcctttcac gtaatgcaca	10320
aattcttcca gctgatccgc acggctggcc agacgatctt cttcctggcc cagatacgc	10380
tgacgcgctt ccagaatcac cggtgatac tgcgcggcga gacgttccat ggcccaatcc	10440
gccgccacat ctttcggcgc aattttgccg gtcacgcgc tataccaaat acggctcagg	10500
gtcagcacca cgttgcgctt atcaccgcc caatccggcg ggctgttcca cagggtcagg	10560
gtttcgttca gcgcttcaaa caggctcgtt tccggcacgc gatcaaacag ttcttcgct	10620
gccggaccaa ccagcgccac gctatgttca cgcgctttgg tcagcagaat ggccagatcg	10680
atatcaatgg tcgccggttc aaaaatgccg gccagaatat cgttacgctg ccattcgcca	10740

-continued

```

aactgcagct cacgttttgc cggataacgc cacggaataa tgtcatcatg caccacgatg 10800
gtcacttcca cgcacgcag aatttcgctt tcaccggggc tcgcgctggt ttccagcaga 10860
tcgttaatca cgcacgcagc ggtgggtttca tccagacgca cggtcacggt caccagcaga 10920
tcaatatcgc tatgcgggtt caggccgcga tccaccgcgc tgccatacag atgcacggcc 10980
agcagggtcg gttccagatg acgttcaatc acgcccacca cttcagacag ctgggtgctc 11040
acttccgcaa tcaccgcttc acgcataatg ttcagtttgg ttttcggacg gctgccctgc 11100
tgggtcacat cgttgcgtct ccacagcatc aggcaacggc ccaccgcata acgcgcctgc 11160
tgtttgctcg cacgcggcat gctctgcacg ccaaaaaagc aatcatagca cgccataaac 11220
accgccaccg caccattacc gccgctacgt tcggtcaggg tacggctcca attacggcta 11280
cgcatactct tcttttttca atattattga agcatttata agggttattg tctcatgagc 11340
ggatacatat ttgaatgtat ttgaaaaat aaacaaatag gggttccgcg cacatttccc 11400
cgaaaagtgc cac 11413

```

```

<210> SEQ ID NO 11
<211> LENGTH: 4768
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: mpdh deletion cassette

<400> SEQUENCE: 11

```

```

aaagacatga accaggctgg agacgagata gtcaggcacc agatggcgag atgggacgag 60
gaagaacat ggacgaggcg aagaatgagg caaagtggga atgacggact gcagtgagt 120
gatggagttc aaagtacctg ggtacctaca tactgcgtgc atgggccacg ataccggcaa 180
ggaccagcga caggcgaggt acgaagcacg aaccagcaag caggaaccag catcgccgt 240
aatgcaggt gcgttctgc cttctagacg gcggtaccag ataagagacg ctgcaccgca 300
cataacggaa ggtcactcga cccgctttca agtggtctct tgcgccacgg cctgctggaa 360
gccacgcccc atgattggcc agcagagagg ctcattgccac tggacagctg tactggtacc 420
tgctgcttct cacctcatgc cctgatgcgt tggtggactg atgggatggt ctcacccgta 480
cgacgtagca cgaggcctgg accactctgg cattgacacg gcaactcaatg caacaatggc 540
atccgatggc cgctgtatgt tgtatcatgt agagcatcga gtacttcac tcgtacctgt 600
acctgcgctc tcgaatgggg accccccctg tggatctggc ggtacaagct gcatgtactc 660
cctggactcg gtccacaccc cccctgctgt gacgatggc gatcagcttc aaagtccctc 720
catcttcttc gcctgttga cccgattgag gcgccatcgt catcgggacc ccctacctga 780
cctctggcct gtaatactgt cctggctgta tgtcttctca ccacctggag cagcctttcg 840
gtggctctcg tgcttctcgt ggactctgtc aaactcatcc tcgaattatt ataaagctgg 900
ccggactccc gctcctccta cttgactctt agttcctctc catcatcaac cagctttcac 960
cacacaaaaa ctccatttc tctttctctt tctcttctc ctacaaccac cacatcgcca 1020
actcttctt ctgcaactcc aactccttca cttctcttct ctgttctact cttctccacc 1080
tacttcacct accgcaatcg agagctacct tacatcaata tggccagcac ctcttcggcg 1140
atacatactc gccaccccag cggggcgat tgtgtgtact aggtaggctc gtactatacc 1200
agcaggagag gtgctgcttg gcaatcgtgc tcagctgtta ggttgactt gtatgtgtact 1260

```

-continued

tgtaagggtgg	tcatgcagtt	gctaagggtac	ctagggagg	attcaacgag	cctgcttcc	1320
aatgtccatc	tggataggat	ggcggtg	ggggccgaag	ctgggaactc	gccaacagtc	1380
atatgtaata	gctcaagttg	atgataccgt	tttgccagga	ttaggatg	agaagcagca	1440
tgaatgtcgc	tcatccgatg	ccgcattcacc	gttgtgtcag	aaacgaccaa	gctaagcaac	1500
taaggtacct	taccgtccac	tatctcaggt	aaccagggtac	taccagctac	cctacctgcc	1560
gtgctacct	gcttagtat	taatctttcc	acctccctcc	tcaatcttct	tttccctcct	1620
ctcctctttt	ttttttcttc	ctcctcttct	tctccataac	cattcctaac	aacatcgaca	1680
ttctctccta	atcaccagcc	tcgcaaatcc	tcagggttagt	attactacta	ctacaatcat	1740
caccacgatg	ctccgccga	cgatgcggct	tctgttcgcc	tgccctcct	ctcactcgtg	1800
cccttgacga	gctaccgccg	cagactctcc	tgcgtcacca	atttttttcc	ctattttacc	1860
ctcctccctc	tctccctctc	gtttcttct	aacaaacaac	caccaccaa	atctctttgg	1920
aagctcacga	ctcacgcaag	ctcaattcgc	agatacaaat	ctagaatgaa	aaagcctgaa	1980
ctcaccgcga	cgtctgtcga	gaagttttctg	atcgaaaagt	tcgacagcgt	ctccgacctg	2040
atgcagctct	cggagggcga	agaatctcgt	gctttcagct	tcgatgtagg	agggcgtgga	2100
tatgtcctgc	gggtaaatag	ctgcgccgat	ggtttctaca	aagatcgta	tgtttatcgg	2160
cactttgcat	cggccgcgct	ccgattccg	gaagtgttg	acattgggga	attcagcgag	2220
agcctgacct	attgcatctc	ccgccgtgca	cagggtgtca	cgttgcaaga	cctgcctgaa	2280
accgaactgc	ccgctgttct	gcgaccggtc	gcggaggcca	tgatgcgat	cgctgcggcc	2340
gatcttagcc	agacgagcgg	gttcggccca	tccgaccgc	aaggaatcgg	tcaatacact	2400
acatggcgtg	atttcatatg	cgcgattgct	gatcccatg	tgtatcactg	gcaaactgtg	2460
atggacgaca	ccgtcagtcg	gtccgtcgcg	caggctctcg	atgagctgat	gctttgggcc	2520
gaggactgcc	ccgaagtcg	gcacctcgtg	cacgcggatt	tcggctccaa	caatgtcctg	2580
acggacaatg	gccgcataac	agcggtcatt	gactggagcg	aggcgatgtt	cggggattcc	2640
caatacgagg	tcgccaat	cttcttctgg	aggccgtggt	tggcttgat	ggagcagcag	2700
acgcgctact	tcgagcggag	gcatccggag	cttgacggat	cgccgcggct	ccgggcgtat	2760
atgctccgca	ttggtcttga	ccaactctat	cagagcttgg	ttgacggcaa	tttcgatgat	2820
gcagcttggg	cgcagggtcg	atgcgacgca	atcgtccgat	ccggagccgg	gactgtcggg	2880
cgtacacaaa	tcgcccgcag	aagcgcggcc	gtctggaccg	atggctgtgt	agaagtactc	2940
gccgatagtg	gaaaccgacg	ccccagcact	cgtccgagg	caaaggaata	atgcattgtc	3000
tgtgttctc	agaatgggcc	ccagaaggcc	gtcgagcatt	gtctatgaat	gcaaacaaaa	3060
atagtaaata	aatagtaatt	ctggccatga	cgaatagagc	caatctgctc	cacttgacta	3120
tccttgtagc	tgtatcgat	gtcgaacct	tgactgccca	ttcaacaat	tgtaaaggaa	3180
tatgagctac	aagttatgtc	tcacgtttgc	gtgcgagccc	gtttgtacgt	tattttgaga	3240
aagcgttgcc	atcacatgct	cacagtcact	tggcttacga	tcattgttgc	gatctttcgg	3300
taagaataca	cagagtaacg	attatacatc	catcgctttc	tatgattagg	tactcagaca	3360
acacatggga	aacaagataa	ccatcgcatg	caaggtcgat	tccaatcatg	atctggactg	3420
gggtattcca	tctaagccat	agtacccttg	cgatggctct	tctggcgctc	tagctcgcca	3480
tcactcacgg	cgctgagtgg	tgaccggcct	actgcactcg	gctttgggct	tcttttgatc	3540

-continued

```

ttttacttct actttttctc ttctacaact tgtgttaaga atgcacgaat gaacggcaag 3600
gcgctgggtt cggggttaat agagcatcgt ctcttttggt ttctattttt ttgtactct 3660
gctagggcat tccttgacgg gaacaaaagg tgctcgggat atcttcatgg gaagcttcag 3720
tctatagatg gactgagcag gcagcataac gggcaagat accacatcaa taaaagcaga 3780
cttgatcatg accttccttt tcatccatca ttgcaatgtc ctctagatg tagctctagt 3840
agcaggcaat ctaaatgtga ggtcaccatg atctcaagct gaagtttcca cagccagtag 3900
tagacttctt aacgatgcaa atgcagccac cccgtccaat acggctcggg aggaacattt 3960
cctccaatcc gtcaaccaac ccctctcagc ttccggctcg gcagetaatt aatcagccca 4020
ggccatgaca ggcgacacat tggatggaag ttccgtcaca cacatccgct aaagcatgcc 4080
ataccctttg ctagcacacc acccacacc aaccagtcac gccgttcctg cccaccaccc 4140
cctcttaatc ttgggtacgt atatggatgg atggatggac agacggctgg aaggacggat 4200
ggatggatgc atccctgaga attaccaat taggaagggc acgatgtgtg gccggctgat 4260
tggtccctga gcatttcatt cctagtgcct ttctttttcc ttcttctctt attttctcaa 4320
ggggcagcct tacgctaggt attccaatac gtaacatgtt gtctcgaggt atgtcactgc 4380
agagcaggtc aaagcaagac aaggctgagt ctacaggcac ctgcctccat tctcgtctct 4440
ccgcatagta gtagtactac ctattttaacc aaatcccga ctacgccgac catggacaac 4500
atgatgtgag acacatctgg acagaagtgg gcaatgccat ttacagcctc atgcatgcca 4560
cttcagcagc ttgagccctt gctgcagcat tatcacagca cctactggga cagtacatag 4620
gtagtaggta ggaagtatgc aggtagcact acatggacat gccccctcgt gggccccctat 4680
ttctgcgcgc atttcacgca cgacaattcc tcacacgagt tggggaacgg aaagattact 4740
gacaggccct tgttttatgc cgactcgg 4768

```

<210> SEQ ID NO 12

<211> LENGTH: 1615

<212> TYPE: DNA

<213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 12

```

ctacaaccac cacatcgcca actcttctt ctgcaactcc aactccttca cttctcttct 60
ctgttctact cttctccacc tacttcacct accgcaatca tggccaacgc aactggcaag 120
aaggcagtc actttggcgc cggcaacatc ggccgtgggt tcgttgcttg ctctctccac 180
aactcgggct atgaggtcgt ttttgccgac gtcaacggag acctgattga taccatcaac 240
tctacccct cctacaaagt catcgaggtc ggctccgagg gcacctcaga aaacgtcatc 300
accaactacc gggccatcaa ctctcggact cagaggaag acctcgttga ggagatccgc 360
actgccgacg tcgtcacctg ctccgtcggc cccaacatcc tcaagttcat cgctcccgtc 420
atcgcaaagg gcacgaccg ccgctccgac gacgagacgc ccctgcaagt cattgcctgc 480
gagaacgcca ttggcgccac cgacacgctg gccagctaca tcaaggatgc caagaacacg 540
gatctgccc gtctcgagac gcaccacctg cgtgcccgt acgccaactc ggctattgac 600
cggatcgtgc cggcgcagga ccccaatgcc ggcttgacg tgaagctgga gaagttcttt 660
gagtgggttg tcgacaagac cccctttgag gacattggca tcccgcctat tgagggcac 720
cagtgggtcg acaacctgga gcccttcac gagcgcaagc tgtacacggt caaactggc 780

```

-continued

```

cacgctacgg ctgcctacca cggtactac cgcagaaagc gcacggtcta cgatgctctc   840
caggacaagg agatcatggc tgaggtecggt ggcgccctga ccgagaccaa gaacctgctc   900
gttgccaagc acggcatcga cgaggaggcc cagctcgctt acatgaacaa gattatcaag   960
cgcatcggca acccccattt ggaggatgct gtcgagcgtg tcggccgcgc tccctgagg  1020
aagctgtcgc gcaaggagcg cttcatcggc cccgcgcgcg aggcgcgcga gaacgacatg  1080
cccgtaagt acctcctcga cgccatcgag atgacgttcc gcttcacgga cgtcgagggt  1140
gacgacgagt ccaaggagct ctctaccatc atgtccgaga acaccccgga ggacgttgct  1200
gccaagggtct gcggcatcaa gccgacggag aagatctacc ctgcctcgtt cgagggttgct  1260
aagcgtgtgc aggctgacag ccaggaggag taacttgcca tggctcttct ggcgttctag  1320
ctcggcatca ctcacggcgc tgagtgggtga ccggcctact gcactcgtgt ttgggttct  1380
tttgatcttt tacttctact ttttctcttc tacaacttgt gttaagaatg cacgaatgaa  1440
cggcaaggcg ctgggttcgc ggtaaataga gcactgtctc ttttgtttct tatttttttt  1500
gtactctgct agggcattcc ttgacgggaa caaaagggtc tcgggatatc ttcattggaa  1560
gcttcagtct atagatggac tgagcaggca gcataacggg tcaagatacc acatc   1615

```

<210> SEQ ID NO 13

<211> LENGTH: 397

<212> TYPE: PRT

<213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 13

```

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

```

-continued

Trp	Val	Asp	Asn	Leu	Glu	Pro	Phe	Ile	Glu	Arg	Lys	Leu	Tyr	Thr	Val
210						215					220				
Asn	Thr	Gly	His	Ala	Thr	Ala	Ala	Tyr	His	Gly	Tyr	Tyr	Arg	Arg	Lys
225					230					235					240
Arg	Thr	Val	Tyr	Asp	Ala	Leu	Gln	Asp	Lys	Glu	Ile	Met	Ala	Glu	Val
			245						250					255	
Arg	Gly	Ala	Leu	Thr	Glu	Thr	Lys	Asn	Leu	Leu	Val	Ala	Lys	His	Gly
		260						265					270		
Ile	Asp	Glu	Glu	Ala	Gln	Leu	Ala	Tyr	Met	Asn	Lys	Ile	Ile	Lys	Arg
	275						280					285			
Ile	Gly	Asn	Pro	His	Leu	Glu	Asp	Ala	Val	Glu	Arg	Val	Gly	Arg	Ala
	290					295					300				
Pro	Leu	Arg	Lys	Leu	Ser	Arg	Lys	Glu	Arg	Phe	Ile	Gly	Pro	Ala	Ala
305					310					315					320
Glu	Ala	Ala	Glu	Asn	Asp	Met	Pro	Val	Lys	Tyr	Leu	Leu	Asp	Ala	Ile
			325						330					335	
Glu	Met	Thr	Phe	Arg	Phe	Gln	Asp	Val	Glu	Gly	Asp	Asp	Glu	Ser	Lys
		340					345						350		
Glu	Leu	Ser	Thr	Ile	Met	Ser	Glu	Asn	Thr	Pro	Glu	Asp	Val	Val	Ala
	355						360					365			
Lys	Val	Cys	Gly	Ile	Lys	Pro	Thr	Glu	Lys	Ile	Tyr	Pro	Arg	Leu	Val
	370					375					380				
Glu	Val	Val	Lys	Arg	Val	Gln	Ala	Asp	Ser	Gln	Glu	Glu			
385					390					395					

<210> SEQ ID NO 14

<211> LENGTH: 1418

<212> TYPE: DNA

<213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 14

cagacaacta caaacaagat ggttggcggt gacgcagaga ttatcgccga tgataccaga	60
atcctgggtc aggatccttt gatcccccgc gctctgctca tctccgagat ccccatgact	120
cccaatgccc tccagaccgt cgtaagggc cggagagacg ctgccagat catcatgggc	180
cagaacgaca ggctattagt cgtcgtcggt cctgctcca tccacgccc cgaagccgcc	240
cacgagtagc ccgcccgcct caagcagctc tccgagaagc tcgccgacga cctctgcac	300
gtgatgcgcg cctacctgga gaagcctcgc acgaccgtcg gctggaaggc cctgatcaac	360
gatcccgata tcgactcgtc ctccagatc aacaagggcc tgcgcgtctc ccgccggctc	420
ttcgtcgacc tgacctccaa cggcatgccc attgccaccg agatgctcga caccatctcc	480
ccccagttcc tcgccgactg catctccgtc ggcgccattg gcgcgcgaac gaccgagtcc	540
cagctgcacc gtgagctcgc ttccggcctg tcttccccg tgggcttcaa gaacggaacc	600
gacggcagcc ttggcgtcgc cattgatgcc attggcgtg ccgcggccaa gcaccacttc	660
atgggtgtga cgaagcaggc cctggctgcc attaccgcga ccaagggcaa cgagcacggc	720
ttcgtgatcc tgagaggcgg aagcaaggcc cccaactttg acaaggagag cgtccaggcg	780
gccaagaaga tgctgacgga caagggccag aagctggcca ttatgattga ctgttctcac	840
ggcaactctc agaagaacca caacaaccag cccaaggttg ccaaggccat tgccgaccag	900
ctgcgcgagg gcgagcgagc cattgtcggc gtcatgatcg agtccaacat caacgagggc	960

-continued

```

aaccaaaagg tcctgctga ggtccggcc ggtctcaagc gcggcgtag catcacggat 1020
gcatgcatca actgggagga cacagtcacg gtcctcgagg accttgccgc tgcggtccgc 1080
acccgacgcg aggtcaactc ggtgctcct agcgagagg ccaaggtag cactctggag 1140
gaggactaga cgtagtttta ccggaagggg cgtatgtgtg tacgatgaga gagaggaacg 1200
aaaagcatga gggtcgggtt gttgattttg aacattacga tgcgtgagtt tatgcgcgcg 1260
aatctgggtt tcctcctgg ccagcgtaaa taagaaaagt tttagggaga gcattcgaca 1320
cggatatacga gaaattggcg caccaaagga acatgtatga aggcataatat tgttggtgt 1380
ccactttgtc acagacggcg cacacgaacc agtttcag 1418

```

<210> SEQ ID NO 15

<211> LENGTH: 376

<212> TYPE: PRT

<213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 15

```

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

```

-continued

275	280	285	
Lys Ala Ile Ala Asp Gln Leu Arg Glu Gly Glu Arg Ala Ile Val Gly			
290	295	300	
Val Met Ile Glu Ser Asn Ile Asn Glu Gly Asn Gln Lys Val Pro Ala			
305	310	315	320
Glu Gly Pro Ala Gly Leu Lys Arg Gly Val Ser Ile Thr Asp Ala Cys			
	325	330	335
Ile Asn Trp Glu Asp Thr Val Thr Val Leu Glu Asp Leu Ala Ala Ala			
	340	345	350
Val Arg Thr Arg Arg Glu Val Asn Ser Gly Ala Pro Ser Ala Glu Ala			
	355	360	365
Lys Val Ser Thr Leu Glu Glu Asp			
370	375		
<210> SEQ ID NO 16			
<211> LENGTH: 3267			
<212> TYPE: DNA			
<213> ORGANISM: Trichoderma reesei			
<400> SEQUENCE: 16			
atgacaacgc tgctcaacaa ccgcgcggat tcgggtcccc cgccgcgcca gattcgcttc			60
gtacataacc agggccagcc gccgtccaag cgacgcagga tcaatgcagc gtatgttctt			120
ctggggcttc tctcgcttg cttcgctctgt cttgcgcgt cacccttggg ggcaaatggc			180
tgctctagac agataagtgc aagatggcgc gagcatcaac accatgcggt taccagtga			240
gctataagta ttagagcta acctcgtctt gcgcgcagat gcttgacctg tcgaagacga			300
aagacgagat gcgcgcgcca acgaccgctt tgcaccacat gactaagaa cgggcaccaa			360
tgctcgggtt atcccagga caggagagat ggcgaggcag cggcaaggac ggcgcagcag			420
gaaacgacgc caaggagga tgcaaccaa ggcgaggacg acgtgcctc tagaaaaccc			480
gctgcttcgg ggcttgcaa ccgtccgacg tctctgggg gcagcaatct gcaagctacg			540
catgtgtctg cctctgtctg ctctgtttca cggcgcgcgc ggcgcaatat caagcagact			600
atcgaagtgc gcgcgatgc caaccgggc gacgcagagc cgcaacagtc tccgtacta			660
catcaccacc accagagcca ccaccagaat caccatcctc atcatcaagc tcatctctca			720
gtcagctcgc atgacttgcc cagttcgccc accctcgcgc gatcggtccc cagcgcgtgc			780
atcccctatt tccgctaatt tggtccacg gccatcgccc ccggtctcaa gcaaatggtg			840
gtctctgttc gcgaccgtgc ccgttcgacg gcaggtcac attccggcgc ctctcttcat			900
tcgacgcccc gcggggcgat ggccagtggc tcggccattg acagcgacat tgtccgcgaa			960
gacattccct cctacgacct caacaactcc gccccgctc atcccttaat cattcagcta			1020
gtcaataacgt tttttgttca cgtcgggtgc aattaccctt ttctcaaaca agacaagttt			1080
ctgcaacatg taaaggagaa acgggtcgag ccgatacttg tcgatgccgt ctgcgccatc			1140
gcggcgcgct tctcggaacca tcacatcctg acgggcgcca acgacaagat gccccggaca			1200
gagagaggcc aggtctttgc gcaacgggcc aggcaggcca ccgtcgacac ttcccttgc			1260
ccttctgtgg gtgcggtcca ggcttgctg ctgatggcat acgagggtt cggggcggcc			1320
caggatagcg cgctgtggat gtatctgggc ttggccattc gcatggcggg tgatctcggt			1380
ctccaaaagt gcgtgggcgt tcaatatcaa ggcgaaaagg acccttggtg cacccgcat			1440

-continued

cgaagtcggg ccaacgacga tgatcaggac agcccagagg cgcacaaggc ggatgccgtc	1500
gacgccccta gcccgaggga acagaaggag gtagagcagg agcgcatata taccttctgg	1560
tccgtcttca tcctggaccg gctcatctcc tctggcactg gacggccggt cactatccgc	1620
gacggcgact acgagctcgc ctttcccgaa tccagcgtcg acctcgtac cggatggccg	1680
aaccatttc ctgttctcct ccagatcacc caccctctac gtcaagtgtg cgatgtgctg	1740
aacaatctcc acgacgcca agaccttacc caggacaagt gggatcagct atccgagatg	1800
gagcatcgac tcacgagaat gtacaagcac tgggacccgc ggcttcagtt caacgtcaac	1860
aactttaaga cgtacctcgg catgggccag ggcaccaact ttatcctcct ccacttttgg	1920
ttccacgcgc tcttcatcat attgcatcaa ccgactcttc ttacccctt ttgcgaactg	1980
cggagcgagc tccagctgct ctccgacagc cgtgagctga gcacgagcag cgcaaagacc	2040
atctgcgaca ttttgtcctt tgccgacttg atagatccga cgagttttat tggcaaccca	2100
ttcacgaatc agcccattta tatcgccgca tgcgcatttc tcatggaatc gagcgccaac	2160
aatgcatctg agggctcgtc caggaggggc tcgcctccca ctcaagcgag tcgctccagg	2220
caacagacga agcactttaa caagcagctc cggcattcgc ttctcgcgtc ggccgccaac	2280
cagaattacc agcgatgcta caactccctg cagcagatcc aggcgtactg gggcgagta	2340
ggctatatcc tcaccgcgct ggaccagaag tcaaagggca agtgggactg cgagacgtat	2400
accgtggagg agtatgagag caccaagctg cccctcctc ggccgctgat tggtagcag	2460
ctgtcgcggt tcgagaagca gtctcgcgcc aagattgccc gctcaccgat tgcctggagt	2520
ctgtccggca cggccaattc gccaactcc agcctgacac tcatgtatca ggctctcaac	2580
gccagcgagc gcgccacac tggattccct ccgctgtgc agtccgtggc gagggctgcc	2640
aatggttcgc cggggagcat cgcgctggac cccattcgac acacgatgcc ggaagtggga	2700
atgtttacgt cggcggttcc gcagccaaac atgtctgcca ttcgaaactc tcctcggcgg	2760
ccatagtctg ccaattctat gaatgcgcg ggggcgaact cgcggaatca gaatatcatg	2820
ggcttcgaag acattccgga agagtgcgc gatgcaaagg gtctgatgac caattactca	2880
acctctggcc agccctacga tggatataat gcaccccca ccagcggcaa ttcgggcatg	2940
ggcaattcta cagccgtgac gactgcccc agtggaaca cgacgtcgaa tgcgaatacc	3000
cacaacaata acaacaacaa cagcaccacc accaccagca gcagcaacag caacagcatg	3060
tatttcgcgc aaaacatggc atcatatcaa gcctgggggc cgatgacggg caacatggac	3120
gccatcacgt tcgacagcca ggacattgac attgaggcgc tggggctgca gcagccggac	3180
ctgatgagcg ggtggctcga gtatcttccc agtgatgtgc tggggctgtt tgacgagcat	3240
ggcaggtcgg ggcagggcgg ccattga	3267

<210> SEQ ID NO 17

<211> LENGTH: 983

<212> TYPE: PRT

<213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 17

Met	Thr	Thr	Leu	Leu	Asn	Asn	Pro	Pro	Asp	Ser	Gly	Pro	Pro	Pro	Arg
1			5						10						15

Gln	Ile	Arg	Phe	Val	His	Asn	Gln	Gly	Gln	Pro	Pro	Ser	Lys	Arg	Arg
	20						25						30		

-continued

Arg	Ile	Asn	Ala	Ala	Cys	Leu	Thr	Cys	Arg	Arg	Arg	Lys	Thr	Arg	Cys
		35					40					45			
Ala	Gly	Glu	Arg	Pro	Leu	Cys	Thr	Thr	Cys	Thr	Lys	Asn	Gly	His	Gln
	50					55					60				
Cys	Leu	Gly	Tyr	Pro	Glu	Asp	Arg	Arg	Asp	Gly	Glu	Ala	Ala	Ala	Arg
65					70				75						80
Thr	Gly	Asp	Glu	Glu	Thr	Thr	Pro	Arg	Glu	Asp	Val	Asn	Gln	Gly	Glu
			85						90					95	
Asp	Asp	Val	Ala	Ser	Arg	Lys	Pro	Ala	Ala	Ser	Gly	Leu	Ala	Asn	Arg
			100					105					110		
Pro	Thr	Ser	Ser	Gly	Gly	Ser	Asn	Leu	Gln	Ala	Thr	His	Ala	Ala	Ala
		115					120					125			
Ser	Ala	Ala	Ser	Ala	Ser	Arg	Pro	Pro	Arg	Arg	Asn	Ile	Lys	Gln	Thr
	130					135					140				
Ile	Glu	Asn	His	His	Pro	His	His	Gln	Ala	His	Leu	Ser	Val	Thr	Ser
145					150					155					160
Asp	Asp	Leu	Pro	Ser	Ser	Pro	Thr	Leu	Arg	Arg	Ser	Ala	Pro	Ser	Arg
			165						170					175	
Arg	Ile	Pro	Tyr	Phe	Arg	Tyr	Phe	Gly	Ser	Thr	Ala	Ile	Val	Pro	Gly
		180						185					190		
Phe	Lys	Gln	Met	Val	Val	Ser	Val	Arg	Asp	Arg	Arg	Arg	Ser	Thr	Ala
		195					200					205			
Gly	Ser	His	Ser	Gly	Ala	Ser	Leu	His	Ser	Thr	Pro	Ser	Gly	Ala	Met
	210					215					220				
Ala	Ser	Gly	Ser	Ala	Ile	Asp	Ser	Asp	Ile	Val	Arg	Glu	Asp	Ile	Pro
225					230					235					240
Ser	Tyr	Asp	Pro	Asn	Asn	Ser	Ala	Pro	Val	His	Pro	Leu	Ile	Ile	Gln
			245						250					255	
Leu	Val	Asn	Thr	Phe	Phe	Val	His	Val	Gly	Cys	Asn	Tyr	Pro	Phe	Leu
		260						265					270		
Lys	Gln	Asp	Lys	Phe	Leu	Gln	His	Val	Lys	Glu	Lys	Arg	Val	Glu	Pro
		275					280					285			
Ile	Leu	Val	Asp	Ala	Val	Cys	Ala	Ile	Ala	Ala	Arg	Phe	Ser	Asp	His
	290					295					300				
His	Ile	Leu	Thr	Gly	Gly	Asn	Asp	Lys	Met	Pro	Arg	Thr	Glu	Arg	Gly
305					310					315					320
Gln	Val	Phe	Ala	Gln	Arg	Ala	Arg	Gln	Ala	Thr	Val	Asp	Thr	Phe	Pro
			325						330					335	
Cys	Pro	Ser	Val	Gly	Ala	Val	Gln	Ala	Cys	Leu	Leu	Met	Ala	Tyr	Glu
			340					345					350		
Gly	Phe	Gly	Ala	Ala	Gln	Asp	Ser	Ala	Leu	Trp	Met	Tyr	Leu	Gly	Leu
		355					360					365			
Ala	Ile	Arg	Met	Ala	Val	Asp	Leu	Gly	Leu	Gln	Lys	Cys	Val	Gly	Val
	370					375					380				
Gln	Tyr	Gln	Gly	Glu	Lys	Asp	Pro	Trp	Tyr	Thr	Arg	His	Arg	Ser	Arg
385					390					395					400
Ala	Asn	Asp	Asp	Asp	Gln	Asp	Ser	Pro	Glu	Ala	His	Lys	Ala	Asp	Ala
			405						410					415	
Val	Asp	Ala	Pro	Ser	Pro	Glu	Glu	Gln	Lys	Glu	Val	Glu	Gln	Glu	Arg
			420					425					430		
Ile	Tyr	Thr	Phe	Trp	Ser	Val	Phe	Ile	Leu	Asp	Arg	Leu	Ile	Ser	Ser

-continued

435					440					445					
Gly 450	Thr	Gly	Arg	Pro	Val	Thr 455	Ile	Arg	Asp	Gly 460	Asp	Tyr	Glu	Leu	Ala
Phe 465	Pro	Glu	Ser	Ser	Val 470	Asp	Leu	Ala	Thr	Gly 475	Trp	Pro	Asn	Pro	Phe 480
Pro	Val	Leu	Leu	Gln 485	Ile	Ile	His	Leu	Tyr 490	Gly	Gln	Val	Cys	Asp 495	Val
Leu	Asn	Asn	Leu	His 500	Asp	Ala	Gln	Asp 505	Leu	Thr	Gln	Asp	Lys 510	Trp	Asp
Gln	Leu	Ser 515	Glu	Met	Glu	His	Arg 520	Leu	Thr	Arg	Met	Tyr 525	Lys	His	Trp
Asp	Pro 530	Arg	Leu	Gln	Phe	Asn 535	Val	Asn	Asn	Phe	Lys 540	Thr	Tyr	Leu	Gly
Met 545	Gly	Gln	Gly	Thr	Asn 550	Phe	Ile	Leu	Leu	His 555	Phe	Trp	Phe	His	Ala 560
Leu	Phe	Ile	Ile	Leu 565	His	Gln	Pro	Thr	Leu	Leu 570	Thr	Pro	Phe	Cys 575	Glu
Leu	Arg	Ser	Glu	Leu 580	Gln	Leu	Leu	Ser 585	Asp	Ser	Arg	Glu	Leu 590	Ser	Met
Ser	Ser	Ala 595	Lys	Thr	Ile	Cys	Asp 600	Ile	Leu	Ser	Phe	Ala 605	Asp	Leu	Ile
Asp	Pro 610	Thr	Ser	Phe	Ile	Gly 615	Asn	Pro	Phe	Thr	Asn 620	Gln	Pro	Ile	Tyr
Ile 625	Ala	Ala	Cys	Ala	Phe 630	Leu	Met	Glu	Ser	Ser 635	Ala	Asn	Asn	Ala	Ser 640
Glu	Gly	Ser	Ser	Arg 645	Glu	Gly	Ser	Pro	Pro 650	Thr	Gln	Ala	Ser	Arg 655	Ser
Arg	Gln	Gln	Thr 660	Lys	His	Phe	Asn	Lys 665	Gln	Ser	Arg	His	Ser 670	Leu	Leu
Ala	Ser	Ala 675	Ala	Asn	Gln	Asn	Tyr 680	Gln	Arg	Cys	Tyr	Asn 685	Ser	Leu	Gln
Gln	Ile 690	Gln	Ala	Tyr	Trp	Gly 695	Gly	Val	Gly	Tyr	Ile 700	Leu	Thr	Ala	Leu
Asp 705	Gln	Lys	Ser	Lys	Gly 710	Lys	Trp	Asp	Cys	Glu 715	Thr	Tyr	Thr	Val	Glu 720
Glu	Tyr	Glu	Ser	Thr 725	Lys	Leu	Pro	Pro	Pro 730	Arg	Pro	Ser	Ile	Gly 735	Glu
Gln	Leu	Ser	Arg 740	Phe	Glu	Lys	Gln	Ser 745	Ser	Pro	Lys	Ile 750	Ala	Gly	Ser
Pro	Ile	Ala 755	Trp	Ser	Leu	Ser	Gly 760	Thr	Ala	Asn	Ser	Pro 765	Asn	Ser	Ser
Leu	Thr 770	Leu	Met	Tyr	Gln	Ala 775	Leu	Asn	Ala	Ser	Ala 780	Gly	Gly	His	Thr
Gly 785	Phe	Pro	Pro	Pro	Val 790	Gln	Ser	Val	Ala	Arg 795	Ala	Ala	Asn	Gly	Ser 800
Pro	Gly	Ser	Ile	Ala 805	Leu	Asp	Pro	Ile	Arg 810	His	Thr	Met	Pro	Glu	Val 815
Gly	Met	Phe	Thr 820	Ser	Ala	Val	Pro	Gln 825	Pro	Asn	Met	Ser	Ala 830	Ile	Arg
Asn	Ser	Pro 835	Arg	Arg	Pro	Tyr	Ala 840	Ala	Asn	Ser	Met	Asn 845	Ala	Ala	Gly

-continued

Ala Asn Ser Arg Asn Gln Asn Ile Met Gly Phe Glu Asp Ile Pro Glu
850 855 860

Glu Ser Pro Asp Ala Lys Gly Leu Met Ser Asn Tyr Ser Thr Ser Gly
865 870 875 880

Gln Pro Tyr Asp Gly Tyr Asn Ala Ser Pro Thr Ser Gly Asn Ser Gly
885 890 895

Met Gly Asn Ser Thr Ala Val Thr Ser Ala Pro Ser Gly Asn Thr Thr
900 905 910

Asn Ser Met Tyr Phe Ala Gln Asn Met Ala Ser Tyr Gln Ala Trp Gly
915 920 925

Pro Met Thr Gly Asn Met Asp Ala Ile Thr Phe Asp Ser Gln Asp Ile
930 935 940

Asp Ile Glu Ala Leu Gly Leu Gln Gln Pro Asp Leu Met Ser Gly Trp
945 950 955 960

Leu Glu Tyr Leu Pro Ser Asp Val Leu Gly Leu Phe Asp Glu His Gly
965 970 975

Glu Ser Gly Gln Gly Gly His
980

<210> SEQ ID NO 18

<211> LENGTH: 1773

<212> TYPE: DNA

<213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 18

```

atgtcaacca gacacatttt tacgtcatgt gacggcctcg tcaacaaggc catccgaggc    60
atcattacct acaacccttc tctcagccta gacgaaagca accgggtcgt ctatgacacc    120
acatacgaca agtccaaagt cgtcatcata agcggaggcg gctccggaca tgagcctgcg    180
tggacaggct atgtcgggtca aaacatgctc gcagcttcgg ttcaaggaga catctttgcc    240
agccccagca caaagcaaat ccttgagca gtggacgctg ttccgagtga caagggcacc    300
attctgggtca tcacaaaacta cactggtgac tgtctgcact ttgggctggc caacgagaaa    360
gccaatgcca aaggccacaa ctgccggacc atcatctgcg gcgacgatgt ctctgtcggc    420
aagaacggga gcatggttgg tcgcagaggt ttggctggac agatgggtgt cctcaaagtc    480
atgggctggg ctgctggtgc tggaggctcg ctggacgaag tttacaatct tgggtgcgca    540
ttctcgcagc agatttgtgc tattgcggca accttggatc attgccatgt tccaggacgt    600
gctgaacatg gcatgcttca agaaggcgtt gttgaactcg gcaccggccc tcataacgaa    660
cccgataaca agaaactatc cctgcgcca totgcgtctg atctcgtctc ccagattcta    720
caatactgtc tgaatgaaaa ggatccggaa agaggtacg ttcacttcgc ccctagtgat    780
gaagtcatac tacttctcag caactacggc ggcataatcg acctggagat gggcgctctg    840
gttgacgaaa ttctcgagca gcttgcaaga gactggaaga ttgagccggt gagggtatgc    900
gccggcttca tcgagacgtc actcaatgct ccagcctttt ctgtctccgt catcaatgtg    960
actgtcgcgc caaagaattg cagttactct gtccaagaga tcaagtcctt ctttgatgcc   1020
aggacgacga cgtttttggga gtccatgact gggtcgcaga cacggcggag gacgagacag   1080
cagcagattg ttcagccccc ggaggaagag gtcaaggcag ttgatgagtc aaaagatctg   1140
actattgatc ccacgctctt agagaagatg ttgaggaacg cttgtgaagc gctcgtcaag   1200

```

-continued

```

gctgagccag atctcaccaa gtgggatacc gtcattgggtg atggagattg cggagagacg 1260
ctcaggactg gagccacatc ttgtctcacg gcccttgatt cgggattggc gaaatctggg 1320
tcagtgtgta gactactatt ggaactcgaa gacatcgtgg agagcaagat gggcgggtact 1380
ctgggcggca tattgggcat cttttttgtc tccttgctgtg cagctgtcga gaagaatatt 1440
ggactggccc aaagcccggt ctgtctctgg ggaacagctc ttgcaacagc tctggagagt 1500
ctcaggaggt atacgcctgc caaggctcga gatcgtactg tcatggacac actcatcccg 1560
tttgcctacg tcatgtgtga gtcgggcttt gataacggcg tcaaggcggc acaggatgga 1620
gcggatgcga ccaagacgat gaagccgaga ctgggaagag caacgtatgt tggcgtgtcg 1680
agtgagaaca aggagttgcc tctgatccc ggtgcttggg gggccatggt ggccattcgc 1740
ggtttgcgac aaggtctggt tgaagcgatc tag 1773

```

<210> SEQ ID NO 19

<211> LENGTH: 590

<212> TYPE: PRT

<213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 19

```

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

```

-continued

Ala	Pro	Ser	Asp	Glu	Val	Ile	Leu	Leu	Leu	Ser	Asn	Tyr	Gly	Gly	Ile	
			260					265					270			
Ser	His	Leu	Glu	Met	Gly	Ala	Leu	Val	Asp	Glu	Ile	Leu	Glu	Gln	Leu	
		275					280					285				
Ala	Arg	Asp	Trp	Lys	Ile	Glu	Pro	Val	Arg	Val	Cys	Ala	Gly	Phe	Ile	
	290				295						300					
Glu	Thr	Ser	Leu	Asn	Ala	Pro	Ala	Phe	Ser	Val	Ser	Val	Ile	Asn	Val	
305				310					315						320	
Thr	Ala	Ala	Ala	Lys	Asn	Cys	Ser	Tyr	Ser	Val	Glu	Glu	Ile	Lys	Ser	
				325				330						335		
Phe	Phe	Asp	Ala	Arg	Thr	Thr	Thr	Phe	Trp	Glu	Ser	Met	Thr	Gly	Ser	
		340						345					350			
Gln	Thr	Arg	Arg	Arg	Thr	Arg	Gln	Gln	Gln	Ile	Val	Gln	Pro	Pro	Glu	
		355					360					365				
Glu	Glu	Val	Lys	Ala	Val	Asp	Glu	Ser	Lys	Asp	Leu	Thr	Ile	Asp	Pro	
	370					375					380					
Thr	Leu	Leu	Glu	Lys	Met	Leu	Arg	Asn	Ala	Cys	Glu	Ala	Leu	Val	Lys	
385				390						395					400	
Ala	Glu	Pro	Asp	Leu	Thr	Lys	Trp	Asp	Thr	Val	Met	Gly	Asp	Gly	Asp	
			405						410					415		
Cys	Gly	Glu	Thr	Leu	Arg	Thr	Gly	Ala	Thr	Ser	Leu	Leu	Thr	Ala	Leu	
			420				425						430			
Asp	Ser	Gly	Leu	Ala	Lys	Ser	Gly	Ser	Val	Val	Arg	Val	Leu	Leu	Glu	
		435					440					445				
Leu	Glu	Asp	Ile	Val	Glu	Ser	Lys	Met	Gly	Gly	Thr	Leu	Gly	Gly	Ile	
	450					455					460					
Leu	Gly	Ile	Phe	Phe	Val	Ser	Leu	Arg	Ala	Ala	Val	Glu	Lys	Asn	Ile	
465				470					475						480	
Gly	Leu	Ala	Gln	Ser	Pro	Val	Leu	Leu	Trp	Gly	Thr	Ala	Leu	Ala	Thr	
			485					490						495		
Ala	Leu	Glu	Ser	Leu	Arg	Arg	Tyr	Thr	Pro	Ala	Lys	Val	Gly	Asp	Arg	
		500					505						510			
Thr	Val	Met	Asp	Thr	Leu	Ile	Pro	Phe	Ala	Thr	Val	Met	Cys	Glu	Ser	
		515				520						525				
Gly	Phe	Asp	Asn	Gly	Val	Lys	Ala	Ala	Gln	Asp	Gly	Ala	Asp	Ala	Thr	
	530				535						540					
Lys	Thr	Met	Lys	Pro	Arg	Leu	Gly	Arg	Ala	Thr	Tyr	Val	Gly	Val	Ser	
545				550					555						560	
Ser	Glu	Asn	Lys	Glu	Leu	Pro	Pro	Asp	Pro	Gly	Ala	Trp	Gly	Ala	Met	
			565					570						575		
Val	Ala	Ile	Arg	Gly	Leu	Arg	Gln	Gly	Leu	Val	Glu	Ala	Ile			
			580				585						590			

<210> SEQ ID NO 20

<211> LENGTH: 1326

<212> TYPE: DNA

<213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 20

ttagtgtagg	gtgtaaccgc	catcaacgcg	aatatcgga	ccggtcacat	agcggctggc	60
atcagacagc	aggaaagcca	ctggggcccat	gagatcctga	ggcagaccca	tcttgccctg	120
aggaatgagg	gacacccatt	tcttcttcag	gtcaggcttt	tcattcagga	ttttctcggt	180

-continued

```

cctagttaag gatttgagtt agccggccga ggtaacggct accgttcaag aacagagaag   240
agtgggttca actcaciaag ccgtcagcat gtatcccggc gagatacagt tgactcggat   300
gccagcatgc gcccaactcaa cgccgagcga agcccgagaga tggcggactg cggccttgga   360
ggcggttatag ggctagtcc agttcaaccc gttagctgtc ttatgctctg ctcacccatg   420
tctccttttt ttctttcttt cttttttccg tcgtggaagg gactcacgct tgaggctggg   480
ggacgttgac aatggccccc gacatgctgc ctatcatgac catgctacca gggctcttcc   540
tctccatgag gtgacgagca accgatgtag cgaagagata tgtgccgtcg acgttgacgc   600
cccagagctt gcgaactcta tcaatggggg agttgggtgg ctcgaagtcc tcggtgaagc   660
cagccgaggt gaccaggttg tcaatcttgc catgttgctg gatgatctcg gcgatacacg   720
cctcaacgga atcggcatct gagacgtcgg cgtaaatgcgc cgtcactttg ggaacgctag   780
accgggtgtc agccgtctgc ctgcaaagcc gtgttacgtg cagagcggag cggcagctca   840
ctcaccgttt tgcggtgggg ttctccttga gaaaggcatc cagcagcagt ttctgtctgt   900
tctcggcctc ctctctgtgc gctgcgtcag ctgtctggcc tgggggtacga gcgacccgga   960
actggcacgc ggaatctcac tgttcatgtc gacgagggcc aggtcggcgc cagagaagac  1020
cattccctgg ccaatcacca gtcccagccc tcttgccgcg ccggtgacga caccgacgtt  1080
tccatcgagg gagaagccgg caagcgtagg cttgacatac tggccgccgg tgccagagac  1140
gggctcgtg ctcggaagct catgctccgg aggatactcg aactcgacct cggcgtcggg  1200
tcgagagtgg ctgccgggct taccctgccg cggcggcgcc acactatgga acttcctcgc  1260
agcggccagc ggcaggcgtc tggcggggcg ggaggcaaag gttcggcgag ccgcaagccg  1320
cagcat                                           1326

```

<210> SEQ ID NO 21

<211> LENGTH: 334

<212> TYPE: PRT

<213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 21

```

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

```

-continued

Ser Asp Ala Asp Ser Val Glu Ala Cys Ile Ala Glu Ile Ile Gln Gln
145 150 155 160

His Gly Lys Ile Asp Asn Leu Val Thr Ser Ala Gly Phe Thr Glu Asn
165 170 175

Phe Glu Ala Thr Asn Tyr Pro Ile Asp Arg Val Arg Lys Leu Trp Gly
180 185 190

Val Asn Val Asp Gly Thr Tyr Leu Phe Ala Thr Ser Val Ala Arg His
195 200 205

Leu Met Glu Arg Lys Ser Pro Gly Ser Met Val Met Ile Gly Ser Met
210 215 220

Ser Gly Ala Ile Val Asn Val Pro Gln Pro Gln Ala Pro Tyr Asn Ala
225 230 235 240

Ser Lys Ala Ala Val Arg His Leu Ala Ala Ser Leu Ala Val Glu Trp
245 250 255

Ala His Ala Gly Ile Arg Val Asn Cys Ile Ser Pro Gly Tyr Met Leu
260 265 270

Thr Ala Leu Thr Glu Lys Ile Leu Asp Glu Lys Pro Asp Leu Lys Lys
275 280 285

Lys Trp Val Ser Leu Ile Pro Gln Gly Lys Met Gly Leu Pro Gln Asp
290 295 300

Leu Met Gly Pro Val Ala Phe Leu Leu Ser Asp Ala Ser Arg Tyr Val
305 310 315 320

Thr Gly Ala Asp Ile Arg Val Asp Gly Gly Tyr Thr Leu Thr
325 330

<210> SEQ ID NO 22
 <211> LENGTH: 640
 <212> TYPE: DNA
 <213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 22

```

tcatgcctca agatgcccaa tctcttcac cttcttcgca ctgcactcg atggatcaaa    60
ctcatacccc agccactcag tagccaacct cctcgcaact tccagcccaa caaccctctc    120
ccccatacac aacacctgcg cattattgct ctttacctg cgctcaacgc taaagctatc    180
atgtgccgta accgctcgta tgcccttgac cttgttgga gaaatggcga cgcctagacc    240
tgtgccgcag atgagcagag cgcggctcgc tttgcccctg gcgacgaggg gggcgggcgg    300
ggcgggcgcg tgggataggt gtgtcttggt ggagggtcgt tcggatccta cgtcgatgac    360
ttggtctacg cgggcgtcgc tttcaaagtc ttttttgatg gtgtttttgt agctgacgcc    420
ggcggtcctgc ggaagcgtga gtgagctcgt gtcagtgtgt gcccggttctg atgacctaga    480
ttatgcaatc ttgacgatga agagaccgat agtcatgaag acttatgaag ttgacacttg    540
atctcgttgt gatatgatgt gtgctgggtc tggcagccga aacagaagca cttacatcgc    600
agcccatggc gattctccat ttccgtgcca tcgtatatgt    640

```

<210> SEQ ID NO 23
 <211> LENGTH: 156
 <212> TYPE: PRT
 <213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 23

```

Thr Tyr Thr Met Ala Pro Lys Trp Arg Ile Ala Met Gly Cys Asp Asp
1           5           10          15

```

-continued

Ala Gly Val Ser Tyr Lys Asn Thr Ile Lys Lys Asp Phe Glu Ser Asp
 20 25 30

Ala Arg Val Asp Gln Val Ile Asp Val Gly Ser Asp Glu Pro Ser Asp
 35 40 45

Lys Thr Pro Tyr Pro Gln Arg Ala Ala Ala Ala Arg Leu Val Ala
 50 55 60

Glu Gly Lys Ala Asp Arg Ala Leu Leu Ile Cys Gly Thr Gly Leu Gly
 65 70 75 80

Val Ala Ile Ser Ala Asn Lys Val Lys Gly Ile Arg Ala Val Thr Ala
 85 90 95

His Asp Ser Phe Ser Val Glu Arg Ser Val Lys Ser Asn Asn Ala Gln
 100 105 110

Val Leu Cys Met Gly Glu Arg Val Val Gly Leu Glu Val Ala Arg Arg
 115 120 125

Leu Ala Thr Glu Trp Leu Gly Tyr Glu Phe Asp Pro Ser Ser Ala Ser
 130 135 140

Ala Lys Lys Val Glu Glu Ile Gly His Leu Glu Ala
 145 150 155

<210> SEQ ID NO 24
 <211> LENGTH: 726
 <212> TYPE: DNA
 <213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 24

```

tcacgcttca gctacttctt gtatcgtctt gacaaactgc tccggatcgt gtccgaaccg      60
ccccaggaac aatccgtcca gtccgctctt caacttctcg tagagccccc gcccggcact      120
ccccccatac aaaacctctg tcgtccccgc cctcctccca acggactcca acctctgat      180
cccagcaaca acatcgagga tgtgcttctc tccagcaggt tgactggcac caatcgccca      240
tacaggctca tacgcaaaa tgacctctgc ctggcgggc acagcattca tgacagcatc      300
aacctgcacc cgacactctt caaccgcgat atgcacgtcg ccttgagtct tttctccaat      360
gcacacaagc ggtatcatgc cattgcgact tacagcgcca gccttattcg caaccatggt      420
gtcgctctcg ccaaatatgc tctgcgctc tgcattgtcg acttccacga gcttgacgcc      480
tgcgtgcgc aacacggagg ggtgacttc gccgtgaat gcgccttggt ctteccagtg      540
gcagctctga gcgccgtcc aaatggcaac ggagctggac ttaagctggg cgattgtggt      600
ggtcagggag acgaagtcgg ggatgatgaa gatgtcaatg tgggagagga gctctatcgg      660
gacgtctgcc agcttgata agaagctgtt tgtgaagtct cttgtttgct ggagagaaaa      720
gtacat                                           726

```

<210> SEQ ID NO 25
 <211> LENGTH: 241
 <212> TYPE: PRT
 <213> ORGANISM: Trichoderma reesei

<400> SEQUENCE: 25

Met Tyr Phe Ser Leu Gln Gln Thr Arg Asp Phe Thr Asn Ser Phe Leu
 1 5 10 15

Ser Lys Leu Ala Asp Val Pro Ile Glu Leu Leu Ser His Ile Asp Ile
 20 25 30

-continued

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

1. A genetically modified saprotroph comprising at least one gene encoding at least one membrane-bound alditol transporter, at least one gene encoding at least one erythrose reductase and at least one inactivated gene encoding man- nitol 1-phosphate 5-dehydrogenase.

2. The genetically modified saprotroph according to claim 1, wherein the saprotroph is a filamentous fungus selected from the genera *Hypocrea*, *Gibberella*, *Aspergillus* and *Penicillium*.

3. The genetically modified saprotroph according to claim 1, wherein the saprotroph is *Hypocrea jecorina* (*Tricho- derma reesei*).

4. The genetically modified saprotroph according to claim 1, wherein the at least one gene encoding at least one membrane-bound alditol transporter is *fps1*.

5. The genetically modified saprotroph according to claim 1, wherein the at least one gene encoding at least one erythrose reductase is *err1*.

6. The genetically modified saprotroph according to claim 1, wherein the at least one inactivated gene encoding man- nitol 1-phosphate 5-dehydrogenase is *mpdh*.

7. The genetically modified saprotroph according to claim 1, wherein the saprotroph further comprises at least one gene encoding at least one transketolase, in particular *tkl1*.

8. The genetically modified saprotroph according to claim 1, wherein the saprotroph further comprises at least one gene encoding at least one transaldolase, in particular *tal1*.

9. The genetically modified saprotroph according to claim 1, wherein the saprotroph further comprises at least one inactivated gene encoding phospho-2-dehydro-3-deoxyhep- tonate aldolase 1 (*Dhaps-1*).

10. The genetically modified saprotroph according to claim 1, wherein the at least one gene encoding at least one membrane-bound alditol transporter, the at least one gene encoding at least one erythrose reductase, the at least one gene encoding at least one transketolase and/or the at least one gene encoding at least one transaldolase is overex- pressed.

11. A method for the production of erythritol, comprising the steps:

- a) providing at least one genetically modified saprotroph according to claim 1,
- b) culturing the at least one genetically modified sapro- troph provided in step a) in the presence a culture medium, so as to obtain erythritol in the culture medium,
- c) recovering erythritol from the culture medium.

12. The method according to claim 11, wherein the culture medium comprises lignocellulosic biomass, in particular straw, and/or at least one residue of dairy product produc- tion, in particular whey.

13. The method according to claim 11, wherein the culture medium comprises nitrate as a nitrogen source.

14. The method according to claim 11, wherein the culture medium comprises glycerol.

15. The method according to claim **11**, wherein step b) is conducted until a concentration of erythritol in the culture medium of at least 1 g/L, preferably at least 1.5 g/L, preferably at least 2 g/L, preferably at least 2.5 g/L, preferably at least 3 g/L, is reached.

16. The method according to claim **11**, wherein in step b) at least one substrate that causes osmotic stress is added to the culture medium after 14 hours of cultivation.

17. The method according to claim **16**, wherein the substrate that causes osmotic stress is glycerol and/or sodium chloride.

18. Use of a genetically modified saprotroph according to claim **11** for the production of erythritol.

* * * * *