

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 December 2007 (06.12.2007)

PCT

(10) International Publication Number
WO 2007/138049 A2

(51) International Patent Classification:

C12N 9/10 (2006.01) C12R 1/10 (2006.01)

C12N 15/75 (2006.01)

(21) International Application Number:

PCT/EP2007/055190

(22) International Filing Date: 29 May 2007 (29.05.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

PA 2006 00742 31 May 2006 (31.05.2006) DK

(71) Applicant (for all designated States except US):

NOVOZYMES A/S [DK/DK]; Krogshoejvej 36,
DK-2880 Bagsvaerd (DK).

(72) Inventor; and

(75) Inventor/Applicant (for US only): JOERGENSEN,
Steen, Troels [DK/DK]; Prunusvej 5, DK-3450 Allerød
(DK).

(74) Common Representative: NOVOZYMES A/S; Patents,

Krogshoejvej 36, DK-2880 Bagsvaerd (DK).

(81) Designated States (unless otherwise indicated, for every

kind of national protection available): AE, AG, AL, AM,

AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH,
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, MG, MK, MN, MW, MX,
MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO,
RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL,
PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM,
GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished
upon receipt of that report

— with sequence listing part of description published sepa-
rately in electronic form and available upon request from
the International Bureau

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: CHLORAMPHENICOL RESISTANCE SELECTION IN BACILLUS LICHENIFORMIS

(57) Abstract: The present invention relates to a modified *Bacillus licheniformis* host cell, wherein one or more naturally occurring chromosomal chloramphenicol acetyl transferase encoding gene(s), cat, has been inactivated. The inactivation of the chromosomal cat gene(s) allows the use of chloramphenicol as an efficient selective agent in methods for DNA introduction into the modified cell.



WO 2007/138049 A2

TITLE: CHLORAMPHENICOL RESISTANCE SELECTION IN BACILLUS LICHENIFORMIS**SEQUENCE LISTING**

This invention comprises a sequence listing.

5

FIELD OF THE INVENTION

The present invention relates to a modified *Bacillus licheniformis* host cell, wherein naturally occurring chromosomal chloramphenicol acetyl transferase gene(s), *cat*, has been inactivated. The inactivation of the chromosomal *cat* gene(s) allows the use of chloramphenicol as an efficient selective agent in methods for DNA introduction into the modified cell.

10

BACKGROUND OF THE INVENTION

In the industrial production of polypeptides it is of interest to achieve a product yield as high as possible. One way to increase the yield is to increase the copy number of a gene encoding a polypeptide of interest. This can be done by placing the gene on a high copy number plasmid, however plasmids are unstable and are often lost from the host cells if there is no selective pressure during the cultivation of the host cells. Another way to increase the copy number of the gene of interest is to integrate it into the host cell chromosome in multiple copies.

15

20

The *cat*-gene (chloramphenicol acetyl transferase) from plasmid pC194 is routinely used as a selective marker in *Bacillus* transformation. *Bacillus licheniformis*, however, exhibits chloramphenicol resistance to an extent which hampers the use of chloramphenicol resistance genes for selection of transformants, whether by protoplast transformation, conjugation or electroporation.

25

Genome sequencing has revealed that several different *B. licheniformis* strains harbour a sequence in the chromosome which has some homology with known chloramphenicol resistance genes. The homologous sequence is located at position at position 2720850 to 2721500 in the EMBL *Bacillus licheniformis* ATCC14580 sequence entry, accession number cp000002.

30

SUMMARY OF THE INVENTION

A putative *cat*-gene, along with flanking DNA, was identified in the EMBL *Bacillus licheniformis* ATCC14580 sequence entry, accession number cp000002. A schematic drawing is shown in figure 1.

35

We wanted to explore, whether the deletion of this putative *cat*-gene from the chromosome of *B. licheniformis* strains would allow the use of chloramphenicol for selection of DNA introduction into such *cat*-deleted strains.

Homologues to the *cat*-gene were identified in a number of different *B. licheniformis* strains from various sources, and these were inactivated in the chromosomes of those strains to provide modified *cat*-sensitive strains.

The present invention relates to methods of introducing genetic material comprising a chloramphenicol resistance marker into a modified *Bacillus licheniformis* host cells, wherein a naturally occurring chromosomal chloramphenicol resistance gene, *cat*, has been inactivated.

The presence of this gene in its active form otherwise hampers the use of chloramphenicol for selection purposes. Having a natural antibiotic resistance encoding gene in an industrial production host cell may also be a cause for concern in certain applications, e.g., for regulatory purposes.

Accordingly, in a first aspect the invention relates to a method for introducing a recombinant DNA construct into a modified *Bacillus licheniformis* host cell, the method comprising:

- a) providing a modified *Bacillus licheniformis* host cell, wherein one or more endogenous genomic chloramphenicol acetyl transferase encoding gene(s) has been inactivated;
- b) introducing a recombinant DNA construct comprising a selection marker encoding a chloramphenicol acetyl transferase into the modified host cell;
- c) cultivating the host cell of step (b) in a growth-medium comprising an inhibitory concentration of chloramphenicol; and
- d) selecting a host cell comprising the DNA construct that is capable of growing in the medium of step (c).

In a second aspect, the invention relates to a modified *Bacillus licheniformis* host cell, wherein one or more genomic chloramphenicol acetyl transferase gene(s) has been inactivated, comprising a recombinant DNA construct carrying a selection marker encoding a chloramphenicol acetyl transferase.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1 shows a schematic drawing of the *cat* gene region from *B.licheniformis* ATCC14580 with indication of the primers used herein.

Figure 2 shows a schematic drawing of the *cat* gene region from *B.licheniformis* PL1980 with indication of the primers used herein.

DEFINITIONS

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook, Fritsch & Maniatis, Molecular Cloning: A Laboratory Manual, Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (herein "Sambrook et al., 1989") DNA Cloning: A Practical Approach, Volumes I and II /D.N. Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed. 1984); Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds (1985)); Transcription And Translation (B.D. Hames & S.J. Higgins, eds. (1984)); Animal Cell Culture (R.I. Freshney, ed. (1986)); Immobilized Cells And Enzymes (IRL Press, (1986)); B. Perbal, A Practical Guide To Molecular Cloning (1984).

A "polynucleotide" is a single- or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases, the sequence of the polynucleotide is the actual sequence of the bases read from the 5' to the 3' end of the polymer. Polynucleotides include RNA and DNA, and may be isolated from natural sources, synthesized in vitro, or prepared from a combination of natural and synthetic molecules.

A "nucleic acid molecule" or "nucleotide sequence" refers to the phosphate ester polymeric form of ribonucleosides (adenosine, guanosine, uridine or cytidine; "RNA molecules") or deoxyribonucleosides (deoxyadenosine, deoxyguanosine, deoxythymidine, or deoxycytidine; "DNA molecules") in either single stranded form, or a double-stranded helix. Double stranded DNA-DNA, DNA-RNA and RNA-RNA helices are possible. The term nucleic acid molecule, and in particular DNA or RNA molecule, refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary or quaternary forms. Thus, this term includes double-stranded DNA found, inter alia, in linear or circular DNA molecules (e.g., restriction fragments), plasmids, and chromosomes. In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (i.e., the strand having a sequence homologous to the mRNA). A "recombinant DNA molecule" is a DNA molecule that has undergone a molecular biological manipulation.

A DNA "coding sequence" or an "open reading frame (ORF)" is a double-stranded DNA sequence which is transcribed and translated into a polypeptide in a cell in vitro or in vivo when placed under the control of appropriate regulatory sequences. The ORF "encodes" the polypeptide. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic sequences, cDNA from eukaryotic mRNA, genomic DNA sequences from eukaryotic (e.g., mammalian) DNA, and even

synthetic DNA sequences. If the coding sequence is intended for expression in a eukaryotic cell, a polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

An expression vector is a DNA molecule, linear or circular, that comprises a segment encoding a polypeptide of interest operably linked to additional segments that provide for its transcription. Such additional segments may include promoter and terminator sequences, and optionally one or more origins of replication, one or more selectable markers, an enhancer, a polyadenylation signal, and the like. Expression vectors are generally derived from plasmid or viral DNA, or may contain elements of both.

Transcriptional and translational control sequences are DNA regulatory sequences, such as promoters, enhancers, terminators, and the like, that provide for the expression of a coding sequence in a host cell e.g. in eukaryotic cells, polyadenylation signals are control sequences.

A "secretory signal sequence" is a DNA sequence that encodes a polypeptide (a "secretory peptide" that, as a component of a larger polypeptide, directs the larger polypeptide through a secretory pathway of a cell in which it is synthesized. The larger polypeptide is commonly cleaved to remove the secretory peptide during transit through the secretory pathway.

The term "promoter" is used herein for its art-recognized meaning to denote a portion of a gene containing DNA sequences that provide for the binding of RNA polymerase and initiation of transcription. Promoter sequences are commonly, but not always, found in the 5' non-coding regions of genes.

A chromosomal gene is rendered "non-functional" if the polypeptide that the gene encodes can no longer be expressed in a functional form. Such non-functionality of a gene can be induced by a wide variety of genetic manipulations or alterations as known in the art, some of which are described in Sambrook et al. *vide supra*. Partial deletions within the ORF of a gene will often render the gene non-functional, as will mutations e.g. substitutions, insertions, frameshifts etc.

"Operably linked", when referring to DNA segments, indicates that the segments are arranged so that they function in concert e.g. the transcription process takes place via the RNA-polymerase binding to the promoter segment and proceeding with the transcription through the coding segment until the polymerase stops when it encounters a transcription terminator segment.

"Heterologous" DNA in a host cell, in the present context refers to exogenous DNA not originating from the cell.

As used herein the term "nucleic acid construct" or "DNA construct" is intended to indicate any nucleic acid molecule of cDNA, genomic DNA, synthetic DNA or RNA origin. The

term "construct" is intended to indicate a nucleic acid segment which may be single- or double-stranded, and which may be based on a complete or partial naturally occurring nucleotide sequence encoding a polypeptide of interest. The construct may optionally contain other nucleic acid segments.

5 The nucleic acid construct of the invention encoding the polypeptide of the invention may suitably be of genomic or cDNA origin, for instance obtained by preparing a genomic or cDNA library and screening for DNA sequences coding for all or part of the polypeptide by hybridization using synthetic oligonucleotide probes in accordance with standard techniques (cf. Sambrook et al., supra).

10 The nucleic acid construct of the invention encoding the polypeptide may also be prepared synthetically by established standard methods, e.g. the phosphoamidite method described by Beaucage and Caruthers, Tetrahedron Letters 22 (1981), 1859 - 1869, or the method described by Matthes et al., EMBO Journal 3 (1984), 801 - 805. According to the phosphoamidite method, oligonucleotides are synthesized, e.g. in an automatic DNA
15 synthesizer, purified, annealed, ligated and cloned in suitable vectors.

 Furthermore, the nucleic acid construct may be of mixed synthetic and genomic, mixed synthetic and cDNA or mixed genomic and cDNA origin prepared by ligating fragments of synthetic, genomic or cDNA origin (as appropriate), the fragments corresponding to various parts of the entire nucleic acid construct, in accordance with standard techniques. The nucleic
20 acid construct may also be prepared by polymerase chain reaction using specific primers, for instance as described in US 4,683,202 or Saiki et al., Science 239 (1988), 487 - 491.

 The term nucleic acid construct may be synonymous with the term "expression cassette" when the nucleic acid construct contains the control sequences necessary for expression of a coding sequence of the present invention

25 The term "control sequences" is defined herein to include all components that are necessary or advantageous for expression of the coding sequence of the nucleic acid sequence. Each control sequence may be native or foreign to the nucleic acid sequence encoding the polypeptide. Such control sequences include, but are not limited to, a leader, a polyadenylation sequence, a propeptide sequence, a promoter, a signal sequence, and a
30 transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the nucleic acid sequence encoding a polypeptide.

 The control sequence may be an appropriate promoter sequence, a nucleic acid
35 sequence that is recognized by a host cell for expression of the nucleic acid sequence. The promoter sequence contains transcription and translation control sequences that mediate the expression of the polypeptide. The promoter may be any nucleic acid sequence that shows

transcriptional activity in the host cell of choice and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

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

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

The control sequence may also be a signal peptide-coding region, which codes for an amino acid sequence linked to the amino terminus of the polypeptide which can direct the expressed polypeptide into the cell's secretory pathway of the host cell. The 5' end of the coding sequence of the nucleic acid sequence may inherently contain a signal peptide-coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted polypeptide. Alternatively, the 5' end of the coding sequence may contain a signal peptide-coding region which is foreign to that portion of the coding sequence which encodes the secreted polypeptide. A foreign signal peptide-coding region may be required where the coding sequence does not normally contain a signal peptide-coding region. Alternatively, the foreign signal peptide coding region may simply replace the natural signal peptide coding region in order to obtain enhanced secretion of the polypeptide relative to the natural signal peptide coding region normally associated with the coding sequence. The signal peptide-coding region may be obtained from a glucoamylase or an amylase gene from an *Aspergillus* species, a lipase or proteinase gene from a *Rhizomucor* species, the gene for the alpha-factor from *Saccharomyces cerevisiae*, an amylase or a protease gene from a *Bacillus* species, or the calf preprochymosin gene. However, any signal peptide coding region capable of directing the expressed polypeptide into the secretory pathway of a host cell of choice may be used in the present invention.

The control sequence may also be a propeptide coding region, which codes for an amino acid sequence positioned at the amino 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 mature active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding region may be obtained from the *Bacillus subtilis* alkaline protease gene (*aprE*), the

Bacillus subtilis neutral protease gene (*nprT*), the *Saccharomyces cerevisiae* alpha-factor gene, or the *Myceliophthora thermophilum* laccase gene (WO 95/33836).

It may also be desirable to add regulatory sequences which allow the regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those which cause the 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 would include the *lac*, *tac*, and *trp* operator systems. Other examples of regulatory sequences are those which allow for gene amplification. In eukaryotic systems, these include the dihydrofolate reductase gene which is amplified in the presence of methotrexate, and the metallothionein genes which are amplified with heavy metals. In these cases, the nucleic acid sequence encoding the polypeptide would be placed in tandem with the regulatory sequence.

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

The present invention also relates to recombinant expression vectors comprising a nucleic acid sequence of the present invention, a promoter, and transcriptional and translational stop signals. The various nucleic acid and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleic acid sequence encoding the polypeptide at such sites. Alternatively, the nucleic acid sequence of the present invention may be expressed by inserting the nucleic acid sequence 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, and possibly secretion.

The recombinant expression vector may be any vector (e.g., a plasmid or virus) which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the nucleic acid sequence. 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 vectors may be linear or closed circular plasmids. The vector may be an autonomously replicating vector, i.e., a vector which 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 which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. The vector system may be a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the host cell, or a transposon.

The vectors of the present invention preferably contain one or more "selectable markers" which permit easy selection of transformed cells. A selectable marker is a gene the product of which provides for biocide, antibiotic or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

A "conditionally essential gene" may function as a "non-antibiotic selectable marker". Non-limiting examples of bacterial conditionally essential selectable markers are the *dal* genes from *Bacillus subtilis* or *Bacillus licheniformis*, that are only essential when the bacterium is cultivated in the absence of D-alanine. Also the genes encoding enzymes involved in the turnover of UDP-galactose can function as conditionally essential markers in a cell when the cell is grown in the presence of galactose or grown in a medium which gives rise to the presence of galactose. Non-limiting examples of such genes are those from *B. subtilis* or *B. licheniformis* encoding UTP-dependent phosphorylase (EC 2.7.7.10), UDP-glucose-dependent uridylyltransferase (EC 2.7.7.12), or UDP-galactose epimerase (EC 5.1.3.2). Also a xylose isomerase gene such as *xylA*, of *Bacilli* can be used as selectable markers in cells grown in minimal medium with xylose as sole carbon source. The genes necessary for utilizing gluconate, *gntK*, and *gntP* can also be used as selectable markers in cells grown in minimal medium with gluconate as sole carbon source. Other non-limiting examples of conditionally essential genes are given below.

Antibiotic selectable markers confer antibiotic resistance to such antibiotics as ampicillin, kanamycin, chloramphenicol, erythromycin, tetracycline, neomycin, hygromycin or methotrexate.

Furthermore, selection may be accomplished by co-transformation, e.g., as described in WO 91/17243, where the selectable marker is on a separate vector.

The vectors of the present invention preferably contain an element(s) that permits stable integration of the vector, or of a smaller part of the vector, into the host cell genome or autonomous replication of the vector in the cell independent of the genome of the cell.

The vectors, or smaller parts of the vectors, may be integrated into the host cell genome when introduced into a host cell. For chromosomal integration, the vector may rely on the nucleic acid sequence encoding the polypeptide or any other element of the vector for stable integration of the vector into the genome by homologous or nonhomologous recombination.

Alternatively, the vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the host cell. The additional nucleic acid sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should preferably contain a sufficient number of nucleic acids, such as 100 to 1,500 base pairs, preferably 400 to 1,500 base pairs, and most preferably 800 to 1,500 base pairs, which are highly homologous with 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 nucleic acid sequences.

The copy number of a vector, an expression cassette, an amplification unit, a gene or indeed any defined nucleotide sequence is the number of identical copies that are present in a host cell at any time. A gene or another defined chromosomal nucleotide sequence may be present in one, two, or more copies on the chromosome. An autonomously replicating vector may be present in one, or several hundred copies per host cell.

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, pACYC184, pUB110, pE194, pTA1060, and pAM β 1. The origin of replication may be one having a mutation which makes its functioning temperature-sensitive in the host cell (see, e.g., Ehrlich, 1978, Proceedings of the National Academy of Sciences USA 75:1433).

The present invention also relates to recombinant host cells, comprising a nucleic acid sequence of the invention, which are advantageously used in the recombinant production of the polypeptides. The term "host cell" encompasses any progeny of a parent cell which is not identical to the parent cell due to mutations that occur during replication.

The cell is preferably transformed with a vector comprising a nucleic acid sequence of the invention followed by integration of the vector into the host chromosome.

"Transformation" means introducing a vector comprising a nucleic acid sequence of the present invention into a host cell so that the vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector. Integration is generally considered to be an advantage as the nucleic acid sequence is more likely to be stably maintained in the cell.

5 Integration of the vector into the host chromosome may occur by homologous or non-homologous recombination as described above.

The transformation of a bacterial host cell may, for instance, be effected by protoplast transformation (see, e.g., Chang and Cohen, 1979, *Molecular General Genetics* 168:111-115), by using competent cells (see, e.g., Young and Spizizin, 1961, *Journal of Bacteriology* 10 81:823-829, or Dubnar and Davidoff-Abelson, 1971, *Journal of Molecular Biology* 56:209-221), by electroporation (see, e.g., Shigekawa and Dower, 1988, *Biotechniques* 6:742-751), or by conjugation (see, e.g., Koehler and Thorne, 1987, *Journal of Bacteriology* 169:5771-5278).

The transformed or transfected host cells described above are cultured in a suitable nutrient 15 medium under conditions permitting the expression of the desired polypeptide, after which the resulting polypeptide is recovered from the cells, or the culture broth.

The medium used to culture the cells may be any conventional medium suitable for growing the host cells, such as minimal or complex media containing appropriate supplements. Suitable media are available from commercial suppliers or may be prepared 20 according to published recipes (e.g. in catalogues of the American Type Culture Collection). The media are prepared using procedures known in the art (see, e.g., references for bacteria and yeast; Bennett, J.W. and LaSure, L., editors, *More Gene Manipulations in Fungi*, Academic Press, CA, 1991).

If the polypeptide is secreted into the nutrient medium, the polypeptide can be 25 recovered directly from the medium. If the polypeptide is not secreted, it is recovered from cell lysates. The polypeptide are recovered from the culture medium by conventional procedures including separating the host cells from the medium by centrifugation or filtration, precipitating the proteinaceous components of the supernatant or filtrate by means of a salt, e.g. ammonium sulphate, purification by a variety of chromatographic procedures, e.g. ion 30 exchange chromatography, gelfiltration chromatography, affinity chromatography, or the like, dependent on the type of polypeptide in question.

The polypeptides may be detected using methods known in the art that are specific for the polypeptides. These detection methods may include use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, an enzyme 35 assay may be used to determine the activity of the polypeptide.

The polypeptides of the present invention may be purified by a variety of procedures known in the art including, but not limited to, chromatography (e.g., ion exchange, affinity,

hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (e.g., preparative isoelectric focusing (IEF), differential solubility (e.g., ammonium sulfate precipitation), or extraction (see, e.g., Protein Purification, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

5

DETAILED DESCRIPTION OF THE INVENTION

As mentioned above, the invention relates in a first aspect to a method for introducing a recombinant DNA construct into a modified *Bacillus licheniformis* host cell, the method comprising:

- 10 a) providing a modified *Bacillus licheniformis* host cell, wherein one or more endogenous genomic chloramphenicol acetyl transferase encoding gene(s) has been inactivated;
- b) introducing a recombinant DNA construct comprising a selection marker encoding a chloramphenicol acetyl transferase into the modified host cell;
- c) cultivating the host cell of step (b) in a growth-medium comprising an inhibitory
15 concentration of chloramphenicol; and
- d) selecting a host cell comprising the DNA construct that is capable of growing in the medium of step (c).

Another aspect of the invention relates to a modified *Bacillus licheniformis* host cell, wherein one or more genomic chloramphenicol acetyl transferase gene(s) has been
20 inactivated, comprising a recombinant DNA construct carrying a selection marker encoding a chloramphenicol acetyl transferase.

It is to be understood that the preferred embodiments listed below are by no means exhaustive, they can be combined in various ways, as the skilled person would immediately appreciate. The skilled person would be well aware of the possibility of combining these
25 preferred embodiments with other technical features well-known in the art of introducing genetic material into *Bacillus* cells, but not listed explicitly herein, such as described in various patent disclosures, e.g., WO 2002/000907, WO 2001/090393, WO 1993/010249, WO 1999/043835, WO 1994/019471, WO 2003/055996, WO 1991/009129, WO 1994/014968, and/or WO 1996/023073, all of which are incorporated herein by reference.

30 In preferred embodiments of the invention the recombinant DNA construct is a plasmid or a linearized plasmid, preferably a linearized and concatamerized plasmid.

In another preferred embodiment of the invention the recombinant DNA construct further comprises a polynucleotide encoding a polypeptide of interest, preferably an enzyme, more preferably an amylolytic enzyme, a lipolytic enzyme, a proteolytic enzyme, a cellulytic
35 enzyme, an oxidoreductase or a plant cell-wall degrading enzyme, and even more preferably an enzyme with an activity selected from the group consisting of aminopeptidase, amylase, amyloglucosidase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase,

cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, galactosidase, beta-galactosidase, glucoamylase, glucose oxidase, glucosidase, haloperoxidase, hemicellulase, invertase, isomerase, laccase, ligase, lipase, lyase, mannosidase, oxidase, pectinase, peroxidase, phytase, phenoloxidase, polyphenoloxidase, protease, ribonuclease, transferase,
5 transglutaminase, or xylanase.

Still another preferred embodiment relates to the first and second aspects, wherein the one or more genomic chloramphenicol acetyl transferase gene(s) encodes a chloramphenicol acetyl transferase comprising an amino acid sequence which is at least 80% identical, 85% identical, 90% identical, 92% identical, 94% identical, 96% identical, 98%
10 identical, or at least 99% identical to the sequence shown in SEQ ID NO: 2, SEQ ID NO: 13, SEQ ID NO: 22, or SEQ ID NO: 24.

A preferred embodiment relates to the first and second aspects, wherein the one or more genomic chloramphenicol acetyl transferase gene(s) comprises a polynucleotide sequence which is at least 80% identical, 85% identical, 90% identical, 92% identical, 94%
15 identical, 96% identical, 98% identical, or at least 99% identical to the sequence shown in SEQ ID NO: 1, SEQ ID NO: 12, SEQ ID NO: 21, or SEQ ID NO: 23.

In another preferred embodiment of the invention the selection marker of the DNA construct encodes a chloramphenicol acetyl transferase comprising an amino acid sequence at least 80% identical, 85% identical, 90% identical, 92% identical, 94% identical, 96%
20 identical, 98% identical, or at least 99% identical to the sequence shown in SEQ ID NO: 2, SEQ ID NO: 13, SEQ ID NO: 22, or SEQ ID NO: 24.

Specific restriction enzymes or resolvases that excise portions of DNA, if it is flanked on both sides by certain recognition sequences known as resolvase sites or *res*-sites, are well known in the art, see e.g. WO 96/23073 (Novo Nordisk A/S) which is included herein by
25 reference. It is preferred, that the selection marker of the DNA construct is flanked by resolvase-sites, preferably resolvase sites are *res*.

In a preferred embodiment of the first aspect, the selection marker of the DNA construct is flanked by nucleotide sequences that are recognized by a specific resolvase, preferably the nucleotide sequences are *res*; even more preferably the at least one marker
30 gene is excised from the DNA construct by the action of a resolvase enzyme subsequent to selecting the host cell that grows under the selective conditions.

Preferably, the DNA construct of the invention further comprises a polynucleotide region that is sufficiently homologous with a genomic DNA region of the host cell to effect integration into the genome of the DNA construct by homologous recombination.

35 As exemplified below, it is also preferred that the inhibitory concentration of chloramphenicol in the growth-medium is at least 1 microgram/ml, preferably at least 2

microgram/ml, more preferably at least 3 microgram/ml, 6 microgram/ml, 9 microgram/ml, or at least 12 microgram/ml.

EXAMPLES

5 Example 1 - *cat* deletion in *Bacillus licheniformis* PL1980

Construction of deletion plasmid

Primers were designed, based on the ATCC14580 genome sequence, to allow PCR amplification of a 0.5 kb segment 5' of the *cat* gene, and another 0.5 kb segment 3' of the *cat* gene. The DNA sequence of the *cat*-gene region of *B. licheniformis* ATCC14580 is shown in
10 SEQ ID NO: 1, and the encoded protein is shown in SEQ ID NO: 2.

PCR amplifications were performed using DNA from *B. licheniformis* PL1980 as template, and READY-TO-GO™ PCR beads from Amersham Pharmacia Biotech, Inc., as described by the supplier, in the following program on a PTC-200 PCR machine from MJ Research:

- 15
- 1 cycle 95 °C 2 minutes
 - 30 cycles 94 °C 30 seconds
 - annealing X °C 1 minute (different temperatures were used in this step)
 - 72 °C 2 minutes, followed by 1 cycle of 72 °C for 5 minutes, whereafter reactions
were cooled to 10 °C.

20

The following primers were used:

SEQ ID NO: 3 (#449000): 5' cagtgaattcggttaaccggcaattcttc

SEQ ID NO: 4 (#449001): 5' cagtgtcgacgtctcttaacatctctcac

SEQ ID NO: 5 (#448309): 5' cagtgtcgacctgcttcactgattccg

25 SEQ ID NO: 6 (#448315): 5' cagtaagcttaagattcctcgatgatttc

Initial PCR amplifications were not succesful, no correctly sized fragments were obtained with either of the two primer combinations #449000 + #449001 or #448309 + #448315. Longer primers were designed:

30 SEQ ID NO: 7 (#450168): 5' cagtgaattcggttaaccggcaattcttcaaacgtc

SEQ ID NO: 8 (#450169): 5' cagtgtcgacgtctcttaacatctctcactgctgtg

SEQ ID NO: 9 (#450170): 5' cagtgtcgaccgcttcactgattccggcaatattc

SEQ ID NO: 10 (#450171): 5' cagtaagcttaagattcctcgatgattccaccacac

SEQ ID NO: 11 (#455882): 5' cagtgaattccgctttcgtttcaaatgacgg

35

These primers were used in a PCR reaction with an annealing temperature, which was ramped down from 62 °C to 53 °C in steps of 1 °C, then kept at 57 °C for further 20 cycles.

The #450170 + #450171 reaction gave a 0.5 kb fragment, which was digested with *EcoRI* + *SaII*, and ligated to *EcoRI* + *SaII* digested pUC19 (Yanisch-Perron et al., 1985, Gene 33(1): 103-119). The ligation mixture was transformed into *E. coli* SJ2 (Diderichsen et al., 1990, J. Bacteriol. 172(8): 4315-4321) by electroporation, and then selecting for ampicillin resistance. A clone was isolated having a sequence which was found to contain 23 differences as compared to the ATCC14580 sequence; this clone was kept and designated SJ7905 (SJ2/pSJ7905). The DNA sequence of the *cat*-gene region of *B. licheniformis* PL1980 is shown in SEQ ID NO: 12, and the encoded protein is shown in SEQ ID NO: 13.

The PCR with #450168 + #450169 was not successful. More PCR reactions across the entire *cat* gene region, with combinations of the above primers and a new primer, #455882, were then attempted, using the same annealing temperature regimen as in the above reactions. Primer combination #450171 + #449000 gave an approximately 1.8 kb fragment; #450171 + #450168 gave two fragments of approximately 1.7 kb and 1.1 kb, respectively; and #450171 + #455882 gave a fragment of approximately 1.0 kb.

These four fragments were purified and each used as template in a reaction with primers #450170 + #450171, using the annealing temperature regimen as above. All reactions gave a 0.5 kb fragment, as also found when this primer combination was used directly with chromosomal DNA as template, indicating that the four fragments were indeed fragments from the *cat* gene region.

The four long fragments were DNA sequenced. A schematic representation is given in figure 2. Based on the sequence, a new primer for the 3'-fragment was designed: SEQ ID NO: 14 (#457708): 5' cagtgtcgacctgcttagcgggtcttactg (corresponds to ATCC14580 *cat* 2kb pos. 606-588, but with several mismatches).

The primer combination #457708 + #450168 was used for amplification with the same annealing temperature regimen as above. The resulting 0.5 kb fragment was digested with *SaII* + *EcoRI* and cloned into *SaII* + *EcoRI* digested pUC19 vector which was transformed into *E. coli*. Two correct transformants of *E. coli* SJ2 were isolated and designated SJ8003 (SJ2/pSJ8003) and SJ8004 (SJ2/pSJ8004).

A pUC19 based plasmid, containing the assembled upstream and downstream fragments, was made by excision of the 0.5 kb *SaII* + *HindIII* fragment from pSJ7905, and insertion of this fragment into *SaII* + *HindIII* digested pSJ8003. Resulting transformants of *E. coli* SJ2 were isolated and designated SJ8011 (SJ2/pSJ8011) and SJ8012 (SJ2/pSJ8012).

The 5'cat – 3'cat segment was then transferred onto a temperature sensitive, mobilizable integration vector by excision of the 0.9 kb *EcoRI* – *HindIII* fragment from pSJ8011, and ligation of this to the 4.3 kb *EcoRI*-*HindIII* fragment of pSJ2739 (US Patent 6,100,063) to provide the so-called deletion plasmid.

Correct transformants of *B. subtilis* DN1885 (Diderichsen et al., 1990, J. Bacteriol. 172(8), 4315-4321) comprising the deletion plasmid were isolated and designated SJ8017 (DN1885/pSJ8017) and SJ8018 (DN1885/pSJ8018).

Construction of a *B. licheniformis* PL1980-derived strain with a deletion in the *cat* gene

B. subtilis conjugation donor strain PP289-5 (PP289-5 (US Patent 6,066,473) was transformed with the deletion plasmids, resulting in SJ8039 and SJ8040 (with pSJ8017), and in SJ8041 and SJ8042 (with pSJ8018).

B. licheniformis strain PP1897-3 is a derivative of *B. licheniformis* PL1980 which has an internal deletion in the alcalase gene, a deletion of most of the C-component protease coding region, and an artificially inserted promoter at the *amyL*, *xylA* and *gntP* loci. It was used as recipient in conjugations with the above two donor strains SJ8039 and SJ8040.

Conjugations were performed essentially as described in US Patent 6,066,473, selecting for erythromycin resistance. Tetracycline sensitive transconjugants were isolated, and chloramphenicol and erythromycin sensitive strains were isolated following integration and excision of the deletion plasmid.

Transconjugants were streaked onto LB PSG plates with erythromycin (5 mikrogram/ml) which were incubated at 50 °C overnight. The selection for erythromycin resistance at high temperature ensures that colonies formed have arisen by integration of the deletion plasmid into the *B. licheniformis* host strain chromosome by homologous recombination at either the 5' *cat* or the 3' *cat* sequence, because the plasmid is unable to replicate as a free plasmid at this temperature. Colonies were inoculated into TY liquid cultures and incubated at 30 °C overnight. This temperature allows the replication of the integrated plasmid, which facilitates its excision from the chromosome, and ultimately loss from the cell (indicated by erythromycin sensitivity).

The plasmid may excise via the same region of homology as was used for integration, in which case the resulting cell is identical to the host strain. Or the plasmid may alternatively excise via the other region of homology (e.g. integration via 5' *cat*, and excision via 3' *cat*, or vice versa) in which case the *cat* gene is deleted from the chromosome.

Aliquots of the overnight cultures were used to inoculate fresh TY cultures, and plated on LB PSG plates. Plates were incubated at 30 °C overnight, whereafter plates were replica plated to plates with and without erythromycin or chloramphenicol.

2 erythromycin sensitive and chloramphenicol sensitive strains were kept as SJ8071 (from donor SJ8039) and SJ8072 (from donor SJ8041). There was absolutely no growth of these deletion strains when reisolated onto LPBSG plates with 10 mikrogram/ml chloramphenicol, whereas there was some growth of a control strain reisolated onto the same plates.

Example 2 – *cat* gene deletion in the ATCC14580 strain

A few attempts were made to introduce the *cat*-deletion into the ATCC14580 chromosome like it was done in the above, via integration and excision of a deletion plasmid introduced via conjugation from SJ8039-SJ8042. We obtained transconjugants and subsequently strains with integrated plasmids, but could not isolate the desired excision derivatives carrying the *cat* deletion. As this might be due to the sequence differences between the two different *B. lichenformis* strains in the *cat* gene region, a new deletion vector was constructed.

Construction of deletion plasmid

A boiled cell suspension of ATCC14580 was used as template in PCR reactions with primer combination #450168 + #450169 (to give a fragment 3' of *cat*), and with primer combination #450170 + #450171 (to give a fragment 5' of *cat*), using 57 °C as annealing temperature.

The 3' fragment was digested with *EcoRI* + *SalI* and ligated to *EcoRI* + *SalI* digested pUC19 vector, and the ligation mixture was transformed into *E. coli* SJ2 by electroporation. Transformants with the correct sequence of the cloned DNA were kept as SJ8151 (SJ2/pSJ8151) and SJ8152 (SJ2/pSJ8152).

The 5' fragment was digested with *SalI* + *HindIII* and ligated to *SalI* + *HindIII* digested pUC19 vector, and the ligation mixture transformed into *E. coli* SJ2 by electroporation. Transformants with the correct sequence of the cloned DNA were kept as SJ8153 (SJ2/pSJ8153) and SJ8154 (SJ2/pSJ8154).

The 5'*cat* – 3'*cat* segment was assembled in a temperature sensitive, mobilizable integration vector by excision of the 0.5 kb *EcoRI* – *SalI* fragment from pSJ8151, and the 0.5 kb *SalI*-*HindIII* fragment from pSJ8153, and ligation of these two fragments to the 4.4 kb *EcoRI*-*HindIII* fragment of pSJ2739. Correct transformants of *B. subtilis* DN1885 were kept as SJ8209 (DN1885/pSJ8209) and SJ8210 (DN1885/pSJ8210).

Construction of *B. licheniformis* ATCC 14580 strain with a deletion in the *cat* gene

B. subtilis conjugation donor strain PP289-5 was transformed with the integration vectors, resulting in SJ8221 and SJ8222 (with pSJ8209), and in SJ8223 and SJ8224 (with pSJ8210).

B. licheniformis strain ATCC14580 was used as recipient in conjugations with the above donor strains. Conjugations were performed as previously described, selecting erythromycin resistance.

Tetracycline sensitive transconjugants were isolated, and chloramphenicol and erythromycin sensitive strains isolated following integration and excision of the plasmid, as previously described. Two such strains were kept as SJ8343 (from donor SJ8222) and SJ8344 (from donor SJ8223).

Example 3 - Test of chloramphenicol resistance

Strains PP1897-3, SJ8071 (= PP1897-3 delta *cat*), ATCC14580, and SJ8344 (= ATCC14580 delta *cat*) were propagated in 10 ml TY (WO 94/14968, p16) supplemented with 0.5 % glucose at 30°C overnight, with shaking. 50 microliter aliquots were used to inoculate further 10 ml TY + glucose cultures, supplemented with chloramphenicol (cam) to concentrations: 1 microgram/ml, 2 microgram/ml, 3 microgram/ml, 6 microgram/ml, 9 microgram/ml, or 12 microgram/ml, and incubation with shaking continued at 30°C. Growth was scored by visual inspection of turbidity after approximately 7 hours (table 1), and again after overnight incubation (table 2).

Table 1. After 7 hours incubation: ++ good growth, + weak growth, - no growth.

Strain	1 µg/ml cam	2 µg/ml cam	3 µg/ml cam	6 µg/ml cam	9 µg/ml cam	12 µg/ml cam
PP1897-3	++	++	+	-	-	-
SJ8071	+	-	-	-	-	-
ATCC14580	++	+	+	-	-	-
SJ8344	+	-	-	-	-	-

Table 2. After overnight incubation: ++ good growth, + weak growth, - no growth.

Strain	1 µg/ml cam	2 µg/ml cam	3 µg/ml cam	6 µg/ml cam	9 µg/ml cam	12 µg/ml cam
PP1897-3	++	++	++	++	++	+

SJ8071	++	++	++	+	+	+
ATCC14580	++	++	++	++	++	+
SJ8344	++	++	++	+	+	+

It is apparent that deletion of the putative chloramphenicol resistance gene, as done in strains SJ8071 and SJ8344, significantly decreases the chloramphenicol resistance of the strains.

5

Example 4 - Use of chloramphenicol resistance for transconjugant selection

Construction of test plasmid with chloramphenicol and erythromycin resistance genes

To investigate the potential use of chloramphenicol for transconjugant selection, use was made of plasmid pSJ7976. This plasmid is based on the mobilizable, temperature-sensitive vector backbone pSJ2739 mentioned previously, and contains the erythromycin resistance gene from pE194. It contains, in addition, the coding sequence of the pC194 *cat* gene inserted between segments, that flank the *B. licheniformis* C-component protease (U.S. Patent No. 5,459,064, accession no. D10060, Kakudo *et al.*, 1992, *J. Biol. Chem.* 267: 23782) gene 5' and 3'. It was constructed as follows:

15 A fragment 5' of the C-component protease gene was made by PCR amplification using chromosomal DNA from PL1980 as template, and an annealing temperature of 57°C with the following primers:

SEQ ID NO: 15 (#448697): 5' gactaagcttagatcttcacttccttattttgttgaagta

SEQ ID NO: 16 (#448698): 5' gactgaattcgtcgatcacttctgccactc

20

The PCR amplified fragment was digested with *EcoRI* + *HindIII*, ligated to *EcoRI* + *HindIII* digested pUC19, and the ligation mixture was transformed, by electroporation, into *E. coli* SJ2. Two correct transformants were kept as SJ7867 (SJ2/pSJ7867) and SJ7868 (SJ2/pSJ7868).

25 A fragment 3' of the C-component protease gene was made by PCR amplification using chromosomal DNA from PL1980 as template, and an annealing temperature of 57°C with the following primers:

SEQ ID NO: 17 (#448699): 5' gactgaattcagatctgctagcacgcgtgcggccgcacgaagacagcccgttc

SEQ ID NO: 18 (#448700): 5' gactaagcttcataaattctcgatacaacac

30

The PCR amplified fragment was digested with *EcoRI* + *HindIII*, ligated to *EcoRI* + *HindIII* digested pUC19, and the ligation mixture transformed, by electroporation, into *E. coli*

SJ2. Two correct transformants were kept as SJ7869 (SJ2/pSJ7869) and SJ7870 (SJ2/pSJ7870).

The C-component 5'- and 3'-flanking fragments were assembled in a temperature-sensitive, mobilizable integration vector by excision of the 0.5 kb EcoRI-BglII fragment from pSJ7867 and the 0.5 kb BglII-HindIII fragment from pSJ7869, and ligation of these fragments to the 4.4 kb EcoRI-HindIII fragment of pSJ2739. The ligation mixture was transformed into competent cells of *B. subtilis* DN1885, selecting erythromycin resistance (5 microgram/ml) at 30°C. Two correct transformants were kept as SJ7909 (DN1885/pSJ7909) and SJ7910 (DN1885/pSJ7910).

The chloramphenicol resistance gene of pC194 was amplified by PCR from plasmid pDN1050 (Diderichsen et al., 1993, Plasmid 30, 312-315), using an annealing temperature of 57°C with primers:

SEQ ID NO: 19 (#448716): 5' gactggatccatgaactttaataaaattgatttagac

SEQ ID NO: 20 (#448718): 5' gactgaattcgctagcacgcgtataaaagccagtcattaggcc

The PCR amplified fragment was digested with BamHI + EcoRI, ligated to BamHI + EcoRI digested pUC19, and the ligation mixture transformed, by electroporation, into *E. coli* SJ2. A correct transformant was kept as SJ7887 (SJ2/pSJ7887).

The pC194 *cat* gene was inserted between the 5' and 3' C-component flanking segments, by excision of the 0.7 kb BamHI-MluI fragment from pSJ7887 and ligation of this to the 5.4 kb BglII-MluI fragment of pSJ7909. The ligation mixture was transformed into competent *B. subtilis* DN1885 cells, selecting erythromycin (2 microgram/ml) and chloramphenicol (6 microgram/ml) resistance at 30°C. Two transformants were kept, SJ7975 (DN1885/pSJ7975) and SJ7976 (DN1885/pSJ7976).

Plasmid introduction by conjugation

The test plasmid, pSJ7976, which contains both chloramphenicol and erythromycin resistance genes, was transformed into the conjugative donor strain *B. subtilis* PP289-5, resulting in strain SJ8000.

B. licheniformis strains PP1897-3, the *cat*-deleted derivative SJ8071, strain ATCC14580, and the *cat*-deleted derivative SJ8344 were used as recipients in conjugations with donor strain SJ8000, essentially as described before.

For each donor/recipient combination, donor and recipient colonies were scraped off overnight plates and mixed in a small volume of liquid TY medium, and equal volumes from this mixture was plated on each of two LBPGS plates with D-alanine. Following 6 hours incubation at 30°C, one plate was replicaplated onto LBPGS with erythromycin (2 microgram/ml), and the other plate replica plated onto LPPGS with chloramphenicol (6

microgram/ml). These plates were incubated at 30°C overnight, then left at room temperature (20-25 °C) for further 3 days.

After the overnight incubation, distinct colonies were observed on the plates with erythromycin for all donor/recipient combinations. There were approximately equal numbers on the plates with ATCC14580 and SJ8344, and there were about twice the number of colonies on the plates with PP1897-3 as compared to SJ8071.

The result using chloramphenicol selection was quite different. No distinct transconjugant colonies were observed with the two recipient strains containing intact, chromosomal chloramphenicol resistance genes, but the replica plates contained an even smear of cells.

Distinct transconjugant colonies were observed with the *cat*-deleted recipient strains. The number of transconjugants were the same when using chloramphenicol selection as when using erythromycin selection, for the SJ8071 recipient as well as for the SJ8344 recipient.

From each of the plates that contained distinct transconjugant colonies, whether by erythromycin or by chloramphenicol selection, 10 colonies were transferred to new plates with chloramphenicol or erythromycin, and all colonies were found to grow on both types of plates.

The use of *cat*-deleted recipient strains thus allows the successful use of chloramphenicol resistance as selection for DNA introduction into *B. licheniformis*.

Example 5 - DNA sequence of the *cat*-gene of *B.licheniformis* 9945A

Bacillus licheniformis strain 9945A (Strain 5A2 from the Bacillus Genetic Stock Center; J. Bacteriology (1954) 68:307.) was obtained from Curtis B. Thorne and kept as JA102.

The chromosomal *cat*-gene region was PCR amplified using primers #450168 + #450171, and the amplified fragment DNA was sequenced using the same primers. The resulting sequence is shown in SEQ ID NO: 21, and the encoded protein in SEQ ID NO: 22.

Example 6 – DNA sequence of the *cat*-gene DNA in yet another *B. licheniformis* strain

International patent publication WO 03/083125 (Genencor Int. Inc) discloses the DNA and protein sequences of a chromosomal chloramphenicol resistance gene, on pages 21 and 22 of the text, with the DNA sequence given as SEQ ID NO:58 and the protein sequence given as SEQ ID NO:59 of that publication. The two sequences are also provided herein as SEQ ID NO: 23 and SEQ ID NO: 24, respectively.

Example 7 – % Identity alignments

The four different DNA sequences of the *cat* genes: *cat* ATCC14580, *cat* PL1980, *cat* 9945A, and *cat* WO 03/083125, were aligned in table 3. The matrix was calculated based on

"all against all" alignments of the supplied sequences. The entry in row i and column j in the matrix is calculated as the number of exact matches in the alignment between sequence i and sequence j divided by the total length of the alignment minus the total length of the gaps in the alignment. Each alignment was done using the needle program from the EMBOSS package version 2.8.0. The program needle implements the global alignment algorithm described in the two references:

Needleman, S. B. and Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453.

Kruskal, J. B. (1983) An overview of sequence comparison In D.

Sankoff and J. B. Kruskal, (ed.), Time warps, string edits and macromolecules: the theory and practice of sequence comparison, pp. 1-44 Addison Wesley.

The alignments used the following parameters:

Gap opening penalty: 10.00

Gap extension penalty: 0.50

Substitution matrix: EBLOSUM62

Table 3. Alignment of the four *Bacillus licheniformis* cat-genes.

%-Identity	ATCC14580	PL1980	9945A	WO 03/083125
ATCC14580	100.00	96.30	93.67	99.67
PL1980		100.00	92.59	96.33
9945A			100.00	93.32
WO 03/083125				100.00

The four different protein sequences encoded by the above DNA sequences were also aligned as shown in table 4, using the same software and settings.

Table 4. Identity-alignment of the four proteins encoded by the DNA in table 3.

%-Identity	ATCC14580	PL1980	9945A	WO 03/083125
ATCC14580	100.00	95.37	92.13	94.39
PL1980		100.00	90.28	90.82
9945A			100.00	87.24
WO 03/083125				100.00

CLAIMS

1. A method for introducing a recombinant DNA construct into a modified *Bacillus licheniformis* host cell, the method comprising:
 - a) providing a modified *Bacillus licheniformis* host cell, wherein one or more endogenous
5 genomic chloramphenicol acetyl transferase encoding gene(s) has been inactivated;
 - b) introducing a recombinant DNA construct comprising a selection marker encoding a chloramphenicol acetyl transferase into the modified host cell;
 - c) cultivating the host cell of step (b) in a growth-medium comprising an inhibitory concentration of chloramphenicol; and
 - 10 d) selecting a host cell comprising the DNA construct that is capable of growing in the medium of step (c).
2. The method of claim 1, wherein the recombinant DNA construct is a plasmid or a linearized plasmid, preferably a linearized and concatamerized plasmid.
- 15 3. The method of claim 1 or 2, wherein the recombinant DNA construct further comprises a polynucleotide encoding a polypeptide of interest, preferably an enzyme.
4. The method of any of claims 1 - 3, wherein the one or more genomic chloramphenicol
20 acetyl transferase gene(s) encodes a chloramphenicol acetyl transferase comprising an amino acid sequence which is at least 80% identical to the sequence shown in SEQ ID NO: 2, SEQ ID NO: 13, SEQ ID NO: 22, or SEQ ID NO: 24.
5. The method of any of claims 1 - 4, wherein the one or more genomic chloramphenicol
25 acetyl transferase gene(s) comprises a polynucleotide sequence which is at least 80% identical to the sequence shown in SEQ ID NO: 1, SEQ ID NO: 12, SEQ ID NO: 21, or SEQ ID NO: 23.
6. The method of any of claims 1 - 5, wherein the selection marker of the DNA construct
30 encodes a chloramphenicol acetyl transferase comprising an amino acid sequence at least 80% identical to the sequence shown in SEQ ID NO: 2, SEQ ID NO: 13, SEQ ID NO: 22, or SEQ ID NO: 24.
7. The method of any of claims 1 - 6, wherein the selection marker of the DNA construct
35 is flanked by resolvase-sites.

8. The method of any of claims 1 – 7, wherein the DNA construct further comprises a polynucleotide region that is sufficiently homologous with a genomic DNA region of the host cell to effect integration into the genome of the DNA construct by homologous recombination.
- 5 9. The method of any of claims 1 – 8, wherein the inhibitory concentration of chloramphenicol in the growth-medium is at least 1 microgram/ml, preferably at least 2 microgram/ml, more preferably at least 3 microgram/ml, 6 microgram/ml, 9 microgram/ml, or at least 12 microgram/ml.
- 10 10. A modified *Bacillus licheniformis* host cell, wherein one or more genomic chloramphenicol acetyl transferase gene(s) has been inactivated, comprising a recombinant DNA construct carrying a selection marker encoding a chloramphenicol acetyl transferase.
11. The host cell of claim 10, wherein the recombinant DNA construct is a plasmid or a
15 linearized plasmid, preferably a linearized and concatamerized plasmid.
12. The host cell of claim 10 or 11, wherein the recombinant DNA construct further comprises a polynucleotide encoding a polypeptide of interest, preferably an enzyme.
- 20 13. The host cell of any of claims 10 - 12, wherein the one or more genomic chloramphenicol acetyl transferase gene(s) encodes a chloramphenicol acetyl transferase comprising an amino acid sequence which is at least 80% identical to the sequence shown in SEQ ID NO: 2, SEQ ID NO: 13, SEQ ID NO: 22, or SEQ ID NO: 24.
- 25 14. The host cell of any of claims 10 – 13, wherein the one or more genomic chloramphenicol acetyl transferase gene(s) comprises a polynucleotide sequence which is at least 80% identical to the sequence shown in SEQ ID NO: 1, SEQ ID NO: 12, SEQ ID NO: 21, or SEQ ID NO: 23.
- 30 15. The host cell of any of claims 10 - 14, wherein the selection marker of the DNA construct encodes a chloramphenicol acetyl transferase comprising an amino acid sequence at least 80% identical to the sequence shown in SEQ ID NO: 2, SEQ ID NO: 13, SEQ ID NO: 22, or SEQ ID NO: 24.
- 35 16. The host cell of any of claims 10 – 15, wherein the selection marker of the DNA construct is flanked by resolvase-sites.

17. The host cell of any of claims 10 – 16, wherein the DNA construct further comprises a polynucleotide region that is sufficiently homologous with a genomic DNA region of the host cell to effect integration into the genome of the DNA construct by homologous recombination.

Figure 1. Schematic drawing of the *cat* gene region from ATCC14580.

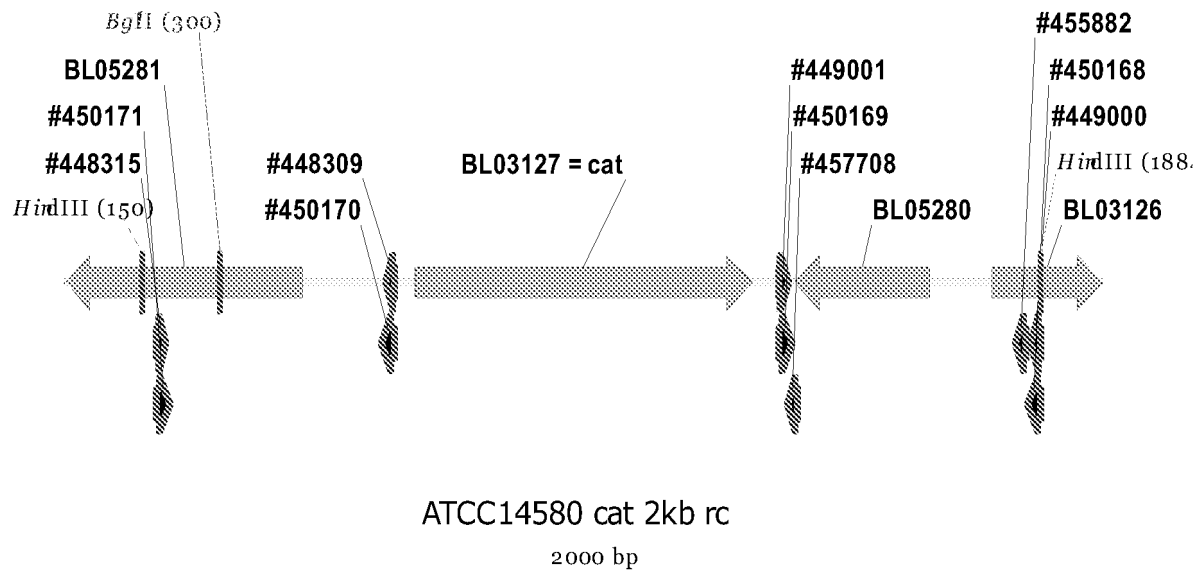


Figure 2. Schematic drawing of the *cat* gene region from PL1980.

