



(43) International Publication Date
27 September 2012 (27.09.2012)

- (51) International Patent Classification:
C07K 14/43 (2006.01)
- (21) International Application Number:
PCT/EP2012/055123
- (22) International Filing Date:
22 March 2012 (22.03.2012)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
11159299.4 23 March 2011 (23.03.2011) EP
- (71) Applicant (for all designated States except US): **NO-VOZYMES A/S** [DK/DK]; Krogshoejvej 36, DK-2880 Bagsvaerd (DK).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): **BREÜNER, Anne** [DK/DK]; Strandvejen 503, DK-2930 Klampenborg (DK).
- (74) Common Representative: **NOVOZYMES A/S**; Krogshoejvej 36, DK-2880 Bagsvaerd (DK).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
— with sequence listing part of description (Rule 5.2(a))

(54) Title: SWEET-TASTING POLYPEPTIDE FROM GRAM-POSITIVE BACTERIA

(57) Abstract: The present invention relates to a Gram-positive host cell transformed with a polynucleotide encoding a sweet-tasting Brazzein polypeptide as well as methods of producing the Brazzein polypeptide using said host cell.



WO 2012/127002 A1

TITLE: SWEET-TASTING POLYPEPTIDE FROM GRAM-POSITIVE BACTERIA**Reference to sequence listing**

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

FIELD OF THE INVENTION

The present invention relates to recombinant Gram-positive bacteria that efficiently
produces and secretes a heterologous sweet-tasting polypeptide as well as methods for
10 producing said polypeptide using said bacteria.

BACKGROUND OF THE INVENTION

Due to increasing attention to the negative health effects of obesity, a market demand
for food and beverage products having alternative nutritional characteristics, including, for
15 example, reduced calorie content, has increased as well. There is a market demand to replace
the high-calorie sweeteners typically used in food and beverage products, such as sucrose and
high fructose corn syrup (HFCS), with non-nutritive sweeteners. A number of such non-nutritive
sweeteners have been identified. Some of these are small proteins which are naturally found in
plants. Such proteins are often referred to as sweet proteins.

20 Brazzein is a sweet protein which can be extracted from the fruit of the West African
climbing plant *Pentadiplandra brazzeana* Baillon (WO9531547). It has been characterized as a
monomer protein having a 3-dimensional structure with four evenly spaced disulfide bonds.
Three forms of the protein seem to exist in nature differing only at the N-terminal amino acid
residue. One corresponds to the predicted 54-amino acid translation product containing a
25 glutamine at its N-terminus. This form has been shown to be short lived as the N-terminal
glutamine undergoes natural conversion to pyroglutamate, resulting in the second form. The
loss of the N-terminal glutamine or pyroglutamate yields the 53-amino acid form which has been
reported to be twice as sweet as the form having an N-terminal pyroglutamate (Lamphear, Barry
J. *et al.* (2005) *Plant Biotechnology Journal* 3(1): 103-114).

30 On weight basis, Brazzein is 500 to 2,000 times sweeter than sugar. Its sweet
perception is quite similar to that of sucrose with a clean sweet taste with lingering aftertaste.
Further, it has been shown to be stable over a broad pH range (2.5 to 8) and it can withstand
heat which makes it suitable for many industrial food manufacturing processes. As a protein it is
safe for diabetics and very soluble in water (>50 mg/ml). Brazzein thus represents a very good
35 alternative to other available low calorie sweeteners.

Natural sourcing from *P. brazzeana* is difficult and expensive, though, and therefore
alternatives for sourcing of the protein are being searched for. Brazzein can be chemically

synthesised, which is useful for production of, e.g., variants in small scale, but not suitable for large scale production. Heterologous production in *E. coli* of Brazzein or variants of Brazzein has been described (see, e.g., Assadi-Porter, F. M. *et al.* (2000) *Arch. Biochem. Biophys.* **376**(2): 252-258; WO2008112475). The bacterial system is ideal for structural investigations because it is easy to manipulate genetically and well suited for isotopic labelling. Expression of plant proteins from *E. coli* may be performed using the advantage of forming inclusion bodies inside the cell that can quite easily be separated from host proteins. However, proteins unfold and precipitate in inclusion bodies, and controlling correct S-S bond formation after resolubilization and refolding is not always easy (see, e.g., Tamás and Shewry (2006) *Journal of Cereal Science* **43**: 259-274). Also, heterologous protein expression using inclusion bodies normally requires a number of purification steps which make such systems less suitable for industrial scale production.

Genetic engineering into plants, such as maize or wheat, has also been described. Brazzein-containing germ flour from maize has been demonstrated useful as a low-intensity sweetener providing a low-calorie alternative to sucrose, which also gives the intrinsic bulking properties necessary to replace the volume lost on removal of sugar. Also, a high-intensity sweetener based on corn-expressed Brazzein could potentially be provided from enrichment of such material, which would extend the range of product applications (Lamphear, Barry J. *et al.* (2005) *Plant Biotechnology Journal* **3**(1): 103-114).

Commercial-scale production of purified Brazzein to be used as high-intensity sweetener may be more economic from a microbial system, though. But, in general, industrial production of plant proteins from microbial systems is not straightforward, and obtaining a correctly folded plant protein in high yield from a microorganism may be challenging and is not always possible. One factor which may complicate heterologous expression of Brazzein from a microbial system is the small size of the protein. Another factor is the existence of numerous cysteine residues because of, e.g., the possibility of non-native S-S bond formation, possibly leading to loss of function or to intermolecular aggregation, e.g., during secretion. Thaumatin, which is another sweet protein originating from a plant, has been successfully expressed intracellularly in yeast in considerable quantities; however the yeast cells lacked the ability to process this molecule into a functional protein having sweet taste (Lee, J.-H. *et al.* (1988) *Biochemistry* **27**: 5101-5107).

Plasmid-carried Brazzein-encoding genes have been expressed in *Lactococcus lactis* and *Lactococcus reuteri* to produce sweet-tasting lactic acid bacteria. (Berlec A. *et al.*, 2006, *Appl Microbiol Biotechnol* **73**:158-165; Berlec A. *et al.*, 2008, *Lett Appl Microbiol* **46**: 227-231; Berlec and Strukelj, 2009, *Lett Appl Microbiol* **48**(6): 750-755; as well as published Korean patent application KR 2011135728).

Brazzein has been expressed both with and without a signal peptide in *E. coli* but only the form without the signal peptide was sweet. A single amino-acid substitution (His31→Ala)

mutant and non-secreted form of Brazzein was expressed without a signal peptide in *Bacillus subtilis*, it was necessary to break the cells by sonication to release the Brazzein (Zhang Y. *et al.*, (2009) J Fudan University (Natural Science) 48(1): 135-141).

Brazzein has been expressed in a filamentous fungal *Aspergillus* host cell (WO 2011/015633; Novozymes A/S). However, attempts to express genetically optimized synthetic Brazzein in yeast were unsuccessful (Guan, Z., Hellekant, G., and Yan, W. 1995. Chem. Senses 20, 701).

The full-length amino acid sequence of Brazzein is available from Uniprot Accession number: P56552.

One purpose for the present inventors has been to identify a method for production of Brazzein to be used as a high-intensity sweetener which can be performed in industrial scale. Industrial scale production at low cost requires that the protein is expressed in high yield and can easily be purified in a functionally folded 3-dimensional form possessing at least the same sweetness as the natural product.

New methods for efficient improved production of Brazzein remain desirable.

SUMMARY OF THE INVENTION

The invention provides, in a first aspect, a Gram-positive host cell transformed with a polynucleotide encoding a sweet-tasting polypeptide, selected from the group consisting of:

- (a) a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the mature polypeptide of SEQ ID NO:48;
- (b) a polypeptide encoded by a polynucleotide that hybridizes under medium stringency conditions, medium-high stringency conditions, high stringency conditions, or very high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO:47, or (ii) the full-length complement of (i);
- (c) a polypeptide encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the mature polypeptide coding sequence of SEQ ID NO:47;
- (d) a variant of the mature polypeptide of SEQ ID NO:48 comprising a substitution, deletion, and/or insertion at one or more positions; and
- (e) fragment of the polypeptide of (a), (b), (c) or (d) that is sweet-tasting.

In a second aspect, the invention relates to a method of producing a sweet tasting polypeptide, comprising:

- (a) cultivating the host cell of the first aspect under conditions conducive for production of the polypeptide; and
- (b) recovering the polypeptide.

5 DEFINITIONS

Coding sequence: The term "coding sequence" means a polynucleotide, which directly specifies the amino acid sequence of a polypeptide. The boundaries of the coding sequence are generally determined by an open reading frame, which begins with a start codon such as ATG, GTG, or TTG and ends with a stop codon such as TAA, TAG, or TGA. The coding sequence
10 may be a genomic DNA, cDNA, synthetic DNA, or a combination thereof.

Control sequences: The term "control sequences" means nucleic acid sequences necessary for expression of a polynucleotide encoding a mature polypeptide of the present invention. Each control sequence may be native (*i.e.*, from the same gene) or foreign (*i.e.*, from a different gene) to the polynucleotide encoding the polypeptide or native or foreign to each
15 other. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the
20 coding region of the polynucleotide encoding a polypeptide.

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

Expression vector: The term "expression vector" means a linear or circular DNA
25 molecule that comprises a polynucleotide encoding a polypeptide and is operably linked to control sequences that provide for its expression.

Fragment: The term "fragment" means a polypeptide having one or more (*e.g.*, several) amino acids deleted from the amino and/or carboxyl terminus of a mature polypeptide; wherein the fragment is sweet tasting.

PrsA protein: The PrsA protein of the secretion machinery of *B. subtilis* was disclosed
30 in Kontinen, V.P. and Sarvas, M. (1988, J. Gen. Microbiol., 134:2333-2344) and Kontinen, V.P., et al. (1991, Mol. Microbiol. 5:1273 1283) as well as in WO 94/019471 (Novozymes A/S). The *prsA* gene, which encodes the PrsA protein, was initially defined by nonlethal mutations that decreased the secretion of several exoproteins (Kontinen, V.P. and Sarvas, M., (1988) J. Gen.
35 Microbiol., 134:2333-2344). Based on the DNA sequence of the cloned *prsA* gene and subsequent work with this gene and protein, it was asserted that *prsA* encodes a protein (PrsA) that acts as a chaperone, and is translocated across the cytoplasmic membrane. The PrsA

protein has been found to possess a limited amount of sequence homology (about 30%) with the PrtM protein of *Lactococcus lactis*, a protein proposed to assist the maturation of an exported serine protease (Haandrikman, A.J., et al, (1989) *J. Bacteriol.*, 171:2789-2794; Vos, P., et al., (1989) *J. Bacteriol.*, 171:2795-2802).

Host cell: The term "host cell" means any cell type that is susceptible to transformation, transfection, transduction, or the like with a nucleic acid construct or expression vector comprising a polynucleotide of the present invention. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication.

Isolated or purified: The term "isolated" or "purified" means a polypeptide or polynucleotide that is removed from at least one component with which it is naturally associated. For example, a polypeptide may be at least 1% pure, *e.g.*, at least 5% pure, at least 10% pure, at least 20% pure, at least 40% pure, at least 60% pure, at least 80% pure, at least 90% pure, or at least 95% pure, as determined by SDS-PAGE, and a polynucleotide may be at least 1% pure, *e.g.*, at least 5% pure, at least 10% pure, at least 20% pure, at least 40% pure, at least 60% pure, at least 80% pure, at least 90% pure, or at least 95% pure, as determined by agarose electrophoresis.

Mature polypeptide: The term "mature polypeptide" means a polypeptide in its final form following translation and any post-translational modifications, such as N-terminal processing. In one aspect, the mature polypeptide is the 1-53 of SEQ ID NO:48. It is known in the art that a host cell may produce a mixture of two or more different mature polypeptides (*i.e.*, with a different C-terminal and/or N-terminal amino acid) expressed by the same polynucleotide.

Mature polypeptide coding sequence: The term "mature polypeptide coding sequence" means a polynucleotide that encodes a mature sweet-tasting polypeptide.

Nucleic acid construct: The term "nucleic acid construct" means a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or is modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature or which is synthetic, which comprises one or more control sequences.

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

Sequence identity: The relatedness between two amino acid sequences or between two nucleotide sequences is described by the parameter "sequence identity".

For purposes of the present invention, the sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *J. Mol. Biol.* 48: 443-453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000,

Trends Genet. 16: 276-277), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix. The output of Needle labeled "longest identity" (obtained using the *–nobrief* option) is used as the percent identity and is calculated as follows:

5 (Identical Residues x 100)/(Length of Alignment – Total Number of Gaps in Alignment)

For purposes of the present invention, the sequence identity between two deoxyribonucleotide sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *supra*) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *supra*), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled "longest identity" (obtained using the *–nobrief* option) is used as the percent identity and is calculated as follows:

10 (Identical Deoxyribonucleotides x 100)/(Length of Alignment – Total Number of Gaps in Alignment)

Subsequence: The term "subsequence" means a polynucleotide having one or more (*e.g.*, several) nucleotides deleted from the 5' and/or 3' end of a mature polypeptide coding sequence; wherein the subsequence encodes a sweet-tasting fragment.

Variant: The term "variant" means a sweet-tasting polypeptide comprising an alteration, *i.e.*, a substitution, insertion, and/or deletion of one or more (*e.g.*, several) amino acid residues at one or more positions. A substitution means a replacement of the amino acid occupying a position with a different amino acid; a deletion means removal of the amino acid occupying a position; and an insertion means adding an amino acid adjacent to the amino acid occupying a position.

25

BRIEF DESCRIPTION OF DRAWINGS

Fig. 1 shows the outline of the various PCR constructs made in the Examples section with a reference to the relevant Example no. for each construct.

30 DETAILED DESCRIPTION OF THE INVENTION

Sweet-Tasting Polypeptides

The first aspect of the invention relates to a Gram-positive host cell transformed with a polynucleotide encoding a sweet-tasting polypeptide, selected from the group consisting of:

35 (a) a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least

94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the mature polypeptide of SEQ ID NO:48;

(b) a polypeptide encoded by a polynucleotide that hybridizes under medium stringency conditions, medium-high stringency conditions, high stringency conditions, or very high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO:47, or (ii) the full-length complement of (i);

(c) a polypeptide encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the mature polypeptide coding sequence of SEQ ID NO:47;

(d) a variant of the mature polypeptide of SEQ ID NO:48 comprising a substitution, deletion, and/or insertion at one or more positions; and

(e) a fragment of the polypeptide of (a), (b), (c) or (d) that is sweet-tasting.

A preferred embodiment relates to the host cell of the first aspect, wherein the polypeptide comprises the amino acid sequence of the mature polypeptide shown in SEQ ID NO:48; preferably the polypeptide consists of the amino acid sequence of the mature polypeptide shown in SEQ ID NO:48 which is also known as des-Brazzein or des-pGln1 Brazzein, where the *N*-terminal Glutamine amino acid is missing compared to SEQ ID NO:33. Des-Brazzein is what is produced, after maturation, from the synthetic construct of SEQ ID NO:47 and it is sweeter than full-length Brazzein.

Even more preferably, the polypeptide comprises the amino acid sequence shown in SEQ ID NO:33 or consists of the amino acid sequence shown in SEQ ID NO:33 which is the full-length form of Brazzein.

In another preferred embodiment, the polypeptide comprises the amino acid substitution (His31→Ala).

Another preferred embodiment relates to the host cell of the first aspect, wherein the expression of at least one endogenous protease enzyme has been reduced or inactivated; preferably the expression of at least one intracellular, extracellular or cell-wall-associated protease has been reduced or inactivated; more preferably the expression of intracellular serine protease, LspA, has been reduced or inactivated and/or the expression of one or more extracellular protease selected from the group of AprE, NprE, Bpf, Mpr, Epr, NprB, Bpr and Vpr, has been reduced or inactivated and/or the expression of at least the cell-wall-associated protease, WprA, has been reduced or inactivated.

Further, in a preferred embodiment, the expression of at least one of the following endogenous gene-products have been reduced or inactivated: SigF, AmyE and surfactin.

One very well-known way of reducing or inactivating the expression of a polypeptide in a host cell is to partially or fully delete the corresponding encoding gene(s). Accordingly, it is preferred, that said expression of the above-mentioned activities has been reduced or inactivated by partial or full deletion of the respective encoding gene(s).

5 A preferred embodiment relates to the host cell of the first aspect, wherein the chaperone protein, PrsA, is expressed; preferably the chaperone protein, PrsA, is over-expressed, such as, outlined in the Examples below.

In an embodiment, the present invention relates to isolated polypeptides having a sequence identity to the mature polypeptide of SEQ ID NO:48 of at least 60%, e.g., at least
10 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%, which are sweet-tasting. In one aspect, the polypeptides differ by no more than ten amino acids, e.g., nine amino acids, eight amino acids, seven amino acids, six amino acids, five amino acids, four amino acids, three amino acids, two amino acids, or one amino acid from
15 the mature polypeptide of SEQ ID NO:48.

A polypeptide of the present invention preferably comprises or consists of the amino acid sequence of SEQ ID NO:48 or an allelic variant thereof; or is a sweet-tasting fragment thereof. In another aspect, the polypeptide comprises or consists of the mature polypeptide of SEQ ID NO:48.

20 In another embodiment, the present invention relates to isolated sweet-tasting polypeptides that are encoded by a polynucleotide that hybridizes under very low stringency conditions, low stringency conditions, medium stringency conditions, medium-high stringency conditions, high stringency conditions, or very high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO:47, or (ii) the full-length complement of (i)
25 (Sambrook *et al.*, 1989, *Molecular Cloning, A Laboratory Manual*, 2d edition, Cold Spring Harbor, New York).

The polynucleotide of SEQ ID NO:47 or a subsequence thereof, as well as the mature polypeptide of SEQ ID NO:48 or a fragment thereof, may be used to design nucleic acid probes to identify and clone DNA encoding sweet-tasting polypeptides from strains of different genera or species according to methods well known in the art. In particular, such probes can be used
30 for hybridization with the genomic DNA or cDNA of a cell of interest, following standard Southern blotting procedures, in order to identify and isolate the corresponding gene therein. Such probes can be considerably shorter than the entire sequence, but should be at least 15, e.g., at least 25, at least 35, or at least 70 nucleotides in length. Preferably, the nucleic acid
35 probe is at least 100 nucleotides in length, e.g., at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, at least 500 nucleotides, at least 600 nucleotides, at least 700 nucleotides, at least 800 nucleotides, or at least 900 nucleotides in length. Both DNA and

RNA probes can be used. The probes are typically labeled for detecting the corresponding gene (for example, with ^{32}P , ^3H , ^{35}S , biotin, or avidin). Such probes are encompassed by the present invention.

A genomic DNA or cDNA library prepared from such other strains may be screened for DNA that hybridizes with the probes described above and encodes a sweet-tasting polypeptide. Genomic or other DNA from such other strains may be separated by agarose or polyacrylamide gel electrophoresis, or other separation techniques. DNA from the libraries or the separated DNA may be transferred to and immobilized on nitrocellulose or other suitable carrier material. In order to identify a clone or DNA that is homologous with the mature peptide-encoding part of SEQ ID NO:47 or a subsequence thereof, the carrier material is preferably used in a Southern blot.

For purposes of the present invention, hybridization indicates that the polynucleotide hybridizes to a labeled nucleic acid probe corresponding to (i) SEQ ID NO:1; (ii) the mature polypeptide coding sequence of SEQ ID NO:47; (iii) the full-length complement thereof; or (v) a subsequence thereof; under very low to very high stringency conditions. Molecules to which the nucleic acid probe hybridizes under these conditions can be detected using, for example, X-ray film.

In one aspect, the nucleic acid probe is a polynucleotide that encodes the polypeptide of SEQ ID NO:48; the mature polypeptide thereof; or a fragment thereof. In another aspect, the nucleic acid probe is mature polypeptide-encoding part of SEQ ID NO:47.

For probes of at least 100 nucleotides in length, very low stringency conditions are defined as prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 25% formamide, following standard Southern blotting procedures for 12 to 24 hours optimally. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 45°C.

For probes of at least 100 nucleotides in length, low stringency conditions are defined as prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 25% formamide, following standard Southern blotting procedures for 12 to 24 hours optimally. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 50°C.

For probes of at least 100 nucleotides in length, medium stringency conditions are defined as prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 35% formamide, following standard Southern blotting procedures for 12 to 24 hours optimally. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 55°C.

For probes of at least 100 nucleotides in length, medium-high stringency conditions are defined as prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200

micrograms/ml sheared and denatured salmon sperm DNA, and either 35% formamide, following standard Southern blotting procedures for 12 to 24 hours optimally. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 60°C.

For probes of at least 100 nucleotides in length, high stringency conditions are defined as prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 50% formamide, following standard Southern blotting procedures for 12 to 24 hours optimally. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 65°C.

For probes of at least 100 nucleotides in length, very high stringency conditions are defined as prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 50% formamide, following standard Southern blotting procedures for 12 to 24 hours optimally. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 70°C.

In another embodiment, the present invention relates to isolated sweet-tasting polypeptides encoded by polynucleotides having a sequence identity to the mature polypeptide coding sequence of SEQ ID NO:47 of at least 60%, e.g., at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%.

In another embodiment, the present invention relates to variants of the mature polypeptide of SEQ ID NO:48 comprising a substitution, deletion, and/or insertion at one or more (e.g., several) positions. Preferably, amino acid changes are of a minor nature, that is conservative amino acid substitutions or insertions that do not significantly affect the folding and/or activity of the protein; small deletions, typically of one to about 30 amino acids; small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue; a small linker peptide of up to about 20-25 residues; or a small extension that facilitates purification by changing net charge or another function, such as a poly-histidine tract, an antigenic epitope or a binding domain.

Examples of conservative substitutions are within the groups of basic amino acids (arginine, lysine and histidine), acidic amino acids (glutamic acid and aspartic acid), polar amino acids (glutamine and asparagine), hydrophobic amino acids (leucine, isoleucine and valine), aromatic amino acids (phenylalanine, tryptophan and tyrosine), and small amino acids (glycine, alanine, serine, threonine and methionine). Amino acid substitutions that do not generally alter specific activity are known in the art and are described, for example, by H. Neurath and R.L. Hill, 1979, *In, The Proteins*, Academic Press, New York. Common substitutions are Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, and Asp/Gly.

Alternatively, the amino acid changes are of such a nature that the physico-chemical properties of the polypeptides are altered. For example, amino acid changes may improve the thermal stability of the polypeptide, alter the substrate specificity, change the pH optimum, and the like.

5 Essential amino acids in a polypeptide can be identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, 1989, *Science* 244: 1081-1085). In the latter technique, single alanine mutations are introduced at every residue in the molecule, and the resultant mutant molecules are tested for sweetness to identify amino acid residues that are critical to the activity of the molecule. See
10 also, Hilton *et al.*, 1996, *J. Biol. Chem.* 271: 4699-4708. The active site of the enzyme or other biological interaction can also be determined by physical analysis of structure, as determined by such techniques as nuclear magnetic resonance, crystallography, electron diffraction, or photoaffinity labeling, in conjunction with mutation of putative contact site amino acids. See, for example, de Vos *et al.*, 1992, *Science* 255: 306-312; Smith *et al.*, 1992, *J. Mol. Biol.* 224: 899-
15 904; Wlodaver *et al.*, 1992, *FEBS Lett.* 309: 59-64. The identity of essential amino acids can also be inferred from an alignment with a related polypeptide.

Single or multiple amino acid substitutions, deletions, and/or insertions can be made and tested using known methods of mutagenesis, recombination, and/or shuffling, followed by a relevant screening procedure, such as those disclosed by Reidhaar-Olson and Sauer, 1988,
20 *Science* 241: 53-57; Bowie and Sauer, 1989, *Proc. Natl. Acad. Sci. USA* 86: 2152-2156; WO 95/17413; or WO 95/22625. Other methods that can be used include error-prone PCR, phage display (e.g., Lowman *et al.*, 1991, *Biochemistry* 30: 10832-10837; U.S. Patent No. 5,223,409; WO 92/06204), and region-directed mutagenesis (Derbyshire *et al.*, 1986, *Gene* 46: 145; Ner *et al.*, 1988, *DNA* 7: 127).

25 Mutagenesis/shuffling methods can be combined with high-throughput, automated screening methods to detect activity of cloned, mutagenized polypeptides expressed by host cells (Ness *et al.*, 1999, *Nature Biotechnology* 17: 893-896). Mutagenized DNA molecules that encode active polypeptides can be recovered from the host cells and rapidly sequenced using standard methods in the art. These methods allow the rapid determination of the importance of
30 individual amino acid residues in a polypeptide.

In an embodiment, the number of amino acid substitutions, deletions and/or insertions introduced into the mature polypeptide of SEQ ID NO:48 is not more than 10, e.g., 1, 2, 3, 4, 5, 6, 7, 8 or 9.

The polypeptide may be a hybrid polypeptide in which a region of one polypeptide is
35 fused at the N-terminus or the C-terminus of a region of another polypeptide.

The polypeptide may be a fusion polypeptide or cleavable fusion polypeptide in which another polypeptide is fused at the N-terminus or the C-terminus of the polypeptide of the

present invention. A fusion polypeptide is produced by fusing a polynucleotide encoding another polypeptide to a polynucleotide of the present invention. Techniques for producing fusion polypeptides are known in the art, and include ligating the coding sequences encoding the polypeptides so that they are in frame and that expression of the fusion polypeptide is under control of the same promoter(s) and terminator. Fusion polypeptides may also be constructed using intein technology in which fusion polypeptides are created post-translationally (Cooper *et al.*, 1993, *EMBO J.* 12: 2575-2583; Dawson *et al.*, 1994, *Science* 266: 776-779).

A fusion polypeptide can further comprise a cleavage site between the two polypeptides. Upon secretion of the fusion protein, the site is cleaved releasing the two polypeptides. Examples of cleavage sites include, but are not limited to, the sites disclosed in Martin *et al.*, 2003, *J. Ind. Microbiol. Biotechnol.* 3: 568-576; Svetina *et al.*, 2000, *J. Biotechnol.* 76: 245-251; Rasmussen-Wilson *et al.*, 1997, *Appl. Environ. Microbiol.* 63: 3488-3493; Ward *et al.*, 1995, *Biotechnology* 13: 498-503; and Contreras *et al.*, 1991, *Biotechnology* 9: 378-381; Eaton *et al.*, 1986, *Biochemistry* 25: 505-512; Collins-Racie *et al.*, 1995, *Biotechnology* 13: 982-987; Carter *et al.*, 1989, *Proteins: Structure, Function, and Genetics* 6: 240-248; and Stevens, 2003, *Drug Discovery World* 4: 35-48.

In a preferred embodiment, the host cell of the present invention is not a *Lactococcus lactis* cell, a *Lactobacillus reuteri* cell or a *Bacillus subtilis* cell transformed with at least one extrachromosomal element or plasmid comprising a polynucleotide encoding Brazzein.

Sources of Sweet-Tasting Polypeptides

A sweet-tasting polypeptide of the present invention may be obtained from plants or other organisms or microorganisms of any genus. For purposes of the present invention, the term "obtained from" as used herein in connection with a given source shall mean that the polypeptide encoded by a polynucleotide is produced by the source or by a strain in which the polynucleotide from the source has been inserted. In one aspect, the polypeptide obtained from a given source is secreted extracellularly.

The polypeptide may be a bacterial polypeptide. For example, the polypeptide may be a Gram-positive bacterial polypeptide such as a *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, or *Streptomyces* polypeptide, or a Gram-negative bacterial polypeptide such as a *Campylobacter*, *E. coli*, *Flavobacterium*, *Fusobacterium*, *Helicobacter*, *Ilyobacter*, *Neisseria*, *Pseudomonas*, *Salmonella*, or *Ureaplasma* polypeptide.

In one aspect, the polypeptide is a *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, or *Bacillus thuringiensis* polypeptide.

In another aspect, the polypeptide is a *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, or *Streptococcus equi* subsp. *Zooepidemicus* polypeptide.

In another aspect, the polypeptide is a *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, or *Streptomyces lividans* polypeptide.

The polypeptide may also be a fungal polypeptide. For example, the polypeptide may be a yeast polypeptide such as a *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* polypeptide; or a filamentous fungal polypeptide such as an *Acremonium*, *Agaricus*, *Alternaria*, *Aspergillus*, *Aureobasidium*, *Botryosphaeria*, *Ceriporiopsis*, *Chaetomidium*, *Chrysosporium*, *Claviceps*, *Cochliobolus*, *Coprinopsis*, *Coptotermes*, *Corynascus*, *Cryphonectria*, *Cryptococcus*, *Diplodia*, *Exidia*, *Filibasidium*, *Fusarium*, *Gibberella*, *Holomastigotoides*, *Humicola*, *Irpex*, *Lentinula*, *Leptosphaeria*, *Magnaporthe*, *Melanocarpus*, *Meripilus*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Piromyces*, *Poitrasia*, *Pseudoplectania*, *Pseudotrichonympha*, *Rhizomucor*, *Schizophyllum*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trichoderma*, *Trichophaea*, *Verticillium*, *Volvariella*, or *Xylaria* polypeptide.

In another aspect, the polypeptide is a *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, or *Saccharomyces oviformis* polypeptide.

In another aspect, the polypeptide is an *Acremonium cellulolyticus*, *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium inops*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium merdarium*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium tropicum*, *Chrysosporium zonatum*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola grisea*, *Humicola insolens*, *Humicola lanuginosa*, *Irpex lacteus*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium funiculosum*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Thielavia achromatica*, *Thielavia albomyces*, *Thielavia albopilosa*, *Thielavia australeinsis*, *Thielavia fimeti*, *Thielavia microspora*, *Thielavia ovispora*, *Thielavia peruviana*, *Thielavia setosa*, *Thielavia spededonium*, *Thielavia subthermophila*, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* polypeptide.

It will be understood that for the aforementioned species the invention encompasses both the perfect and imperfect states, and other taxonomic equivalents, e.g., anamorphs, regardless of the species name by which they are known. Those skilled in the art will readily recognize the identity of appropriate equivalents.

Strains of these species are readily accessible to the public in a number of culture collections, such as the American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Centraalbureau Voor Schimmelcultures (CBS), and Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

The polypeptide may be identified and obtained from other sources including microorganisms or plants isolated from nature (e.g., soil, composts, water, etc.) using the above-mentioned probes. Techniques for isolating microorganisms from natural habitats are well known in the art. A polynucleotide encoding the polypeptide may then be obtained by similarly screening a genomic DNA or cDNA library of another microorganism or plant or of a mixed DNA sample. Once a polynucleotide encoding a polypeptide has been detected with the probe(s), the polynucleotide can be isolated or cloned by utilizing techniques that are well known to those of ordinary skill in the art (see, e.g., Sambrook *et al.*, 1989, *supra*).

Polynucleotides

The present invention also relates to isolated polynucleotides encoding a sweet-tasting polypeptide of the present invention, as described above.

The techniques used to isolate or clone a polynucleotide encoding a polypeptide are known in the art and include isolation from genomic DNA or cDNA, or a combination thereof. The cloning of the polynucleotides from genomic DNA can be effected, e.g., by using the well known polymerase chain reaction (PCR) or antibody screening of expression libraries to detect cloned DNA fragments with shared structural features. See, e.g., Innis *et al.*, 1990, *PCR: A Guide to Methods and Application*, Academic Press, New York. Other nucleic acid amplification procedures such as ligase chain reaction (LCR), ligation activated transcription (LAT) and polynucleotide-based amplification (NASBA) may be used. The polynucleotides may be cloned from a plant, or a related organism and thus, for example, may be an allelic or species variant of the polypeptide encoding region of the polynucleotide.

Modification of a polynucleotide encoding a polypeptide of the present invention may be necessary for synthesizing polypeptides substantially similar to the polypeptide. The term "substantially similar" to the polypeptide refers to non-naturally occurring forms of the polypeptide. These polypeptides may differ in some engineered way from the polypeptide isolated from its native source, e.g., variants that differ in specific activity, thermostability, pH optimum, or the like. The variants may be constructed on the basis of the polynucleotide

presented as the mature polypeptide coding sequence of SEQ ID NO:47, e.g., a subsequence thereof, and/or by introduction of nucleotide substitutions that do not result in a change in the amino acid sequence of the polypeptide, but which correspond to the codon usage of the host organism intended for production of the enzyme, or by introduction of nucleotide substitutions that may give rise to a different amino acid sequence. For a general description of nucleotide substitution, see, e.g., Ford *et al.*, 1991, *Protein Expression and Purification* 2: 95-107.

Nucleic Acid Constructs

The present invention also relates to nucleic acid constructs comprising a polynucleotide of the present invention operably linked to one or more control sequences that direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences.

A polynucleotide may be manipulated in a variety of ways to provide for expression of the polypeptide. Manipulation of the polynucleotide prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying polynucleotides utilizing recombinant DNA methods are well known in the art.

The control sequence may be a promoter sequence, a polynucleotide that is recognized by a host cell for expression of a polynucleotide encoding a polypeptide of the present invention. The promoter sequence contains transcriptional control sequences that mediate the expression of the polypeptide. The promoter may be any polynucleotide that shows transcriptional activity in the host cell of choice including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

Examples of suitable promoters for directing transcription of the nucleic acid constructs of the present invention in a bacterial host cell are the promoters obtained from the *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus subtilis* *xylA* and *xylB* genes, *E. coli* *lac* operon, *E. coli* *trc* promoter (Egon *et al.*, 1988, *Gene* 69: 301-315), *Streptomyces coelicolor* agarase gene (*dagA*), and prokaryotic beta-lactamase gene (Villa-Kamaroff *et al.*, 1978, *Proc. Natl. Acad. Sci. USA* 75: 3727-3731), as well as the *tac* promoter (DeBoer *et al.*, 1983, *Proc. Natl. Acad. Sci. USA* 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in Gilbert *et al.*, 1980, *Scientific American*, 242: 74-94; and in Sambrook *et al.*, 1989, *supra*.

A preferred embodiment relates to the host cell of the first aspect, wherein the polypeptide is expressed from an endogenous or an exogenous promoter; preferably the polypeptide is expressed from a synthetic promoter.

The control sequence may also be a suitable transcription terminator sequence, which is recognized by a host cell to terminate transcription. The terminator sequence is operably linked to the 3'-terminus of the polynucleotide encoding the polypeptide. Any terminator that is functional in the host cell of choice may be used in the present invention.

5 The control sequence may also be a suitable leader sequence, when transcribed is a nontranslated region of an mRNA that is important for translation by the host cell. The leader sequence is operably linked to the 5'-terminus of the polynucleotide encoding the polypeptide. Any leader sequence that is functional in the host cell of choice may be used.

10 The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3'-terminus of the polynucleotide and, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence that is functional in the host cell of choice may be used.

15 The control sequence may also be a signal peptide coding region that encodes a signal peptide linked to the N-terminus of a polypeptide and directs the polypeptide into the cell's secretory pathway. The 5'-end of the coding sequence of the polynucleotide may inherently contain a signal peptide coding sequence naturally linked in translation reading frame with the segment of the coding sequence that encodes the polypeptide. Alternatively, the 5'-end of the coding sequence may contain a signal peptide coding sequence that is foreign to the coding sequence. A foreign signal peptide coding sequence may be required where the coding sequence does not naturally contain a signal peptide coding sequence. Alternatively, a foreign signal peptide coding sequence may simply replace the natural signal peptide coding sequence in order to enhance secretion of the polypeptide. However, any signal peptide coding sequence that directs the expressed polypeptide into the secretory pathway of a host cell of choice may be used.

25 Effective signal peptide coding sequences for bacterial host cells are the signal peptide coding sequences obtained from the genes for *Bacillus* NCIB 11837 maltogenic amylase, *Bacillus licheniformis* subtilisin, *Bacillus licheniformis* beta-lactamase, *Bacillus stearothermophilus* alpha-amylase, *Bacillus stearothermophilus* neutral proteases (*nprT*, *nprS*, *nprM*), and *Bacillus subtilis* *prsA*. Further signal peptides are described by Simonen and Palva, 30 1993, *Microbiological Reviews* 57: 109-137.

A preferred embodiment relates to the host cell of the first aspect, wherein the polypeptide is expressed with a heterologous secretion signal peptide; preferably the secretion signal peptide is the native Savinase secretion signal peptide shown in positions 1-27 of SEQ ID NO:42, the synthetic LQ2 secretion signal peptide shown in positions 1-29 of SEQ ID NO:36 or 35 the altered Savinase signal peptide encoded by SEQ ID NO:49.

The control sequence may also be a propeptide coding sequence that encodes a propeptide positioned at the N-terminus of a polypeptide. The resultant polypeptide is known as

a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to an active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding sequence may be obtained from the genes for *Bacillus subtilis* alkaline protease (*aprE*) or *Bacillus subtilis* neutral protease (*nprT*).

Where both signal peptide and propeptide sequences are present at the N-terminus of a polypeptide, the propeptide sequence is positioned next to the N-terminus of a polypeptide and the signal peptide sequence is positioned next to the N-terminus of the propeptide sequence.

It may also be desirable to add regulatory sequences that regulate expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those that cause expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory systems in prokaryotic systems include the *lac*, *tac*, and *trp* operator systems. Other examples of regulatory sequences are those that allow for gene amplification.

Expression Vectors

The present invention also relates to recombinant expression vectors comprising a polynucleotide of the present invention, a promoter, and transcriptional and translational stop signals. The various nucleotide and control sequences may be joined together to produce a recombinant expression vector that may include one or more convenient restriction sites to allow for insertion or substitution of the polynucleotide encoding the polypeptide at such sites. Alternatively, the polynucleotide may be expressed by inserting the polynucleotide or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

The recombinant expression vector may be any vector (e.g., a plasmid or virus) that can be conveniently subjected to recombinant DNA procedures and can bring about expression of the polynucleotide. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vector may be a linear or closed circular plasmid.

The vector may be an autonomously replicating vector, i.e., a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one that, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector

or plasmid or two or more vectors or plasmids that together contain the total DNA to be introduced into the genome of the host cell, or a transposon, may be used.

The vector preferably contains one or more selectable markers that permit easy selection of transformed, transfected, transduced, or the like cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

Examples of bacterial selectable markers are the *dal* genes from *Bacillus subtilis* or *Bacillus licheniformis*, or markers that confer antibiotic resistance such as ampicillin, chloramphenicol, kanamycin, or tetracycline resistance.

The vector preferably contains an element(s) that permits integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome.

For integration into the host cell genome, the vector may rely on the polynucleotide's sequence encoding the polypeptide or any other element of the vector for integration into the genome by homologous or non-homologous recombination. Alternatively, the vector may contain additional polynucleotides for directing integration by homologous recombination into the genome of the host cell at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should contain a sufficient number of nucleic acids, such as 100 to 10,000 base pairs, 400 to 10,000 base pairs, and 800 to 10,000 base pairs, which have a high degree of sequence identity to the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding polynucleotides. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. The origin of replication may be any plasmid replicator mediating autonomous replication that functions in a cell. The term "origin of replication" or "plasmid replicator" means a polynucleotide that enables a plasmid or vector to replicate *in vivo*.

Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAM β 1 permitting replication in *Bacillus*.

More than one copy of a polynucleotide of the present invention may be inserted into a host cell to increase production of a polypeptide. An increase in the copy number of the polynucleotide can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the

polynucleotide where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the polynucleotide, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

Accordingly, a preferred embodiment relates to the host cell of the first aspect, which comprises multiple polynucleotides encoding the sweet-tasting polypeptide; preferably the multiple polynucleotides are identical copies; more preferably, the multiple polynucleotides employ different codon usages and therefore have different nucleotide sequences while still encoding the same polypeptide; this will improve their genomic stability since they will have a reduced tendency to recombine with each other via homologous recombination.

It is preferred in the host cell of the first aspect, that the polynucleotide(s) encoding the sweet-tasting polypeptide is/are comprised in the genome of the cell or located on an extrachromosomal replicative expression vector.

The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, e.g., Sambrook *et al.*, 1989, *supra*).

Host Cells

The present invention also relates to recombinant host cells, comprising a polynucleotide of the present invention operably linked to one or more control sequences that direct the production of a polypeptide of the present invention. A construct or vector comprising a polynucleotide is introduced into a host cell so that the construct or vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector as described earlier. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication. The choice of a host cell will to a large extent depend upon the gene encoding the polypeptide and its source.

The host cell may be any cell useful in the recombinant production of a polypeptide of the present invention, e.g., a Gram-positive bacterium. Gram-positive bacteria include, but not limited to, *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, and *Streptomyces*.

The bacterial host cell may be any *Bacillus* cell including, but not limited to, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, and *Bacillus thuringiensis* cells.

The bacterial host cell may also be any *Streptococcus* cell including, but not limited to, *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, and *Streptococcus equi* subsp. *Zooepidemicus* cells.

The bacterial host cell may also be any *Streptomyces* cell including, but not limited to, *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, and *Streptomyces lividans* cells.

The introduction of DNA into a *Bacillus* cell may be effected by protoplast transformation (see, e.g., Chang and Cohen, 1979, *Mol. Gen. Genet.* 168: 111-115), using competent cells (see, e.g., Young and Spizizen, 1961, *J. Bacteriol.* 81: 823-829, or Dubnau and Davidoff-Abelson, 1971, *J. Mol. Biol.* 56: 209-221), electroporation (see, e.g., Shigekawa and Dower, 1988, *Biotechniques* 6: 742-751), or conjugation (see, e.g., Koehler and Thorne, 1987, *J. Bacteriol.* 169: 5271-5278). The introduction of DNA into a *Streptomyces* cell may be effected by protoplast transformation and electroporation (see, e.g., Gong *et al.*, 2004, *Folia Microbiol. (Praha)* 49: 399-405), conjugation (see, e.g., Mazodier *et al.*, 1989, *J. Bacteriol.* 171: 3583-3585), or transduction (see, e.g., Burke *et al.*, 2001, *Proc. Natl. Acad. Sci. USA* 98: 6289-6294). The introduction of DNA into a *Pseudomonas* cell may be effected by electroporation (see, e.g., Choi *et al.*, 2006, *J. Microbiol. Methods* 64: 391-397) or conjugation (see, e.g., Pinedo and Smets, 2005, *Appl. Environ. Microbiol.* 71: 51-57). The introduction of DNA into a *Streptococcus* cell may be effected by natural competence (see, e.g., Perry and Kuramitsu, 1981, *Infect. Immun.* 32: 1295-1297), protoplast transformation (see, e.g., Catt and Jollick, 1991, *Microbios* 68: 189-207), electroporation (see, e.g., Buckley *et al.*, 1999, *Appl. Environ. Microbiol.* 65: 3800-3804) or conjugation (see, e.g., Clewell, 1981, *Microbiol. Rev.* 45: 409-436). However, any method known in the art for introducing DNA into a host cell can be used.

Methods of Production

The present invention also relates to methods of producing a polypeptide of the present invention, comprising: (a) cultivating a recombinant host cell of the present invention under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide.

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

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

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

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

15 In an alternative aspect, the polypeptide is not recovered, but rather a host cell of the present invention expressing the polypeptide is used as a source of the polypeptide.

Signal Peptide

20 The present invention also relates to an isolated polynucleotide encoding a signal peptide, non-limiting specimens of which are shown in the Examples. The polynucleotides may further comprise a gene encoding a protein, which is operably linked to the signal peptide. The protein is preferably foreign to the signal peptide and/or propeptide.

 The present invention also relates to nucleic acid constructs, expression vectors and recombinant host cells comprising such polynucleotides.

25 The present invention also relates to methods of producing a protein, comprising: (a) cultivating a recombinant host cell comprising such polynucleotide; and (b) recovering the protein.

30 The protein may be native or heterologous to a host cell. The term "protein" is not meant herein to refer to a specific length of the encoded product and, therefore, encompasses peptides, oligopeptides, and polypeptides. The term "protein" also encompasses two or more polypeptides combined to form the encoded product. The proteins also include hybrid polypeptides and fused polypeptides.

35 Preferably, the protein is a hormone or variant thereof, enzyme, receptor or portion thereof, antibody or portion thereof, or reporter. For example, the protein may be a hydrolase, isomerase, ligase, lyase, oxidoreductase, or transferase, e.g., an aminopeptidase, amylase, carbohydrase, carboxypeptidase, catalase, cellobiohydrolase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, endoglucanase, esterase, alpha-galactosidase, beta-galactosidase, glucoamylase, alpha-glucosidase, beta-glucosidase,

invertase, laccase, lipase, mannosidase, mutanase, oxidase, pectinolytic enzyme, peroxidase, phytase, polyphenoloxidase, proteolytic enzyme, ribonuclease, transglutaminase, xylanase, or beta-xylosidase. The gene may be obtained from any prokaryotic, eukaryotic, or other source.

The present invention is further described by the following examples that should not be construed as limiting the scope of the invention.

EXAMPLES

Materials and methods

Strains

- 10 A164: *Bacillus subtilis* wild type strain ATCC 6051a
- A164Δ5: *Bacillus subtilis* A164 with deletions in the genes *sigF*, *nprE*, *aprE*, *amyE*, and *srfAC*, rendering them all inactive. Sloma, A., and L. Christianson (1999). Nucleic acids encoding a polypeptide having protease activity. U.S. patent 5,891,701.
- A164Δ10: *Bacillus subtilis* A164 with deletions in the genes *sigF*, *nprE*, *aprE*, *amyE*, *srfAC*,
 15 *wprA*, *bpr*, *vpr*, *mpr*, and *epr*, rendering them all inactive. Connelly MB *et al* (2004) "Extracellular Proteolytic activity Plays a Central Role in Swarming Motility in *Bacillus subtilis*" *J. Bacteriol.* 186(13) 4159-4167
- A164Δ11: *Bacillus subtilis* A164Δ10 with a disruption in the *ispA* gene, rendering it inactive.
- PP2877: *Bacillus subtilis* strain with a strong promoter (P4199 EP1062318), the *prsA* gene from
 20 *B. subtilis* and the *spec* gene inserted in the *amyE* locus.
- PP2921-2: *Bacillus subtilis* strain with a heterologous tandem promoter (as disclosed in WO 1999/043835), an amylase gene and the *cat* gene inserted in the *alr* locus.
- SJ8373: *Bacillus licheniformis* strain with the heterologous tandem promoter followed by the Savinase signal peptide fused to a protease gene.
- 25 AEB1141: *Bacillus subtilis* strain with the heterologous tandem promoter and the *neo* gene inserted in the *pel* locus.
- AEB1576: *B. subtilis* A164Δ5 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of the *bra* Bs-1 gene encoding Brazzein which is N-terminally fused to the LQ2 signal peptide (SP_{LQ2}).
- 30 AEB1579: *B. subtilis* A164Δ10 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of SP_{LQ2}_bra Bs-1.
- AEB1581: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of SP_{LQ2}_bra Bs-1.
- AEB1617: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *alr* locus,
 35 directing transcription of the *des-bra* Bs-1 gene encoding the *des*-pGlu1 form of Brazzein, N-terminally fused to SP_{LQ2}.

AEB1627: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of the *bra* Bs-1 gene encoding Brazzein, N-terminally fused to the Savinase signal peptide (SP_{SAV}).

AEB1629: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of the *des-bra* Bs-1 gene encoding the *des*-pGlu1 Brazzein, N-terminally fused to SP_{SAV}.

AEB1633: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *pel* locus, directing transcription of SP_{LQ2_bra} Bs-1.

AEB1635: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of the *des-bra* Bs-2 gene encoding the *des*-pGlu1 form of Brazzein, N-terminally fused to SP_{LQ2}.

AEB1636: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of the *des-bra* Bs-3 gene encoding the *des*-pGlu1 form of Brazzein, N-terminally fused to SP_{LQ2}.

AEB1641: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of SP_{LQ2_bra} Bs-2.

AEB1643: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of SP_{LQ2_bra} Bs-3.

AEB1657: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of the *bra* Bs-4 gene encoding Brazzein, N-terminally fused to SP_{SAV}.

AEB1659: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of the *bra* Bs-1 gene encoding Brazzein, N-terminally fused to SP_{SAV}. The DNA sequence encoding SP_{SAV} in AEB1659 has been altered so that approximately every 20 bp is different from the sequence in the original SP_{SAV}. The resulting sequence is termed SP_{SAValt}.

AEB1665: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of an operon consisting of SP_{SAValt_bra} Bs-1 directly followed by SP_{SAV_bra} Bs-4.

AEB1677: *B. subtilis* A164Δ11 with the heterologous tandem promoter inserted in the *alr* locus, directing transcription of SP_{SAV_bra} Bs-1, and with the heterologous tandem promoter inserted in the *pel* locus, directing transcription of SP_{LQ2_bra} Bs-1.

AEB1678: AEB1627 with a strong promoter (P4199 EP1062318), the *prsA* gene from *B. subtilis* and the *spec* gene inserted in the *amyE* locus.

AEB1680: AEB1629 with a strong promoter (P4199 EP1062318), the *prsA* gene from *B. subtilis* and the *spec* gene inserted in the *amyE* locus.

PCR templates

The polynucleotide denoted bra Bs-1 (SEQ ID NO 34) was cloned by PCR using plasmid pAEB1550 as template (Example 3). pAEB1550 was obtained by ordering SEQ ID NO 37 (which includes bra Bs-1) from a manufacturer of synthetic DNA, which synthesized the fragment and inserted it in a standard *E. coli* cloning vector with the ColE1 origin and the *bla* gene encoding ampicillin resistance.

The polynucleotide denoted bra Bs-2 (SEQ ID NO 39) was cloned by PCR using plasmid pAEB1548 as template (Example 7). pAEB1548 was obtained as described for pAEB1550, except that bra Bs-1 was replaced by bra Bs-2 in the sequence ordered.

The polynucleotide denoted bra Bs-3 (SEQ ID NO 40) was cloned by PCR using plasmid pAEB1552 as template (Example 7). pAEB1552 was obtained as described for pAEB1550, except that bra Bs-1 was replaced by bra Bs-3 in the sequence ordered.

The polynucleotide denoted bra Bs-4 (SEQ ID NO 46) was cloned by PCR using plasmid pAEB1631 as template (Example 14). pAEB1631 was obtained by ordering SEQ ID NO 47 (which includes bra Bs-4) from a manufacturer of synthetic DNA, which synthesized the fragment and inserted it in a standard *E. coli* cloning vector as described for pAEB1550.

pSJ7305: An altered version of the Savinase signal peptide gene (SEQ ID NO 45) inserted in a standard *E. coli* cloning vector.

Table 1. Primer and sequence overview

Primer No./Seq.	SEQ ID NO	Nucleotide Sequence 5'→ or amino acid sequence N-term. →
pab59	1	GGTATAGTAAGGCTATCAAGGGTGTC
pab245	2	GAGGGGCCTGCTGTCCA
pab257	3	TTCGGACGGACTCTACCTCAAATG
pab312	4	TGTAACGCCCTCAAAAGTAAGAAC
pab351	5	CTTTTGCCGCTTCCGTTGTTATTC
pab499	6	AGGACTGCGCTTCGGTGCTCA
pab500	7	GCAACCTGCAAATATGTCTAAAAATCAATG
pab556	8	AGTCCGCTCCTATGCTCGTCCTTC
pab557	9	GGCTTCATCAACCTTATCCATCATTTCA
pab618	10	GAGCTCTATAAAAATGAGGAGGGAACCG
pab645	11	ACCTCCACCAAACCTTACGGATACGAC
pab671	12	TGAACTTTAATAAAATTGATTTAGACAATTGGAAGA
pab672	13	TCTTCCAATTGTCTAAATCAATTTTATTAAAGTTCA

pab673	14	CGGTTCCCTCCTCATTTTTATAGAGCTC
pab688	15	GCATTTAGCTCATCTATTGCATCAGCACAAAGACAAGTGTAAGAAGGT ATACGAGAACTACCC
pab689	16	GCATTTAGCTCATCTATTGCATCAGCAGACAAGTGTAAGAAGGTATA CGAGAACTACCC
pab690	17	TGCTGATGCAATAGATGAGCTAAATGC
pab695	18	GTTTGCCGATTACAAAAACATCAGCCGACAAGTGTAAGAAGGTATAC GAGAACTACCCA
pab696	19	GGCTGATGTTTTTGTAAATCGGCAAAC
pab697	20	GTTTGCCGATTACAAAAACATCAGCCGACAAGTGTAAGAAGGTTTAC GAGAACTAC
pab698	21	TAGATAAAGACTGGCTTGAAAACGCACG
pab699	22	CGTGCGTTTTCAAGCCAGTCTTTATCTATTAGTACTCGCAGTAGTCG CAGATGCAC
pab700	23	GATCCTCTGAATATTTGCGAACCTCTG
pab701	24	CATACGCCTGCCCGCCTTG
pab702	25	TAAGCATTTGAGGTAGAGTCCGTCCGA
pab707	26	GTTTGCCGATTACAAAAACATCAGCCCAAGACAAGTGTAAGAAGGTA TACGAGAACTAC
pab708	27	GTTTGCCGATTACAAAAACATCAGCCCAGGACAAGTGTAAGAAGGTT TACGAGAACTAC
pab717	28	GCATTTAGCTCATCTATTGCATCAGCACAGGACAAATGTAAGAAGGT TTACGAGAACTACCC
pab720	29	GTGCATCTGCGACTACTGCGAGTACTAATATAAAAATGAGGAGGGAA CCGAATGAAAAAACCGC
pab721	30	TTAGTACTCGCAGTAGTCGCAGATGCAC
pab738	31	CATTTTCTTCATCTATTGCAAGCGCTCAGGACAAGTGTAAGAAGGTA TACGAGAACTACCC
pab739	32	AGCGCTTGCAATAGATGAAGAAAATG
Brazzein	33	QDKCKKVYENYPVSKCQLANQCNYDCKLDKHARSGECFYDEKRNLQCI CDYCEY
bra Bs-1	34	CAGGACAAGTGTAAGAAGGTATACGAGAACTACCCAGTAAGCAAGT GCCAACTTGCGAACCAATGCAACTACGACTGCAAGCTTGACAAGCAT GCTCGCTCTGGCGAGTGTTTCTACGACGAGAAGCGCAACCTTCAGT GCATCTGCGACTACTGCGAGTAC
SP-LQ2	35	ATGAAACAACAAAAACGGCTTTACGCCCGATTGGTGCTTATGTGCAC

		GCTGTTATTTGTCAGTTTGCCGATTACAAAAACATCAGCC
SP-LQ2 Brazzein	36	MKQQKRLYARLVLMCTLLFVSLPITKTS AQDKCKKVYENYPVSKCQLAN QCNYDCKLDKHARSGECFYDEKRNLQCICDYCEY
SP-LQ2 bra Bs-1	37	Encoding sequence of the SP-LQ2 signal peptide fused with bra Bs-1 is shown in the sequence listing.
SOE- PCR	38	SOE-PCR fragment for insertion of heterologous tandem promoter-SP-LQ2 bra Bs-1- <i>cat</i> construct in <i>alrA</i> .
bra Bs-2	39	CAAGACAAGTGTAAGAAGGTATACGAGAACTACCCAGTATCAAAGTG CCAACTTGCGAACCAATGCAACTACGACTGTAAGCTTGACAAGCATG CTCGCTCTGGCGAGTGCTTCTACGACGAGAAGCGCAACCTTCAATG CATCTGTGACTACTGCGAGTAC
bra Bs-3	40	CAGGACAAGTGTAAGAAGGTTTACGAGAACTACCCTGTTTCTAAGTG TCAACTTGCGAACCAAGTGCAACTACGACTGCAAGCTTGACAAGCATG CTCGCTCAGGCGAGTGTTTCTACGACGAGAAGCGCAACCTTCAATG TATCTGCGACTACTGTGAGTAC
SP_Sav	41	ATGAAAAAACCGCTGGGAAAAATTGTCGCAAGCACAGCACTTCTTAT TTCAGTGGCATTAGCTCATCTATTGCATCAGCA
SP_Sav Brazzein	42	MKKPLGKIVASTALLISVAFSSSIASAQDKCKKVYENYPVSKCQLANQCN YDCKLDKHARSGECFYDEKRNLQCICDYCEY
SOE- PCR	43	SOE-PCR fragment for insertion of heterologous tandem promoter-SP _{Sav} -bra Bs-1- <i>cat</i> in <i>alrA</i> is shown in the sequence listing.
Con- struct	44	DNA sequence of <i>amyE::spec-P4199-prsA</i> is shown in the sequence listing.
SOE- PCR	45	SOE-PCR fragment for insertion of heterologous tandem promoter-SP _{LQ2} -bra Bs-1- <i>neo</i> in <i>pefB</i> is shown in the sequence listing.
bra Bs-4	46	CAGGACAAATGTAAGAAGGTTTACGAGAACTACCCTGTAAGCAAGTG TCAACTTGCGAACCAAGTGCAACTACGACTGTAAGCTTGACAAACATG CTCGCTCAGGCGAGTGTTTCTACGATGAGAAGCGCAACCTTCAATG CATCTGCGACTACTGTGAGTAC
Con- struct	47	Ordered fragment encoding the fusion of the LQ2 signal peptide with the bra Bs-4 gene encoding the des-pGlu1-Brazzein; shown in the sequence listing.
des- pGlu1 Brazzein	48	Encoded by SEQ ID NO:47: MKQQKRLYARLVLMCTLLFVSLPITKTS A DKCKKVYENYPVSKCQLANQCNYDCKLDKHARSGECFYDEKRNLQCIC DYCEY
SP _{Sav} alt	49	DNA encoding the altered savinase signal peptide: SP _{Sav} alt.

Molecular biological methods

DNA manipulations and transformations were performed by standard molecular biology methods as described in:

- Sambrook et al. (1989): Molecular cloning: A laboratory manual. Cold Spring Harbor laboratory, Cold Spring Harbor, NY.
- Ausubel et al. (eds) (1995): Current protocols in Molecular Biology. John Wiley and Sons.
- Harwood and Cutting (eds) (1990): Molecular Biological Methods for Bacillus. John Wiley and Sons.

Enzymes for DNA manipulation were obtained from New England Biolabs, Inc. and used essentially as recommended by the supplier.

Competent cells and transformation of *B. subtilis* was obtained as described in Yasbin et al. (1975): Transformation and transfection in lysogenic strains of *Bacillus subtilis*: evidence for selective induction of prophage in competent cells. *J. Bacteriol.* **121**, 296-304.

Fed-batch fermentations

All media were sterilized by methods known in the art. Unless otherwise described, tap water was used. The ingredient concentrations referred to in the below recipes are before any inoculation.

First inoculum medium

LB agar: 10 g/l peptone from casein; 5 g/l yeast extract; 10 g/l Sodium Chloride; 12 g/l Bacto-agar adjusted to pH 6.8 to 7.2. Premix from Merck was used.

Transfer buffer

M-9 buffer (deionized water is used): disodium hydrogen phosphate, 2H₂O 8.8 g/l; Potassium dihydrogen phosphate 3 g/l; Sodium Chloride 4 g/l; Magnesium sulphate, 7H₂O 0.2 g/l.

Inoculum shake flask medium

PRK-50: 110 g/l soy grits; disodium hydrogen phosphate, 2H₂O 5 g/l; pH adjusted to 8.0 with NaOH/H₃PO₄ before sterilization.

Make-up medium

Tryptone (Casein hydrolysate from Difco) 30 g/l; Magnesium sulphate, 7H₂O 4 g/l; di-Potassium hydrogen phosphate 7 g/l; di-Sodium hydrogen phosphate, 2H₂O 7 g/l; Di-Ammonium sulphate 4 g/l; Citric acid 0.78 g/l; Vitamins (Thiamin-dichlorid 34.2 mg/l; Riboflavin 2.9 mg/l; Nicotinic

acid 23 mg/l; Calcium D-pantothenate 28.5 mg/l; Pyridoxal-HCl 5.7 mg/l; D-biotin 1.1 mg/l; Folic acid 2.9 mg/l; Trace metals (MnSO₄, H₂O 39.2 mg/l; FeSO₄, 7H₂O 157 mg/l; CuSO₄, 5H₂O 15.6 mg/l; ZnCl₂ 15.6 mg/l); Antifoam (SB2121) 1.25 ml/l; pH adjusted to 6.0 with NaOH/H₃PO₄ before sterilization.

5

Feed medium

Sucrose 708 g/l

Procedure for inoculum steps

10 The strain was grown on LB agar slants 1 day at 37°C.

The agar was washed with M-9 buffer. The optical density (OD) at 650 nm of the resulting cell suspension was measured.

The inoculum shake flask (PRK-50) was inoculated with an inoculum of OD (650 nm) x ml cell suspension = 0.1.

15 The shake flask was incubated at 37°C at 300 rpm for 20 hr.

The fermentation in the main fermentor (fermentation tank) was started by inoculating the main fermentor with the growing culture from the shake flask. The inoculated volume was 11% of the make-up medium (80 ml for 720 ml make-up media).

20 *Fermentor Equipment*

Standard lab fermentors were used equipped with a temperature control system, pH control with ammonia water and phosphoric acid, dissolved oxygen electrode to measure >20% oxygen saturation through the entire fermentation.

25 *Fermentation parameters*

Temperature: 38°C

The pH was kept between 6.8 and 7.2 using ammonia water and phosphoric acid

Control: 6.8 (ammonia water); 7.2 phosphoric acid

Aeration: 1.5 liter/min

30 Agitation: 1500 rpm

Feed strategy

0 hr. 0.05 g/min/kg initial broth after inoculation

8 hr. 0.156 g/min/kg initial broth after inoculation

35 End 0.156 g/min/kg initial broth after inoculation

Experimental setup

The fermentations were run for five days.

Example 1. Designing a synthetic Brazzein gene for expression in *B. subtilis*

The protein sequence of wild type Brazzein from *Pentadiplandra brazzeana* Baillon was acquired from the article by Ming D, Hellekant G (1994) "Brazzein, a new high-potency thermostable sweet protein from *Pentadiplandra brazzeana* B" *FEBS Lett.* **355**(1): 106-108. A synthetic gene was designed from the Brazzein protein sequence using a codon usage table based on codon usage in *Bacillus subtilis*. The gene was named bra Bs-1.

The amino acid sequence of the mature form of Brazzein can be found in SEQ ID NO 33. The nucleotide sequence of bra Bs-1 can be found in SEQ ID NO 34.

Example 2. The Brazzein gene fused to the required elements

The Brazzein gene was ordered on a synthetic DNA fragment. To enable secretion of the Brazzein molecule the gene was N-terminally fused to the LQ2 signal peptide (SP_{LQ2}) which is known to be functional in *B. subtilis*.

The nucleotide sequence of SP_{LQ2} alone can be found in SEQ ID NO 35. The amino acid sequence of Brazzein fused to SP_{LQ2} can be found in SEQ ID NO 36.

To enable cloning by SOE-PCR (Splicing by Overlapping Extension PCR), appropriate fragments upstream and downstream of the Brazzein gene were included. The nucleotide sequence of the entire synthetic fragment ordered is given in SEQ ID NO 37. The plasmid received from the manufacturer with the synthetic fragment of DNA was termed pAEB1550.

Example 3. Obtaining a SOE-PCR fragment with the Brazzein gene and flanking regions for integration in the *alr*-region in *B. subtilis*

By performing SOE-PCR with the primers and templates listed in Table 2, the bra Bs-1 gene fused to SP_{LQ2} was joined to an upstream DNA fragment that contained the region upstream of the *alr* gene in the *B. subtilis* chromosome and a heterologous tandem promoter (as disclosed in WO 1999/043835). Downstream, a DNA fragment was added which contains the *cat* gene for selection in *Bacillus*, the *alr* gene and the region downstream of the *alr* gene in the *B. subtilis* chromosome. The entire SOE-PCR fragment is depicted in Fig. 1. The nucleotide sequence of the fragment can be found in SEQ ID NO 38.

Table 2: SOE-PCR strategy to insert SP_{LQ2}_bra Bs-1 in the *alr* locus

1. PCRs				SOE-PCR	
Reaction	Template DNA from	Primers	Size	Primers	Size
Upstream	PP2921-2	pab499 & pab673	3.5 kb	pab556 & pab557	6.5 kb

Middle	pAEB1550	pab618 & pab672	0.3 kb		
Downstream	PP2921-2	pab500 & pab671	2.9 kb		

Example 4. Transforming the SOE-PCR fragment with *bra Bs-1* into three *B. subtilis* host strains

The SOE-PCR fragment with *bra Bs-1* obtained in Example 3 was transformed into host strain *B. subtilis* A164Δ5, selecting for resistance against chloramphenicol, since the *cat* gene on the SOE-PCR fragment renders the strain resistant to this antibiotic. Correct insertion of the fragment was tested in chloramphenicol resistant transformants by PCR on chromosomal DNA and sequencing of the resulting PCR fragment using appropriate primers. One correct strain was named AEB1576.

To investigate the effect of inactivation of genes encoding proteases on Brazzein expression the fragment inserted in *alr* in AEB1576 was transferred to the host strains *B. subtilis* A164Δ10 and A164Δ11 by transformation of the strains with chromosomal DNA from AEB1576 and selecting for resistance against chloramphenicol. Correct insertion of the fragment in chloramphenicol resistant transformants was verified by PCR on chromosomal DNA using appropriate primers. Correct strains were termed:

AEB1576 (A164Δ5 host)
AEB1579 (A164Δ10 host)
AEB1581 (A164Δ11 host)

Example 5. Evaluation of by SDS-PAGE of Brazzein yield in shake flasks from the three *B. subtilis* strains with *bra Bs-1*

AEB1576, AEB1579, and AEB1581 were fermented in shake flasks for three days and samples were withdrawn. Prior to SDS-PAGE analysis, pH in the samples was lowered by addition of acetic acid to between 4.0 and 5.0, followed by incubation at 90°C for 15 min and spinning down of the samples. This procedure leads to precipitation of most of the protein in the sample except Brazzein. Fresh DTT was added to the supernatants along with sample buffer before application on SDS-gel. The samples were run on NuPAGE® Novex® 4-12% Bis-Tris gels with MES running buffer.

It was observed that while no significant band the size of Brazzein was detectable on the gel the when AEB1576 was applied, a distinct Brazzein band was seen from AEB1579, and with AEB1581 this band was significantly stronger. Thus, one or more of the five protease deletions (*wprA*, *bpr*, *vpr*, *mpr*, and *epr*) leading from A164Δ5 (background strains of AEB1576) to A164Δ10 (background strain of AEB1579) has a positive impact on Brazzein expression, and in

particular the deletion in the protease gene *ispA* leading from A164Δ10 to A164Δ11 (background strain of AEB1581) has a significant positive impact on Brazzein yield. Therefore, A164Δ11 was selected as host for the remaining experiments with Brazzein expression.

5 **Example 6. Fermentation of AEB1579 and AEB1581 in laboratory scale fermentors**

AEB1579 and AEB1581 were fermented in laboratory scale fermentors as described in Materials and methods. Samples from the fermentations were applied on SDS-PAGE as described in Example 5. The results obtained in shake flasks were confirmed, since AEB1579 produced very little or no Brazzein, while the same amount of AEB1581 fermentation broth on
10 gel resulted in a clear band on the gel, corresponding to approximately a 25-fold increase in Brazzein yield as judged from the intensity of the bands on the SDS-PAGE. The yield from the tank fermentation of AEB1581 is expected to be industrially relevant.

Example 7. Expression of Brazzein genes *bra Bs-2* and *bra Bs-3*

Two additional Brazzein genes, *bra Bs-2* and *bra Bs-3*, were designed, based in a codon usage table from *B. subtilis*. The nucleotide sequences of *bra Bs-2* and *bra Bs-3* can be found in SEQ ID NO 39 and 40, respectively. The genes were ordered on synthetic DNA fragments fused to SP_{LQ2} and surrounded by the same upstream and downstream fragments as described in Example 2. The plasmids received with the synthetic fragments were termed pAEB1548 (*bra*
20 *Bs-2*) and pAEB1552 (*bra Bs-3*). The PCR strategy employed for obtaining strains expressing Brazzein encoded by *bra Bs-2* and *bra Bs-3* is shown in Table 3.

Table 3: SOE-PCR strategy to insert SP_{LQ2}-*bra Bs-2* and SP_{LQ2}-*bra Bs-3* in the *alr* locus

1. PCRs				SOE-PCR	
Reaction		Template DNA from	Primers	Size	
Upstream		AEB1617	pab499 & pab696	3.6 kb	pab556 & pab557 6.5 kb
Middle	<i>bra Bs-2</i>	pAEB1548	pab672 & pab707	0.3 kb	
	<i>bra Bs-3</i>	pAEB1552	pab672 & pab708	0.3 kb	
Downstream		AEB1617	pab500 & pab671	2.9 kb	

25 The SOE-PCR fragments (shown in Fig. 1) were inserted in the *alr* locus in *B. subtilis* host strain A164Δ11 (via A164Δ5) as described in Example 4, and the resulting strains were termed AEB1641 (*bra Bs-2*) and AEB1643 (*bra Bs-3*).

The amounts of Brazzein produced from AEB1641 and AEB1643 were evaluated in shake flasks as described in Example 5. The yields were comparable to the yield obtained from AEB1581.

5 Example 8. Expression of the des-pGlu1 form of Brazzein

The major form of Brazzein in *Pentadiplandra brazzeana* Baillon is of 54 aa. However, a minor and twice as sweet form exists that lacks the amino terminal pyroglutamic acid residue, des-pGlu1-Brazzein (Assadi-Porter *et al.* (2000) "Efficient Production of Recombinant Brazzein, a Small, Heat-Stable, Sweet-Tasting Protein of Plant Origin" *Arch. Biochem. Biophys.* **376**(2), 252-258).

The PCR strategy employed to obtain a strain expressing this form of Brazzein encoded by bra Bs-1 is shown in Table 4. The PCR strategy employed to obtain a strain expressing this form of Brazzein encoded by bra Bs-2 and bra Bs-3 is shown in Table 5. The SOE-PCR fragments obtained are depicted in Fig. 1. They are identical to the fragments described in Example 3 designed for construction of AEB1581 (bra Bs-1), and in Example 7 for construction of AEB1641 (bra Bs-2) and AEB1643 (bra Bs-3), except that the Brazzein genes lack the first codon of the mature part of Brazzein, leading to genes encoding des-pGlu1-Brazzein.

Table 4: SOE-PCR strategy to insert SP_{LQ2}_des-bra Bs-1 in *alr*

1. PCRs				SOE-PCR	
Reaction	Template DNA from	Primers	Size	Primers	Size
Upstream	AEB1576	pab499 & pab696	3.6 kb	pab556 & pab557	6.5 kb
Downstream	AEB1576	pab500 & pab695	3.1 kb		

Table 5: SOE-PCR strategy to insert SP_{LQ2}_des-bra Bs-2 and -bra Bs-3 in the *alr* locus

1. PCRs				SOE-PCR	
Reaction	Template DNA from	Primers	Size	Primers	Size
Upstream	AEB1617	pab499 & pab696	3.6 kb	pab556 & pab557	6.5 kb
Middle	bra Bs-2	pAEB1548	pab672 & pab695		
	bra Bs-3	pAEB1552	pab672 & pab697		
Downstream	AEB1617	pab500 & pab671	2.9 kb		

The SOE-PCR fragments were inserted in the *alr* locus in *B. subtilis* host strain A164Δ11 (via A164Δ5) as described in Example 4, and the resulting strains were termed:

Brazzein gene	Strain
<i>des-bra</i> Bs-1	AEB1617
<i>des-bra</i> Bs-2	AEB1635
<i>des-bra</i> Bs-3	AEB1636

5

The amount of Brazzein produced from the three strains was evaluated in shake flasks as described in Example 5. Very little *des*-pGlu1-Brazzein was visible on the SDS-gel from any of the strains.

10 **Example 9. Replacing the LQ2 signal peptide with the signal peptide from Savinase**

By performing SOE-PCR with appropriate primers and templates (shown in Table 6) a fragment was obtained (Fig. 1) similar to the one used during construction of AEB1581, but with SP_{LQ2} in front of the full-length *bra* Bs-1 gene replaced by the Savinase signal peptide (SP_{SAV}), which is also known to be functional in *Bacillus*. The nucleotide sequence of SP_{SAV} is found in
 15 SEQ ID NO 41. The amino acid sequence of Brazzein fused to SP_{SAV} can be found in SEQ ID NO 42. The DNA sequence of the entire fragment amplified by SOE-PCR is given in SEQ ID NO 43.

Table 6: SOE-PCR strategy to insert SP_{SAV}*_bra* Bs-1 in the *alr* locus

1. PCRs				SOE-PCR	
Reaction	Template DNA from	Primers	Size	Primers	Size
Upstream	AEB1576	pab59 & pab499	3.2 kb	pab556 & pab557	6.5 kb
Middle	SJ8373	pab245 & pab690	0.8 kb		
Downstream	AEB1576	pab645 & pab688	2.5 kb		

20

One correct strain obtained by insertion of the SOE-PCR fragment in *alr* in host *B. subtilis* A164Δ11 (via A164Δ5) as described in Example 4 was termed AEB1627.

In a similar manner, SP_{LQ2} in AEB1617 (containing the shorter form of *bra* Bs-1 which
 25 encodes *des*-pGlu1-Brazzein) was replaced by SP_{SAV}. Primers and templates used are shown in Table 7. The resulting strain was termed AEB1629.

Table 7: SOE-PCR strategy to insert SP_{SAV}*_des*-pGlu1-*bra* Bs-1 in the *alr* locus

1. PCRs				SOE-PCR	
Reaction	Template	Primers	Size	Primers	Size

	DNA from				
Upstream	AEB1576	pab59 & pab499	3.2 kb	pab556 & pab557	6.5 kb
Middle	SJ8373	pab245 & pab690	0.8 kb		
Downstream	AEB1576	pab645 & pab689	2.5 kb		

Example 10. Production of Brazzein with the Savinase signal peptide in shake flasks

Brazzein production from AEB1627 and AEB1629 was evaluated in shake flasks as described in Example 5. For full length Brazzein (AEB1581 compared to AEB1627) replacing SP_{LQ2} with SP_{SAV} led to an approximate doubling of Brazzein yield, as judged from the intensity of the bands on SDS-PAGE. For *des*-pGlu1-Brazzein, replacing SP_{LQ2} with SP_{SAV} did not seem to lead to an increase in Brazzein expression (AEB1617 compared to AEB1629).

Example 11. Production of Brazzein with the Savinase signal peptide in laboratory scale fermentors

AEB1627 and AEB1629 were fermented in lab tanks as described in Example 6. Fermentation broth was applied on SDS-PAGE as described in Example 5. Under these conditions, the yield of Brazzein obtained from AEB1627 was approximately 1.8 fold higher than the yield from AEB1581. The yield from AEB1629 was approximately 5 fold lower than from AEB1581 and thus approximately 7.5 fold lower than from AEB1627. Thus, the signal peptide is an important parameter that should be considered when optimizing Brazzein expression.

Example 12. Strains with overexpression of the PrsA chaperone during Brazzein expression

The effect of overexpression of the PrsA chaperone on Brazzein yield was tested by introducing a cassette encompassing a strong promoter (P4199 in EP1062318) controlling transcription of the *prsA* gene from *B. subtilis* in the *amyE* locus in strains AEB1627 (SP_{SAV_bra} Bs-1) and AEB1629 (SP_{SAV_des-pGlu1-bra} Bs-1). This was obtained by transformation of the strains with chromosomal DNA from PP2877. Correct transformants could be identified by selection for spectinomycin, since the *spec* gene, encoding resistance to this antibiotic, is located next to the *prsA* cassette in the *amyE* locus in PP2877.

The DNA sequence of the *spec*-P4199-*prsA* cassette in *amyE* in PP2877 is found in SEQ ID NO 44.

Spectinomycin resistant transformants were tested by PCR, and correct strains were named AEB1678 (based on AEB1627) and AEB1680 (based on AEB1629).

The yield of Brazzein from the strains with co-expression of PrsA was evaluated in shake flasks as described in Example 5. Extra PrsA did not seem to have an effect on

expression of full length Brazzein (AEB1678), but led to a more than 10-fold increase in the yield of the *des*-pGlu1 form of Brazzein (not shown).

Example 13. Effect of overexpression of the *PrsA* chaperone in laboratory scale

5 fermentors

AEB1678 and AEB1680 were fermented in lab scale fermentors as described in Example 6. Fermentation broth was applied on SDS-PAGE as described in Example 5. Under these conditions, co-expression of *PrsA* led to an approximate 1.3-fold decrease in yield of full length Brazzein (AEB1678 compared to AEB1627). But the yield of the *des*-pGlu1 form of
 10 Brazzein obtained from AEB1680 was increased 2 fold relative to the yield in AEB1629. Thus, co-expression of the *PrsA* chaperone has a significant positive effect on expression of the *des*-pGlu1-form of Brazzein.

Example 14. Constructing a strain with a fourth synthetic Brazzein gene

15 A fourth synthetic gene, bra Bs-4, was designed, based on a codon usage table from *B. subtilis*. The nucleotide sequence of the gene can be found in SEQ ID NO 46. The gene was designed so that every 15 to 20 bp are different compared to the nucleotide sequence of bra Bs-1. This low level of homology between the genes impairs homologous recombination between the two genes, thus increasing their genetic stability in the genome. The *des*-pGlu1
 20 form of bra Bs-4 was ordered with an N-terminal fusion to SP_{LQ2}. Somewhat different upstream and downstream regions were added to enable cloning, but the flanking regions are identical in the resulting strain. The entire fragment ordered is found as SEQ ID NO 47. The plasmid received from the manufacturer with bra Bs-4 was termed pAEB1631.

Using SOE-PCR with appropriate primers (Table 8) the full length version of bra Bs-4
 25 fused to SP_{SAV} was joined to upstream and downstream DNA fragments as described in Example 3, leading to the gene being functionally linked to a heterologous tandem promoter, followed by the *cat* gene, and surrounded by DNA fragments that enabled integration in the *Bacillus alr* chromosomal region (Fig. 1).

30 **Table 8:** SOE-PCR strategy to insert SP_{SAV}-bra Bs-4 in the *alr* locus

1. PCRs				SOE-PCR	
Reaction	Template DNA from	Primers	Size	Primers	Size
Upstream	AEB1627	pab499 & pab690	3.6 kb	pab556 & pab557	6.5 kb
Middle	pAEB1631	pab672 & pab717	0.3 kb		
Downstream	AEB1627	pab500 & pab671	2.9 kb		

The SOE-PCR fragment was inserted in the *alr* locus in *B. subtilis* A164Δ11 as described in Example 4. Chloramphenicol resistant transformants were further verified by PCR and sequencing using appropriate primers. A correct strain was termed AEB1657.

Expression of Brazzein from AEB1657 was evaluated in shake flasks as described in Example 5. The amount of Brazzein expressed from AEB1627 and AEB1657 appeared to be similar.

Example 15. Replacing the DNA sequence encoding the Savinase signal peptide in

AEB1627 with a synthetic version

Homologous recombination between bra Bs-1 and bra Bs-4 is impaired due to every 15 to 20 bp being different. However the codons encoding SP_{SAV} fused to the Brazzein genes in AEB1627 (bra Bs-1) and AEB1657 (bra Bs-4) are identical. By SOE-PCR with appropriate primers and templates (Table 9) a fragment was obtained that was similar to the one used when constructing AEB1627 (Example 9) except that the part of the gene encoding SP_{SAV} in front of bra Bs-1 was replaced with a gene (SP_{SAV}alt) still encoding SP_{SAV} but with low enough homology to the wt SP_{SAV} gene to prevent recombination (Fig. 1). The DNA sequence encoding SP_{SAV}alt is given in SEQ ID NO 49.

Table 9: SOE-PCR strategy to insert SP_{SAV}alt_bra Bs-1 in the *alr* locus

1. PCRs				SOE-PCR	
Reaction	Template DNA from	Primers	Size	Primers	Size
Upstream	AEB1627	pab499 & pab702	3.1 kb	pab556 & pab557	6.5 kb
Middle	pSJ7305	pab257 & pab739	0.5 kb		
Downstream	AEB1627	pab500 & pab738	3.1 kb		

The SOE-PCR fragment was inserted in the *alr* locus in A164Δ11 (via A164Δ5) as described in Example 4, and a correct strain was named AEB1659. The amount of Brazzein expressed from AEB1627 and AEB1659 was similar when evaluated in shake flask fermentation as described in Example 5.

Example 16. Constructing a strain with an operon consisting of two Brazzein genes in tandem

The SP_{SAV}_Brazzein gene fusions in AEB1657 and AEB1659 are different enough to impair homologous recombination between them. Thus, an operon consisting of these two

genes in tandem will be stable. Such an operon was constructed by amplifying the upstream *alr* region, heterologous tandem promoter and *bra* Bs-4 from AEB1659 and the downstream *alr* region, *cat* gene, and *bra* Bs-1 from AEB1657 and joining the two fragments by SOE-PCR. The resulting fragment is shown in Fig. 1, the primers and templates used are given in Table 10.

5

Table 10: SOE-PCR strategy to insert SP_{sav}alt_ *bra* Bs-1-SP_{sav}_ *bra* Bs-4 in *alr*

1. PCRs				SOE-PCR	
Reaction	Template DNA from	Primers	Size	Primers	Size
Upstream	AEB1659	pab499 & pab721	3.7 kb	pab556 & pab557	6.8 kb
Downstream	AEB1657	pab500 & pab720	3.2 kb		

The SOE-PCR fragment was inserted in the *alr* locus in A164Δ11 (via A164Δ5) as described in Example 4, and a correct strain was named AEB1665.

10

Example 17. *Brazzein expression from AEB1665 during tank fermentation*

AEB1665 was fermented in lab tanks as described in Example 6. Fermentation broth was applied on SDS-PAGE as described in Example 5. Judged from the gels, the yield of Brazzein obtained from AEB1665 appeared to be slightly decreased relative to the yield from the one-copy strain AEB1627. Since this was unexpected, AEB1677 and AEB1627 were fermented in lab scale fermentors by a slightly different method, and this time the yield of AEB1677 was increased at least 50% compared to AEB1627 (not shown). Thus, fermentation conditions are an important parameter that should be considered when optimizing Brazzein expression.

20

Example 18. *Expression of the Brazzein gene from the pel locus*

By using the primers and templates given in Table 11, a SOE-PCR fragment was obtained that contained the heterologous tandem promoter and *bra* Bs-1 gene fused to SP_{LQ2} (identical to the construct in AEB1581), surrounded by upstream and downstream fragments allowing insertion in the *pel* locus in *B. subtilis* A164. The fragment also contained the *neo* gene, encoding resistance to Kanamycin resistance. The SOE-PCR fragment is shown in Fig. 1.

25

Table 11: SOE-PCR strategy to insert SP_{LQ2}_ *bra* Bs-1 in the *pel* locus

1. PCRs				SOE-PCR	
Reaction	Template DNA from	Primers	Size	Primers	Size

Upstream	AEB1141	pab59 & pab351	2.9 kb	pab351 & pab701	7.8 kb
Middle	AEB1576	pab312 & pab699	0.8 kb		
Downstream	AEB1141	pab698 & pab700	4.5 kb		

The nucleotide sequence of the SOE-PCR fragment can be found in SEQ ID NO 45. By transforming the SOE-PCR fragment into *B. subtilis* A164Δ11 (via A164Δ5, as described in Example 4) and selecting for Kanamycin resistance a strain with the upstream heterologous tandem promoter and the open reading frame with the bra Bs-1 gene fused to SP_{LQ2} could be isolated. The strain was further verified by PCR and sequencing using appropriate primers. A correct strain was named AEB1633.

Expression of Brazzein from AEB1633 was evaluated in shake flasks as described in Example 5. The amount of Brazzein expressed from AEB1633 was similar to the amount expressed from AEB1581.

Example 19. Constructing a strain with two copies of the Brazzein gene, one in the *alr* and one in the *pel* chromosomal region

AEB1633 (constructed in Example 18) contains a Brazzein expression cassette inserted in the *amyE* locus along with the *neo* gene. To obtain a strain with two such expression cassettes, AEB1633 was transformed with chromosomal DNA from AEB1627 which contains a similar cassette inserted in the *alr* locus along with the *cat* gene (constructed in Example 9). By selecting for chloramphenicol resistance, a strain that had the expression cassette inserted in the *alr* locus could be selected. The resulting strain, AEB1677, contains one Brazzein encoding gene (fused to SP_{SAV}) in the *alr* locus and one (fused to SP_{LQ2}) in the *pel* locus, both genes under transcriptional control of the heterologous tandem promoter.

Example 20. Production of Brazzein in tanks from the two-copy Brazzein strain AEB1677

AEB1677 was fermented in laboratory scale tanks as described in Example 6. Fermentation broth was applied on SDS-PAGE as described in Example 5. The yield of Brazzein obtained from the two-copy strain AEB1677 appeared to be approximately 85% of the yield from the one-copy strain AEB1627. We expect that this unexpectedly low yield is caused by the fermentation conditions as is also observed for the two-copy strain AEB1665 in Example 17.

CLAIMS

1. A Gram-positive host cell transformed with at least one polynucleotide encoding a sweet-tasting polypeptide, selected from the group consisting of:

(a) a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the mature polypeptide of SEQ ID NO:48;

(b) a polypeptide encoded by a polynucleotide that hybridizes under medium stringency conditions, medium-high stringency conditions, high stringency conditions, or very high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO:47, or (ii) the full-length complement of (i);

(c) a polypeptide encoded by a polynucleotide comprising a nucleotide sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the mature polypeptide coding sequence of SEQ ID NO:47;

(d) a variant of the mature polypeptide of SEQ ID NO:48 comprising a substitution, deletion, and/or insertion at one or more positions; and

(e) a fragment of the polypeptide of (a), (b), (c) or (d) that is sweet-tasting.

2. The host cell of claim 2, wherein the polypeptide comprises the amino acid sequence of the mature polypeptide shown in SEQ ID NO:48.

3. The host cell of claim 1 or 2, wherein the polypeptide consists of the amino acid sequence of the mature polypeptide shown in SEQ ID NO:48.

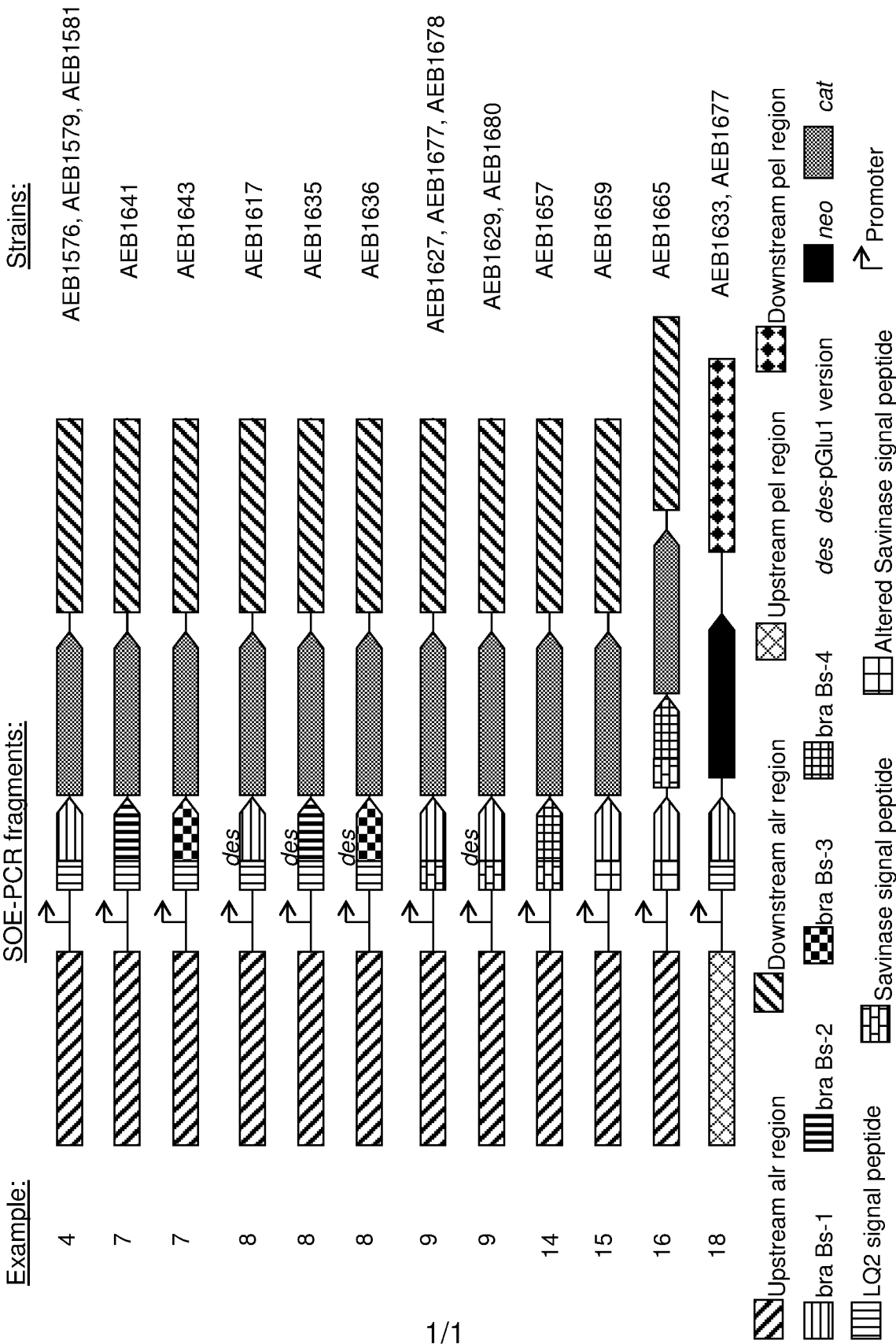
4. The host cell of claim 1 or 2, wherein the polypeptide comprises the amino acid sequence shown in SEQ ID NO:33.

5. The host cell of claim 4, wherein the polypeptide consists of the amino acid sequence shown in SEQ ID NO:33.

6. The host cell of any of claims 1-5, which is a *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, or *Streptomyces* cell.

7. The host cell of claim 6, which is a *Bacillus* cell; preferably a *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, or *Bacillus thuringiensis* cell.
8. The host cell of claim 7, wherein the expression of at least one endogenous protease enzyme has been reduced or inactivated.
9. The host cell of claim 8, wherein the expression of at least one intracellular, extracellular or cell-wall-associated protease has been reduced or inactivated.
10. The host cell of claim 9, wherein the expression of intracellular serine protease, LspA, has been reduced or inactivated.
11. The host cell of claim 9, wherein the expression of one or more extracellular protease selected from the group of AprE, NprE, Bpf, Mpr, Epr, NprB, Bpr and Vpr, has been reduced or inactivated.
12. The host cell of claim 9, wherein the expression of at least the cell-wall-associated protease, WprA, has been reduced or inactivated.
13. The host cell of any of claims 1-7, wherein the expression of at least one of the following endogenous gene-products have been reduced or inactivated: SigF, AmyE and surfactin.
14. The host cell of any of claims 8-13, wherein said expression has been reduced or inactivated by partial or full deletion of the respective encoding gene(s).
15. The host cell of any of claims 1-14, wherein the polypeptide is expressed with a heterologous secretion signal peptide; preferably the secretion signal peptide is the native Savinase secretion signal peptide shown in positions 1-27 of SEQ ID NO:42, the synthetic LQ2 secretion signal peptide shown in positions 1-29 of SEQ ID NO:36 or the altered Savinase signal peptide encoded by SEQ ID NO:49.
16. The host cell of any of claims 1-15, wherein the polypeptide is expressed from an endogenous or an exogenous promoter; preferably wherein the polypeptide is expressed from a synthetic promoter.

17. The host cell of any of claims 1-16, wherein the chaperone protein, PrsA, is expressed.
18. The host cell of any of claims 1-17, which comprises multiple polynucleotides encoding
5 the sweet-tasting polypeptide.
19. The host cell of claim 18, wherein the multiple polynucleotides are identical copies.
20. The host cell of claim 19, wherein the multiple polynucleotides employ different codon
10 usages and therefore have different nucleotide sequences while still encoding the same polypeptide.
21. The host cell of any of claims 1-20, wherein the polynucleotide(s) encoding the sweet-tasting polypeptide is/are comprised in the genome of the cell or located on an
15 extrachromosomal replicative expression vector.
22. A method of producing a sweet-tasting polypeptide, comprising:
- (a) cultivating the host cell of any of claims 1-21 under conditions conducive for production of the polypeptide; and
- 20 (b) recovering the polypeptide.



1/1
Figure 1

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2012/055123

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/43
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data, FSTA

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BERLEC A ET AL: "Large increase in brazzein expression achieved by changing the plasmid/strain combination of the NICE system in Lactococcus lactis", LETTERS IN APPLIED MICROBIOLOGY, vol. 48, no. 6, June 2009 (2009-06), pages 750-755, XP055002564, ISSN: 0266-8254	1-7,15, 16,18-22
Y	cited in the application page 751, column 1, paragraph 3 - column 2, paragraph 2 page 754, column 1, paragraph 4 - column 2, paragraph 1 ----- -/--	8-14,17



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

14 June 2012

Date of mailing of the international search report

20/06/2012

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

van Heusden, Miranda

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2012/055123

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BERLEC A ET AL: "Optimization of fermentation conditions for the expression of sweet-tasting protein brazzein in Lactococcus lactis", LETTERS IN APPLIED MICROBIOLOGY, WILEY-BLACKWELL PUBLISHING LTD, GB, vol. 46, no. 2, 1 February 2008 (2008-02-01), pages 227-231, XP002563579, ISSN: 0266-8254, DOI: DOI:10.1111/J.1472-765X.2007.02297.X cited in the application	1-7,15, 16,18-22
Y	abstract page 228, column 1, paragraph 3 -----	8-14,17
X	ALES BERLEC ET AL: "Expression of the sweet-tasting plant protein brazzein in Escherichia coli and Lactococcus lactis: a path toward sweet lactic acid bacteria", APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, SPRINGER, BERLIN, DE, vol. 73, no. 1, 16 May 2006 (2006-05-16), pages 158-165, XP019441709, ISSN: 1432-0614, DOI: DOI:10.1007/S00253-006-0438-Y cited in the application	1-7,15, 16,18-22
Y	abstract page 159, column 1, paragraph 4 - column 2, paragraph 1 page 161, column 2, paragraph 1 - paragraph 4 page 162, column 1, paragraph 3 - page 163, column 1, paragraph 2 -----	8-14,17
A	WO 2011/015633 A1 (NOVOZYMES AS [DK]; VIND JESPER [DK]; TAMS JEPPE WEGENER [DK]; BEIER LA) 10 February 2011 (2011-02-10) cited in the application the whole document -----	1-22
A	GUAN ET AL: "Expression of sweet protein brazzein by Saccharomyces cerevisiae", CHEMICAL SENSES, IRL PRESS, OXFORD, GB, vol. 20, no. 6, 1 April 1995 (1995-04-01), page 701, XP002092297, ISSN: 0379-864X cited in the application abstract ----- -/--	1-22

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2012/055123

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MASUDA T ET AL: "Developments in biotechnological production of sweet proteins", JOURNAL OF BIOSCIENCE AND BIOENGINEERING, ELSEVIER, AMSTERDAM, NL, vol. 102, no. 5, 1 November 2006 (2006-11-01), pages 375-389, XP025182978, ISSN: 1389-1723, DOI: DOI:10.1263/JBB.102.375 [retrieved on 2006-11-01] page 376, column 2, paragraph 2 page 381, column 1, paragraph 3 - column 2, paragraph 1 -----	1,6-9,22
Y	M. BINDEL CONNELLY ET AL: "Extracellular Proteolytic Activity Plays a Central Role in Swarming Motility in Bacillus subtilis", JOURNAL OF BACTERIOLOGY, vol. 186, no. 13, 1 July 2004 (2004-07-01) , pages 4159-4167, XP55029535, ISSN: 0021-9193, DOI: 10.1128/JB.186.13.4159-4167.2004 cited in the application abstract; tables 1,2 -----	8-14
Y	WO 94/19471 A (THE FINNISH NATIONAL PUBLIC HEALTH INSTITUTE) 1 September 1994 (1994-09-01) cited in the application page 4 -----	17
X,P	KR 2011 0135728 A (INHA IND PARTNERSHIP INST [KR]) 19 December 2011 (2011-12-19) cited in the application abstract -----	1,6

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2012/055123

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011015633 A1	10-02-2011	CN 102498126 A EP 2462156 A1 US 2012107873 A1 WO 2011015633 A1	13-06-2012 13-06-2012 03-05-2012 10-02-2011
WO 9419471 A	01-09-1994	AT 335083 T AU 6142694 A CN 1120853 A DE 69434807 T2 DK 0686195 T3 EP 0686195 A1 JP 4202985 B2 JP H08507207 A JP 2004350691 A US 5780261 A WO 9419471 A1	15-08-2006 14-09-1994 17-04-1996 29-03-2007 27-11-2006 13-12-1995 24-12-2008 06-08-1996 16-12-2004 14-07-1998 01-09-1994
KR 20110135728 A	19-12-2011	NONE	