

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
27 December 2002 (27.12.2002)

PCT

(10) International Publication Number
WO 02/103010 A1

(51) International Patent Classification⁷: **C12N 15/10**,
15/65, 15/76, 15/90, 1/21 // C12R 1:465

(21) International Application Number: PCT/GB02/02798

(22) International Filing Date: 14 June 2002 (14.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0114535.8 14 June 2001 (14.06.2001) GB
0200477.8 9 January 2002 (09.01.2002) GB

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): **PLANT BIOSCIENCE LIMITED** [GB/GB]; Norwich Research Park, Colney Lane, Norwich, Norfolk NR4 7UH (GB).

Declaration under Rule 4.17:

— *of inventorship (Rule 4.17(iv)) for US only*

(72) Inventors; and

Published:

— *with international search report*

(75) Inventors/Applicants (*for US only*): **GUST, Bertolt** [DE/GB]; 75 Alexandra Road, Norwich, Norfolk NR2 3EA (GB). **CHATER, Keith, Frederick** [GB/GB]; 6 Coach House Court, Norwich, Norfolk NR4 7QR (GB). **KIESER, Tobias, Edwin** [GB/GB]; 7 The Hedgerows, Norwich, Norfolk NR5 9BP (GB).

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

(74) Agents: **DENISON, Christopher, M.** et al.; Mewburn Ellis, York House, 23 Kingsway, London, Greater London WC2B 6HP (GB).

(54) Title: METHODS AND MATERIALS FOR TARGETED GENE DISRUPTION IN ACTINOMYCETE BACTERIA

(57) **Abstract:** A method is provided for generating a targeted nucleic acid disruption in an actinomycete, the method comprising: (1) providing a first nucleic acid construct comprising: (a) a nucleic acid sequence (the "marker sequence") encoding a marker; (b) a nucleic acid sequence (the "origin sequence") comprising an origin of transfer; (c) a pair of nucleic acid sequences (the "recombining sequences") which are capable of undergoing recombination with each other, thereby to eliminate nucleic acid lying between them and leave only a single recombining sequence; and (d) a pair of nucleic acid sequences (the "targeting sequences") capable of undergoing homologous recombination with a nucleic acid of interest of said actinomycete, wherein parts (a) and (b) lie between the pair (c) of recombining sequences and parts (a), (b) and (c) lie between the pair of targeting sequences; (2) recombining the first nucleic acid construct with a first plasmid bearing at least part of the actinomycete target nucleic acid of interest to form a "recombined plasmid"; (3) transferring the recombined plasmid by conjugation from a first bacterial cell into an actinomycete bacterial cell; (4) integrating nucleic acid of the first nucleic acid construct from the transferred plasmid into the target actinomycete nucleic acid by homologous recombination; and (5) eliminating the origin sequence and marker sequence from the integrated nucleic acid, to leave an insertion in the actinomycete nucleic acid, the insertion having a single recombining sequence of the original pair of recombining sequences. Also provided are related materials, methods for producing them, and libraries of cells having such targeted nucleic acid disruptions.



WO 02/103010 A1

Methods and Materials for Targeted Gene Disruption in
Actinomycete Bacteria

The present invention relates to methods and materials
5 for generating targeted gene disruptions, especially
multiple targeted gene disruptions, in actinomycete
bacteria, especially *Streptomyces* spp..

The generation of gene disruptions in actinomycetes poses
10 several problems. For example, like most bacteria,
actinomycetes are not readily transformable with linear
DNA because of the presence of intracellular exonucleases
that degrade linear DNA.

15 Because of this, to inactivate chromosomal genes of
S.coelicolor, a gene disruption is usually created on a
suitable plasmid in *E.coli* before introducing the
construct into *S.coelicolor* for recombination with the
chromosome. Alternatively single crossover integration of
20 a phage vector or suicide plasmids has been used (Kieser
et al., 2000).

Moreover, the generation of multiple knock-out mutants
tends to require the use of a corresponding number of
25 different selectable markers, leading to undesirable
complication, e.g. in the preparation of suitable
culturing media.

These problems have, however, been addressed to some
30 extent in *E.coli*. The λ RED (*gam*, *bet*, *exo*) functions
promote a greatly enhanced rate of recombination when
using linear DNA. By exploiting this, Datsenko and Wanner
made 13 different disruptions on the *E.coli* chromosome by
replacing a chromosomal sequence with a selectable marker

generated by PCR using primers with 36nt homology extensions. After selection, the resistance gene can also be eliminated by using a helper plasmid expressing the FLP recombinase, which acts on directly repeated FRT (FLP
5 recognition target) sites flanking the resistance gene leaving a remaining FRT site (Datsenko and Wanner, 2000).

It would be desirable to apply a similar approach to actinomycetes such as *S.coelicolor*.

10

The present invention is intended to address this and other problems associated with the generation of targeted gene disruptions in actinomycete bacteria.

15 Accordingly, in a first general aspect, the present invention provides a nucleic acid construct comprising:

- (a) a nucleic acid sequence (the "marker sequence") encoding a marker, preferably a selectable marker;
- 20 (b) a nucleic acid sequence (the "origin sequence") comprising an origin of transfer;
- (c) a pair of nucleic acid sequences (the "recombining sequences") which are capable of undergoing recombination with each other,
25 thereby to eliminate nucleic acid lying between them and leave only a single recombining sequence; and
- (d) a pair of nucleic acid sequences (the "targeting sequences") capable of undergoing
30 homologous recombination with a gene of interest,

wherein parts (a) and (b) lie between the pair (c) of recombining sequences and parts (a), (b) and (c) lie between the pair of targeting sequences, such that

following homologous recombination of the targeting sequences with a target gene and recombination of the recombining sequences, a single recombining sequence is left as an insertion in the target gene, and the marker
5 sequence and origin sequence are eliminated.

Preferably the construct further comprises (e) a pair of nucleic acid sequences (the "intervening sequences") which flank the recombining sequences and which are
10 flanked by the targeting sequences. It is however contemplated that in some embodiments only one intervening sequence may be present (i.e. one end of the construct has an intervening sequence between a
15 other has only adjacent targeting and recombining sequences). Following the above-mentioned recombination events, the intervening sequence(s) (where present) will also form part of the insertion in the target gene.

20 The term "flank" as used herein is not intended to imply that the sequences concerned are immediately adjacent, though this is generally preferred.

In certain preferred embodiments, the insertion comprises
25 a stop codon in at least one reading frame in a forward or backward direction. More preferably, it comprises a stop codon in each of the three forward or in each of the three backward reading frames. Still more preferably it comprises a stop codon in each of the six forward and
30 backward reading frames. It is however contemplated that in some embodiments it may be desirable to encode an in-frame amino acid insertion without terminating the encoded polypeptide.

Preferably the insertion consists of a multiple of three nucleotides, especially where it is desirable to encode an in-frame amino acid insertion without terminating the encoded polypeptide.

5

As will be evident from the foregoing, an intervening sequence may be used to provide an insertion into a gene of interest, e.g. to modify the upstream, non-coding regions of the gene, or to create a gene encoding a fusion protein.

10

In particular, an intervening sequence placed downstream of the recombining sequences may be targeted to the upstream non-coding region of a gene, e.g. to insert a promoter sequence upstream of the coding sequence. In this case, the targeting sequences of the construct would be based on the sequence of the upstream non-coding region of the gene (e.g. the existing promoter sequence of the gene) and optionally (for the downstream targeting sequence) the start of the coding sequence of the gene. After recombination with the target gene and elimination of the marker sequence and origin sequence, the intervening sequence will form part of the insertion in the target gene, downstream of the single remaining recombining sequence, but upstream of the coding region of the gene. This is particularly suited to inserting exogenous control sequences upstream of the coding region, e.g. promoter and/or enhancer or repressor sequences.

25
30

Alternatively, an intervening sequence placed upstream of the recombining sequences may be targeted to the 3' end of the coding region of a gene, e.g. to alter the 3' coding sequence of the gene, particularly to encode a

fusion protein. In this case, the targeting sequences of the construct would be based on the sequence of the 3' end of the coding region of the gene and (for the downstream targeting sequence) the start of the 3' non-coding sequence of the gene. After recombination with the target gene and elimination of the marker sequence and origin sequence, the intervening sequence will form part of an insertion in the 3' end of the target gene, upstream of the single remaining recombining sequence. Preferably such an insertion will replace the normal stop codon of the target gene with a codon for an amino acid, allowing translation to continue into the intervening sequence. Preferably the intervening sequence and/or the recombining sequence provides a stop codon, however, after the desired insertion.

Accordingly, the term "gene" as used herein is not to be interpreted as limited to a coding sequence, but also includes upstream and downstream non-coding regions. Also, the term "disruption" is not to be interpreted as limited to the prevention of functional expression of a gene, but also includes for example the inclusion of exogenous control sequences and/or exogenous coding sequences into the coding or non-coding regions of the target gene.

A sequence encoding any suitable marker (preferably a selective marker) may be used; antibiotic resistance markers are preferred.

The origin sequence is discussed further below.

The recombining sequences are preferably FRT (FLP recognition target) sequences, which may be recombined by

FLP recombinase activity *in trans* (e.g. provided by a helper plasmid, e.g. as described in the Examples). The FLP recombinase system is well known per se to those skilled in the art, e.g. from Hoang et al. (1998).

- 5 Preferably each FRT sequence comprises a palindrome (preferably an interrupted palindrome), having the palindromic half site GAAGTTCCTAT. Preferably the interruption is by a sequence of about 12 nucleotides (e.g. 8-16 nucleotides). Preferably each FRT sequence
- 10 comprises the double stranded core sequence:

5'-GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC-3'

3'-CTTCAAGGATATGAAAGATCTCTTATCCTTGAAG-5'

- However, any other suitable recombination system may also
- 15 find use in the present invention.

The targeting sequences are also discussed further below.

- Such a nucleic acid construct finds use in generating
- 20 targeted gene disruptions in actinomycete bacteria, e.g. using the methods of the invention.

- In a second aspect, the invention provides a method for generating targeted gene disruptions in an actinomycete,
- 25 the method comprising:

- (1) providing a nucleic acid construct according to the first aspect;
- (2) recombining the nucleic acid construct with a plasmid bearing at least part of an actinomycete
- 30 target gene of interest to form a "recombined plasmid";
- (3) transferring the recombined plasmid by conjugation from a first bacterial cell into an actinomycete bacterial cell;

(4) integrating nucleic acid of the nucleic acid construct from the transferred plasmid into the target chromosomal actinomycete gene by homologous recombination; and

5 (5) eliminating the origin sequence and marker sequence from the integrated nucleic acid, to leave an insertion in the actinomycete gene, the insertion having a single recombining sequence of the original pair of recombining sequences.

10

Step (1) may comprise providing the nucleic acid construct by homology extension PCR from a plasmid containing parts (a), (b), (c) and preferably (e) of the nucleic acid construct of the first aspect, the PCR
15 primer (or its complementary strand) providing part (d) and optionally (e). A nucleic acid construct comprising parts (a), (b), (c) and preferably (e) as defined in the first aspect is a further aspect of the invention, as is the above method of producing a construct according to
20 the first aspect.

The 3' end of each primer will be designed to hybridise with a respective recombining sequence and/or, where present, intervening sequence, and the 5' end will be
25 designed to provide a sequence which, when double stranded, is capable of homologous recombination with the target gene. The 5' ends of the primers therefore generate the targeting sequences of the resultant construct. The primers may also generate the intervening
30 sequence(s), where present. The generation of the necessary sequences may occur in multiple PCR steps, in which case, the primers may individually provide incomplete parts of the sequences identified.

The skilled person will be able to determine an appropriate size for the 5' ends of the homology extension PCR primers (i.e. for the targeting sequences of the nucleic acid construct). Preferably the size of
5 the 5' end of each homology extension primer will be at least 20 nucleotides, more preferably at least 25, 30 or 35 nucleotides, still more preferably at least 40 nucleotides.

10 Preferably step (2) comprises recombining the nucleic acid construct with nucleic acid from an actinomycete cosmid or BAC library, though recombination with actinomycete nucleic acid cloned on any suitable plasmid is contemplated. Most preferred is the *S.coelicolor*
15 cosmid of Redenbach et al., 1996.

Preferably the first bacterial cell is an *E.coli* cell. Where it is intended later to conjugate the plasmid into an actinomycete cell possessing a potent methyl-specific
20 restriction system (*S.coelicolor* has such a system, although the related strain *S.lividans* does not, MacNeil et al., 1992), the first bacterial cell is preferably methylation deficient, e.g. *E.coli* strain ET12567 (MacNeil et al., 1992).

25 The nucleic acid construct is preferably linear. Especially when that is the case, step (2) preferably occurs in a bacterial cell (preferably an *E.coli* cell) having λ RED (*gam*, *bet*, *exo*) function, to enhance the
30 rate of recombination by reducing degradation of linear nucleic acid by bacterial intracellular exonucleases.

The nucleic acid construct preferably does not share sufficient sequence homology for homologous recombination

with the plasmid, except in the targeting sequences
(which share homology with the target actinomycete gene).
This assists in preventing erroneous recombination, i.e.
recombination into the plasmid, but outside the target
5 gene.

Prior to step (3), the plasmid may be pre-existing in the
first bacterial cell or may be introduced into the first
bacterial cell following purification from another
10 source.

The provision of the origin sequence in the nucleic acid
construct allows the introduction of the recombined
plasmid into the actinomycete cell by conjugation. The
15 main advantages in using conjugation (especially from
E.coli) as a means of introducing DNA into actinomycetes
are the simplicity of the procedure because it does not
rely on protoplast formation and regeneration, and its
broad applicability in introducing DNA into actinomycetes
20 other than specifically *S.coelicolor* (Matsushima *et al.*,
1994).

Preferably the origin sequence is an *oriT* sequence which
can be mobilised by the helper (or "transfer") plasmids
25 pUZ8002 and pUB307, such as an *oriT* sequence from an
IncP-group plasmid, such as RP4 (also designated RP1/RK2;
Pansegrau *et al.*, 1994), preferably having the nucleic
acid sequence shown in Fig. 1B, residues 1450-1558, or a
variant thereof. However, the use of any other suitable
30 origin of transfer is also contemplated.

For transfer of (in particular) plasmids bearing an RP4
oriT to occur by conjugation, a transfer function should
be supplied *in trans*, e.g. by an *E.coli* donor strain such

as ET12567 carrying the self-transmissible pUB307 (Bennett et al., 1977, Flett et al., 1997) or ET12567 carrying the non-transmissible pUZ8002 (Kieser et al., 2000).

5

Thus step (3) preferably includes such provision *in trans* of transfer function. This step may involve transforming the recombined plasmid into a donor strain carrying a non-transmissible transfer plasmid (e.g.

10 ET12567/pUZ8002), followed by incubation under suitable conditions with the actinomycete cell (e.g. a pre-germinated actinomycete spore).

Alternatively, in a modified version of the conjugation
15 procedure described by Kieser et al., 2000, step (3) may involve a "triparental"-like conjugation in which the first bacterial cell (which may be a cosmid cell in which recombination of the plasmid and construct has previously occurred), a donor strain carrying a non-transmissible
20 transfer plasmid (e.g. ET12567/pUZ8002), a donor strain carrying a self-transmissible transfer plasmid (e.g. ET12567/pUB307) and the actinomycete cell (e.g. a pre-germinated actinomycete spore) are all incubated together under suitable conditions (for which see Kieser et al.,
25 2000 and the examples). In this "triparental" procedure the recombined plasmid is firstly mobilised by the donor strain carrying the self-transmissible transfer plasmid (e.g. ET12567/pUB307) into the donor strain carrying the non-transmissible transfer plasmid (e.g. ET12567/pUZ8002)
30 and then is subsequently mobilised by the latter into the actinomycete cell.

It is thought that direct mobilisation of the plasmid by pUB307 may occur to a small extent, but that this method

of conjugation is inefficient (estimated at less than 1%), owing to the absence of selection in the actinomycete cell.

- 5 The steps of the method need not occur in the order shown above. In particular, step (5) may precede steps (3) and (4). However, it is preferred that step (5) follows steps (3) and (4), as the presence of the marker sequence (from the nucleic acid construct) in the actinomycete
10 cell aids confirmation that the transferring and integrating steps have successfully occurred. Elimination of the marker sequence (and also the origin sequence) may be confirmed by replica plating of exconjugants, elimination being reflected by lack of
15 marker function (e.g. antibiotic susceptibility).

- Where step (5) follows steps (3) and (4), the elimination of the origin and marker sequences may occur by FLP recombinase-mediated recombination, leaving a single FRT
20 sequence (and, where present, the intervening sequence(s)) as an insertion in the target gene.

- FLP recombinase activity is preferably provided *in trans* using a plasmid that shows inducible (e.g. temperature
25 sensitive) replication and induction of FLP synthesis under conditions in which replication will not occur (e.g. at high temperature). Such a plasmid allows for simultaneous induction of FLP synthesis and loss of the plasmid. Preferred plasmids are BT340, as described in
30 the Examples, and pFLP1 from Hoang et al. (1998).

Loss of FLP recombinase activity allows the process to be repeated to target disruptions in another gene of the same actinomycete cell.

However, an alternative method has been developed for eliminating the origin sequence and marker sequence following steps (3) and (4). The method may also make
5 use of the FLP recombinase system, though in a different way from that described above.

This method makes use of a helper plasmid which contains a nucleic acid construct comprising targeting sequences
10 for the same actinomycete gene as the first plasmid, optionally the intervening sequence(s), a single recombining sequence (since the elimination step (5) has already occurred), and no origin sequence between the targeting sequences. Nucleic acid from such a helper
15 plasmid may be integrated into the actinomycete gene, following previous integration of nucleic acid from the first plasmid. Integration from the first plasmid results in an insert in the actinomycete chromosomal gene, the insert comprising the entire nucleic acid
20 construct as defined in the first aspect. Subsequent integration from the helper plasmid into the same gene replaces that insert with an insert having only a single recombining sequence and (optionally) intervening sequence(s). The two plasmids must therefore be capable
25 of homologous recombination with the actinomycete gene at the same site.

In part, the helper plasmid is preferably constructed by a method comprising steps (1), (2) and (5) as shown
30 above. To allow efficient introduction into the actinomycete cell, however, the helper plasmid also comprises a second origin sequence outside the construct defined above, which therefore neither forms part of the insertion into the actinomycete chromosomal gene,

- following the integration step, nor is eliminated if and when the first origin sequence and marker sequence are eliminated in the construction of the helper plasmid. The second origin sequence may be introduced into the
- 5 helper plasmid following elimination of the first origin sequence. Preferably the helper plasmid also comprises a second marker sequence (preferably encoding a selectable marker).
- 10 The second origin sequence and second marker sequence are preferably introduced by recombination between a helper nucleic acid construct and the helper plasmid, the helper construct comprising the second origin sequence and
- 15 second marker sequence in between two sequences (the "helper construct targeting sequences") which share homology with sequences of the helper plasmid. The helper construct targeting sequences will be chosen so that introduction of the helper construct into the helper plasmid does not interfere with the targeting sequences
- 20 of the helper plasmid (which allow integration of the helper plasmid into the actinomycete gene).

In a particularly preferred embodiment, the first plasmid and helper plasmid are derived from the same basic

25 plasmid (preferably a plasmid of an actinomycete library or cosmid). This will allow convenient selection of helper construct targeting sequences. Preferably the first plasmid and helper plasmid are derived from the same basic plasmid with which the same nucleic acid

30 construct of the first aspect of the invention has been recombined. This allows convenient integration of nucleic acid from the helper plasmid into the same site in the chromosomal actinomycete gene as the earlier integration of nucleic acid from the first plasmid.

As will be appreciated, the helper nucleic acid construct and helper plasmid, as well as methods for their production, all as defined above, form further aspects of the invention.

As will be further appreciated, the use of the helper plasmid to eliminate the origin and marker sequences from the actinomycete chromosomal gene falls within the definition of step (5) above.

An alternative nucleic acid construct (the "additional disruption construct") is also provided for the generation of additional disruptions, after a first disruption has been generated using a construct according to the first aspect of the invention. Such a construct has component features as defined for the construct of the first aspect, with the exception of an origin of transfer. The reason for this is that a plasmid which is to be conjugated into the actinomycete cell needs only one origin of transfer, which will be provided by recombination with a construct of the first aspect; further origins are unnecessary.

The additional disruption construct will comprise different marker and targeting sequences from the origin-containing construct with which it is intended to be used. Preferably, it will comprise the same recombining sequences (e.g. FLT sequences) as the origin-containing sequence, to allow elimination of the marker sequence of the additional disruption construct at the same time as the marker sequence and origin sequence of the construct of the first aspect are eliminated.

Accordingly, such constructs may be used in the method of the second aspect of the invention to introduce further disruptions into the actinomycete nucleic acid, by recombination with the "recombined plasmid" of step (2),
5 preferably before step (3) occurs.

The invention also provides vectors containing the nucleic acid constructs of the invention, flanked by suitable restriction sites for excising nucleic acid
10 constructs from the remainder of the vectors.

The invention also provides cells comprising the plasmids and vectors of the invention, and cells comprising gene disruptions as obtainable according to the methods of the
15 invention. Furthermore, the invention provides a method of determining the effect of a gene disruption, the method comprising culturing such a cell and determining the effect of the gene disruption on the cell.

20 The skilled person is well aware of suitable techniques for the preparation of such vectors and cells, e.g. from Sambrook et al. 1989, and Kieser et al, 2000.

In a further aspect, the present invention provides a
25 method for producing a library of actinomycete host cells having respective different gene disruptions, the method comprising:

repeating the methods of the previous aspects to generate different gene disruptions in different
30 respective host cells.

In this aspect, there is also provided such a library of actinomycete host cells having different respective gene disruptions as produced or as producible according to the

method of this aspect. In particular, the invention provides a library of actinomycete host cells respectively having insertions in different chromosomal genes. Preferably the insertions each comprise a single
5 FRT site, optionally with intervening sequence(s) also.

Such a library may be screened for phenotypic characteristics or changes resulting from the gene disruptions. Screening may be followed by the step of
10 identifying the disrupted gene. This may be followed by the step of cloning the undisrupted gene, e.g. from a cosmid used in the generation of the host cells bearing disruptions.

15 In a further aspect, the present invention provides a kit comprising a nucleic acid construct according to the first aspect of the invention (preferably without targeting sequences) and a helper nucleic acid construct or an additional disruption construct as defined above,
20 optionally with one or more other plasmids used in the methods described herein, e.g. a plasmid providing λ RED function, a self-transmissible plasmid providing transfer function, a non-transmissible plasmid providing transfer function and/or a plasmid providing FRP recombinase
25 function.

The actinomycete gene is preferably a native gene, though it is also contemplated that the materials and methods of the invention may be used for targeted disruption of
30 exogeneous genes in an actinomycete cell. As mentioned above, the term "gene" includes not only the coding region, but also upstream and downstream non-coding regions. Moreover, it will be evident that the present invention can also be applied to the generation of

disruptions in any actinomycete nucleic acid of interest, especially chromosomal nucleic acid. References herein to "gene" may therefore be replaced with "nucleic acid" within the scope of the invention.

5

Preferably the actinomycete cell is a streptomycete cell, more preferably a cell of the genus *Streptomyces*, still more preferably of the species *S.coelicolor* or *S.ambofaciens*, most preferably a cell of the strain *S.coelicolor* A3(2).

10

For the avoidance of doubt, the term "plasmid" is herein defined as a genetic element containing nucleic acid and able to replicate independently of its host's chromosome.

15

Generally, the plasmids of the invention will be circular, not linear, plasmids.

Variants

20

References herein to genes and nucleic acids are not to be interpreted as being restricted to genes and nucleic acids having the specific nucleic acid sequences disclosed herein. Rather, genes and nucleic acids having variants of those sequences are also included, provided that the functional interactions of the various components used in the methods and materials of the invention are maintained. Genes and nucleic acids having the specific sequences disclosed are preferred embodiments.

30

The term "variant" as used herein in relation to a particular nucleic acid (the reference nucleic acid) may denote: any nucleic acid having a sequence which is different from that of the reference nucleic acid, but

which is its complement or which shows significant nucleic acid sequence identity with, or hybridisation under stringent conditions to, the reference nucleic acid or its complement or a fragment of the reference nucleic acid or its complement.

Significant nucleic acid sequence identity is preferably at least 50%, more preferably 60%, 70%, 80% or 90%, still more preferably 95%, 98% or 99%.

10

Significant nucleic acid sequence identity is preferably shown between the variant nucleic acid (or a portion thereof) and a fragment of at least 30 residues of the reference nucleic acid, more preferably a fragment of a least 60, 90 or 120 residues, still more preferably the entire reference nucleic acid.

"Percent (%) nucleic acid sequence identity" is defined as the percentage of nucleotide residues in a candidate sequence that are identical with the nucleotide residues in the sequence under comparison, as determined by the BLASTN module of WU BLAST-2 (Altschul et al. (1996); <http://blast.wustl.edu/blast/README.html>), set to the default parameters, with overlap span and overlap fraction set to 1 and 0.125, respectively.

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/ 0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1%

polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

Embodiments of the invention will now be described, by way of example only.

15 Summary of Knock-Out Experiments

The basic strategy for PCR-targeting for mutagenesis of *Streptomyces coelicolor* is to replace a chromosomal sequence within a *S.coelicolor* cosmid (Redenbach et al., 1996) by a knock-out cassette containing a selectable marker. The knock-out cassette is generated by PCR using primers with 40nt homology extensions. Recombination occurs in *E.coli* containing the λ RED system, to enhance recombination efficiency.

25

An *oriT* from the IncP-group plasmid RP4 (also designated RP1/RK2; Pansegrau et al., 1994) within the knock-out cassette allows the introduction of the PCR targeted cosmid DNA into *S.coelicolor* by conjugation. The main advantages in using conjugation from *E.coli* as a means of introducing DNA into *Streptomyces* are the simplicity of the procedure (because it does not rely on the generation of protoplast formation and regeneration) and its broader

applicability in introducing DNA into other actinomycetes (Matsushima *et al.*, 1994).

5 The potent methyl-specific restriction system of *S.coelicolor* A3(2) is circumvented by passaging DNA through the methylation-deficient *E.coli* host ET12567 (MacNeil *et al.*, 1992).

10 To enable transfer of plasmids by conjugation, a transfer function is supplied *in trans* by the *E.coli* donor strain ET12567 carrying the self-transmissible pUB307 (Bennett *et al.*, 1977, Flett *et al.*, 1997) or ET12567 carrying the non-transmissible pUZ8002 (Kieser *et al.*, 2000).

15 To avoid having to transform mutagenised cosmid DNA (recombinant cosmid containing the knock-out cassette including *oriT*) into ET12567/pUZ8002, a "triparental" like conjugation has been developed in which the mutagenised cosmid strain, ET12567/pUZ8002,
20 ET12567/pUB307 and pre-germinated *S.coelicolor* spores are incubated together in a modified version of the conjugation procedure described by Kieser *et al.*, 2000. In this "triparental" procedure the cosmid is firstly mobilised by ET1256/pUB307 into strain ET1256/pUZ8002 and
25 then is subsequently mobilised by ET1256/pUZ8002 into *S.coelicolor*.

In order to allow in-frame deletions within a specific gene or site of interest, different knock-out cassettes
30 have been constructed. pIJ771 and pIJ772 as template plasmids leave behind an 84 bp "scar" sequence with stop codons in all six reading frames after FLP-mediated excision of the knock-out cassette (where the knock-out cassette has previously been integrated into the specific

gene or site of interest). In contrast, the pIJ773
"scar" (81 bp) has stop codons in all three forward but
in only one reverse frame. pIJ773 might be useful for
creating in-frame deletions in which the scar encodes a
5 new 27-residue internal peptide(s).

Both the inclusion of an *oriT* within the knock-out
cassette and the "triparental" like conjugation allow the
development of a high throughput knock-out system in
10 *S.coelicolor* and potentially in other actinomycetes as
well.

As will be evident from the foregoing summary and the
following more detailed description, the scar sequences
15 correspond to the "insertions" defined above, including a
single FRT site and the A and B or C primers.

The FRT sites are examples of "recombining sequences";
the homology extensions (provided by the PCR primers) are
20 examples of "targeting sequences"; the *oriT* site is an
example of an "origin sequence"; the A, B and C primer
sites are examples of "intervening sequences"; the
aac(3)IV gene is an example of a selectable marker
sequence; the *oriT* and *hyg* of the nucleic acid construct
25 which is carried in pIJ774 (and introduced into cosmid A
in step a in Fig. 7) are examples of a "second origin
sequence" and "second marker sequence", respectively; the
Supercos 1 sequences used to target *oriT* and *hyg* into
cosmid A in Fig. 7 are examples of "helper construct
30 targeting sequences".

Summary of Knock-In Experiments

The same technique may be used, not only to make knock-out mutants by inserting one or more stop codons into a target gene, but also to insert other nucleic acid sequences of interest into a target gene, e.g. to insert a control sequences (e.g. exogenous promoter) upstream of the coding sequence of the gene or to insert additional coding sequence at the 3' end of the normal coding sequence of the gene, to generate a gene encoding a fusion protein.

10

The following more detailed description makes reference to the accompanying drawings, in which:

Fig. 1A shows the construction of plasmid pIJ771, which contains a nucleic acid construct according to the first aspect of the invention (also referred to as a "knock-out cassette");

Fig. 1B shows the sequence of the knock-out cassette from pIJ771: the arrows indicate the primer sites A and B (which correspond to intervening sequences as defined previously), the bold arrows the FRT sites (FLP recognition targets, i.e. recombining sequences) and the broken arrows the primers Apraforw and oriTrev to generate the PCR product of pOJ436, used in the construction of plasmid pIJ771; the upper case underlined sequences (312 bp - 1097 bp and 1450 bp - 1558 bp) represent an apramycin resistance gene *aac(3)IV* and an *oriT*, respectively; the lower case sequences (78 bp - 303 bp and 1561 bp - 1612 bp) are homologous to sequences in the vector Supercos 1; the bold upper case sequences at 1106 bp - 1110 bp, at 1130 bp - 1136 bp and at 1154 bp - 1160 bp indicate the ribosomal binding site, the -10 promoter region and the -35 promoter region respectively;

Fig. 2 shows a scheme for the construction of plasmids pIJ772 and pIJ773, which also contain knock-out cassettes according to the first aspect of the invention: *aac(3)IV* = Apramycin resistance gene; FRT = FLP recognition targets; primers: Apraforw (1,5), oriTrev (2), FRTApra (3), FRToriT or oriT773 (4), AoriTrev or AoriT773 (6), AoriTforw (7), Aprarev (8), FRTforw (9), FRTTrev or FRT773 (10), left primer site (A) and right primer site (B); explanation appears in the text;

Fig. 3A shows the structure of pIJ772;

Fig. 3B shows the sequence of the knock-out cassette from pIJ772: the arrows indicate the primer sites A and B and the bold arrows the FRT sites; the underlined sequences (67 bp - 851 bp and 1204 bp - 1313 bp) represent the apramycin resistance gene *aac(3)IV* and the *oriT*, respectively; the bold sequences at 860 bp - 864 bp, at 884 bp - 890 bp and at 908 bp - 914 bp indicate the ribosomal binding site, the -10 promoter region and the -35 promoter region, respectively.

Fig. 4A shows the structure of plasmid pIJ773;

Fig. 4B shows the sequence of the knock-out cassette from pIJ773: the arrows indicate the primer site A and C and the bold arrows the FRT sites; the underlined sequences (67 bp - 851 bp and 1204 bp - 1313 bp) represent the apramycin resistance gene *aac(3)IV* and the *oriT*, respectively; the bold sequences at 860 bp - 864 bp, at 884 bp - 890 bp and at 908 bp - 914 bp indicate the ribosomal binding site, the -10 promoter region and the -35 promoter region, respectively.

Fig. 5 shows the structure of plasmid pIJ774, which carries a helper nucleic acid construct according to the invention. The hygromycin resistance gene *hyg* fused with an *oriT* was cloned into the *SmaI/BglIII* sites of the vector Supercos1 to generate pIJ774. After digestion of pIJ774 with *HpaI/NcoI*, a 3270 bp fragment can be isolated which contains the hygromycin resistance gene *hyg* and an *oriT*, flanked by 889 bp and 431 bp stretches of Supercos 1 DNA respectively, which allow integration of this fragment into a mutagenised cosmid by λ RED mediated recombination;

Fig. 6 shows the sequence remaining from each of the knock-out cassettes after FLP-mediated excision of the antibiotic resistance gene and *oriT*. In pIJ771 and pIJ772, this forms an 84 bp scar, with stop codons in both directions and in all three reading frames. In pIJ773, it forms an 81 bp scar, with stop codons in all three forward reading frames, but only in one reverse reading frame;

Fig. 7 shows a scheme for generating multiple knock-outs. A mutagenised cosmid such as SC1C3::*whiI771*, SC1C3::*whiI772* or SC1C3::*whiI773* (as described in the text; these cosmids contain a *S.coelicolor whiI* gene into which a nucleic acid construct of the invention has been introduced) is subjected to FLP mediated recombination to excise the apramycin *oriT* cassette leaving the predicted "scar" sequence at the site of the original PCR cassette insertion (shown at **A** as a single FLT site). To subsequently add a new resistance marker with an *oriT* to the mutagenised cosmid, pIJ774 was utilised (constructed as described previously and in the text). The

hygromycin-*oriT* cassette of pIJ774 is used to replace the kanamycin resistance gene of the mutagenised cosmid as follows. The 3270 bp *HpaI*/*NcoI* fragment of pIJ774 containing the hygromycin resistance gene *hyg* combined
5 with an *oriT* and flanked by 889 bp and 431 bp sequences corresponding to sequences in Supercos 1 is used to replace the kanamycin resistance gene in Supercos 1 by λ RED mediated recombination (a). In this way, a new *oriT* is introduced in the mutagenised cosmid (generating a
10 helper plasmid according to the invention) allowing conjugation with a *S.coelicolor* mutant (B) which already contains the apramycin resistance gene fused with an *oriT*, generated as described previously. Exconjugants can be selected (b) for single cross-over resulting in
15 hygromycin and apramycin resistant colonies and can be further screened (c) by replica-plating on DNA medium with and without apramycin and hygromycin to select for double cross-over, which leaves colonies lacking both resistance markers. This results in a *S.coelicolor* mutant
20 (C) which is marker free and therefore can be conjugated with another mutagenised cosmid (D) containing the apramycin resistance gene fused with an *oriT* to generate a double mutant. The procedure can be repeated following steps (a), (b) and (c);

25

Fig. 8A shows the structure of an alternative form of plasmid pIJ773 to that shown in Fig. 4, containing the ampicillin resistance gene *amp*, in place of the *bla* gene. As with the plasmid of Fig. 4, the plasmid contains the
30 apramycin resistance gene *aac(3)IV* (AC=X99313) and the *oriT* of plasmid RP4 (=RK2) (AC=L27758), flanked by FRT sites (FLP recognition targets, see Datsenko and Wanner, 2000). The disruption cassette was cloned into the *EcoRV* site of pBluescript SK (+) (rather than pBluescript KS

(+), as in Fig. 4), allowing its isolation as a 1398 bp *EcoRI/HindIII* fragment.

* indicates suitable restriction sites for verification
5 of mutagenised cosmid DNA by restriction analysis (for example: *SstI* generates a 751 bp internal fragment within the disruption cassette);

Fig. 8B shows the sequence of the plasmid pIJ773 shown in
10 Fig. 8A. Sequence annotations are provided in Annex 1;

Fig. 9 shows a scheme for the construction of disruption
cassettes containing alternative antibiotic resistance
markers, with *oriT* (pIJ778 and pIJ780) and without *oriT*
15 (pIJ779 and pIJ781). The genes for alternative antibiotic
resistance were amplified by PCR (description in the
text) and targeted against pIJ773 using λ red
recombination. To generate cassettes with an *oriT*, the
primer design was carried out in such a way to allow
20 replacing the apramycin resistance gene *aac(3)IV* on
pIJ773 by the spectinomycin / streptomycin resistance
gene *aadA* and the viomycin resistance gene *vph* to
generate pIJ779 and pIJ780 respectively. For the
generation of disruption cassettes without *oriT*, the
25 primers for amplification of alternative resistance genes
were designed to generate a PCR product which contains
the resistance gene with its own promoter. This PCR
product is flanked by 40 bp of *oriT*-sequence and 40 bp of
aac(3)IV - sequence to allow replacement of the apramycin
30 resistance gene and the *oriT* by λ red recombination;

Fig. 10A shows the structure of template plasmid pIJ778,
which contains the streptomycin and spectinomycin
resistance gene *aadA* (AC=M60473) and the *oriT* of plasmid

RP4 (=RK2) (AC=L27758), flanked by FRT sites. The disruption cassette was cloned into the *EcoRV* site of pBluescript SK (+) allowing its isolation as a 1438 bp *EcoRI/HindIII* fragment.

5

* indicates suitable restriction sites for verification of mutagenised cosmid DNA by restriction analysis (for example: *NaeI* generates a 782 bp internal fragment within the disruption cassette);

10

Fig. 10B shows the sequence of plasmid pIJ778. Sequence annotations are provided in Annex 2;

Fig. 11A shows the structure of template plasmid pIJ780 containing the viomycin resistance gene *vph* (AC=X99314) and the *oriT* of plasmid RP4 (=RK2) (AC=L27758), flanked by FRT sites (FLP recognition targets, see Datsenko and Wanner, 2000). The disruption cassette was cloned into the *EcoRV* site of pBluescript SK (+) allowing its isolation as a 1510 bp *EcoRI/HindIII* fragment.

20

* indicates suitable restriction sites for verification of mutagenised cosmid DNA by restriction analysis (for example: *HincII* generates a 1137 bp internal fragment within the disruption cassette);

25

Fig. 11B shows the sequence of plasmid pIJ780. Sequence annotations are provided in Annex 3; and

Fig. 12 shows a scheme for the insertion of an exogenous promoter, *tipAp*, into the upstream non-coding sequence of the *glgE* gene of *S. coelicolor*.

30

Bacteria

Escherichia coli strain BW25113 containing the recombination plasmid pKD46, *E.coli* BW25141 containing the template plasmid pKD4 and *E.coli* DH5 α containing the yeast Flp recombinase genes were acquired from the *E.coli* Genetic Stock Centre, Yale University, New Haven, USA (strain numbers CGSC7630, CGSC7632 and CGSC6729 respectively).

10

Plasmids

Recombination plasmids: pKD46, pKD20 and pIJ790

15 The recombination vector pKD46 described by Datsenko and Wanner, 2000 contains the λ RED *gam*, *bet* and *exo* DNA fragments under the control of the *araC*-*P_{araB}* promoter which allows the induction of the λ RED genes by 1 mM L-arabinose in SOB media (Hanahan, 1983).

20

The recombination plasmid pKD20 is very similar to pKD46, and may be used interchangeably with pKD46. pKD20 is, however, slightly preferred, as it is a more efficient recombination vector.

25

The λ Red recombination plasmid pKD20 (*E.coli* Genetic stock center CGSC Strain#: 7637) was modified by replacing the ampicillin resistance gene *bla* by the chloramphenicol resistance gene *cat* by PCR targeting, generating pIJ790, to permit selection for both the cosmid (ampicillin and kanamycin resistance) and the λ Red recombination plasmid pIJ790 (chloramphenicol resistance).

30

The plasmid pIJ790 was constructed as follows:

The chloramphenicol resistance gene *cat* was amplified
5 from plasmid pIJ666 (Kieser and Melton., 1988) using
primers *cat*forw: (5'-
ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTCGGCACGTAAGAGGTTTC
C-3') and *cat*rev: (5'-
TTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATTAAGGGCACCAATAACTG
10 C-3'). Underlined is the 3' end of the primer homologous
to the *cat* gene of pIJ666. The 5' end of the primers are
homologous to the first (*cat*forw) and the last (*cat*rev)
40 bp of the *bla* gene for ampicillin resistance of pKD20.
Amplification was performed in a 50 µl reaction using the
15 Expand High Fidelity PCR System (Roche), 100 pg template
DNA, 200 µM dNTPs, 50 pmol each primer, 5% DMSO and
following protocol: a denaturation step at 94°C for 2 min
followed by 25 cycles with denaturation at 94°C for 45s,
annealing at 55°C for 45 s and extension at 72°C for 1 min
20 30 s. A last elongation step was done at 72°C for 5 min.
The generated 866 bp fragment was purified using the PCR
purification kit (Qiagen) according to the manufacturer's
instructions.

25 *E. coli* BW25113 containing the recombination plasmid pKD20
was grown in 10 ml SOB (Hanahan, 1983) with carbenicillin
(100 µg/ml) and L-arabinose (final concentration of 10mM)
at 30°C to an OD₆₀₀ of ~ 0.6 and then made electro-
competent by washing twice with 10 ml ice-cold 10 %
30 glycerol and resuspended in 100 µl 10 % glycerol. Cells
were electroporated with ~ 100 ng of the purified PCR
product. Electroporation was carried out in 0.2 cm ice-
cold electroporation chambers using a BioRad GenePulser

II set to the following parameters: 200 Ω , 25 μ F and 2,5 kV. Shocked cells were added to 1 ml LB (Luria-Bertani medium; Sambrook et al., 1998), incubated 1 h at 30°C and then spread onto LB agar to select for chloramphenicol^R

- 5 (25 μ g/ml) o/n at 30°C. Chloramphenicol resistant colonies were checked by replica-plating for carbenicillin^S on LB agar plates containing carbenicillin (100 μ g/ml) for the successful replacement of the bla gene by the PCR product containing the cat gene.

10

Template plasmids

The template plasmids pIJ771, pIJ772 and pIJ773 were used to generate the *whiI*-specific knock-out cassette by PCR
15 (see following description). These plasmids are derived from pBluescript SK (+) (Stratagene, USA) and contain the apramycin resistance gene *aac(3)IV* and the origin of transfer *oriT*, flanked by FRT sites (FLP recognition targets) and primer sites.

20

Construction of these template plasmids was as follows:

pIJ771 (Fig. 1)

- 25 The apramycin resistance gene *aac(3)IV* and the origin of transfer *oriT* from RK2 were jointly amplified from pOJ436 (Bierman et al., 1992) using primers Apraforw (5'-TCATGAGCTCAGCCAATC-3') and oriTrev (5'-CGCCAGCCTCGCAGAGCAG-3') (oligo sequences are underlined
30 in Fig. 1B).

Amplification was performed in a 50 μ l reaction using the Expand High Fidelity PCR System (Roche, Germany), 100 pg template DNA, 200 μ M dNTPs, 50 pmol each primer, 5% DMSO

and following protocol: a denaturation step at 94°C for 2 min followed by 25 cycles with denaturation at 94°C for 45 s, annealing at 55°C for 45 s and extension at 72°C for 1 min 30 s. A last elongation step was done at 72°C for 5 min.

The generated 1246 bp fragment was ligated (using standard methods from Sambrook et al., 1998) to the 2227 bp *PvuII*/*NaeI* fragment of pKD4 that contains the FRT sites and *bla* gene, to generate pIJ771, and was subsequently verified by sequencing.

pIJ772 (Figs. 2 and 3)

Bracketed numbers refer to the primers shown in Fig. 2.

The apramycin resistance gene *aac(3)IV* and the origin of transfer *oriT* from RK2 were amplified from pOJ436 using primers Apraforw (1) and oriTrev (2) (see above for primer sequences) which generate a 1246 bp fragment. A second PCR was performed using the 1246 bp fragment as template and the primers FRTApra (3) (5'-CCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCATGAGCTCAGCCAATCG-3') and FRToriT (4) (5'-TATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTGCCAGCCTCGCAGAGCAG-3') to generate a 1325 bp PCR product flanked by partial FRT (FLP recognition target) sites.

Amplification was performed in a 50 µl reaction using the Expand High Fidelity PCR System (Roche, Germany), 100 pg template DNA, 200 µM dNTPs, 50 pmol each primer, 5% DMSO and following protocol: a denaturation step at 94°C for 2 min followed by 10 cycles with denaturation at 94°C for 45

s, annealing at 50°C for 45 s and extension at 72°C for 1 min 30 s followed by 15 cycles with denaturation at 94°C for 45 s, annealing at 55°C for 45 s and extension at 72°C for 1 min 30 s. A last elongation step was done at 72°C
5 for 5 min.

To construct the complete FRT sites, two PCR reactions were done under the same conditions as the FRTApra/FRToriT PCR detailed above to generate the left
10 and right FRT sites separately. The left FRT site was generated by PCR using primers AoriTforw (7) (5'-GTGTAGGCTGGAGCTGCTTCGAAGTTCCTATACTTTCTAGAG-3') and Aprarev (8) (5'-GTGCAATACGAATGGCGA-3') and the right FRT site using primers Apraforw (5) and AoriTrev (6) (5'-
15 CATATGAATATCCTCCTTAGTTCCTATTCCGAAGTTCCTATT-3').

A 851 bp fragment (containing the left FRT site) and a 1309 bp fragment (containing the right FRT site) were gel purified using a QIAquick Kit (Qiagen, Hilden, Germany)
20 according to manufacturers recommendations and digested with *Xho*I (the apramycin resistance gene *aac(3)IV* contains an internal *Xho*I site). Both digests were purified by phenol/chloroform extraction and ligated together using standard conditions (Sambrook et al.,
25 1998).

The final knock-out construct was amplified by PCR (for conditions see FRTApra/FRToriT PCR above) using the ligation mix as template (1µl of the 50µl total ligation
30 volume) with primers FRTforw (9) (5'-GTGTAGGCTGGAGCTGCTTC-3') and FRTrev (10) (5'-CATATGAATATCCTCCTTAGT-3'), generating a 1375 bp fragment which was further cloned into the *Eco*RV site of

pBluescript SK (+) (Stratagene, USA) to generate pIJ772 (Fig. 3).

pIJ773 (Figs. 2 and 4)

5

This plasmid was constructed in the same manner as pIJ772 (Fig. 2) using pOJ436 as template for the first PCR with primers Apraforw (1) and oriTrev (2). The generated 1246 bp fragment was then used as template for a second PCR
10 with primers FRTApra (3) and oriT773 (4) (5'-
AGTCCTATTCTCTAGAAAGTATAGGAAGTTGCCGCCAGCTCGCAGAGC
AG-3'; note that this differs from primer FRToriT used at 4 in the construction of pIJ772) using the same
conditions as stated previously for primers
15 FRTApra/FRToriT (3,4) in the construction of pIJ772.

The complete FRT sites were added by two PCRs separately using primers AoriTforw (7) with Aprarev (8) and Apraforw (5) with AoriT773 (6) (5'-
20 ATTCCGGGGATCCGTCGACCTGCAGTTCGAAGTTCCTATTCTCTAGAAAG-3';
note that this differs from primer Aprarev used at 6 in the construction of pIJ772), generating 851 bp and 1304 bp fragments respectively, which were gel purified (using a QIAquick purification kit (Qiagen, Hilden, Germany)
25 according to manufacturers recommendations), digested with XhoI, purified by phenol/chloroform extraction, and ligated together using standard conditions (Sambrook et al., 1998). The ligation mix was then used as template (1µl of the 50µl total ligation volume) to amplify a 1371
30 bp fragment with primers FRTforw (9) and FRT773 (10) (5'-
ATTCCGGGGATCCGTCGACC-3'; note that this differs from primer FRTrev used at 10 in the construction of pIJ772).
The same conditions were used as described previously for

primers FRTApra/FRToriT (3,4) in the construction of pIJ772.

The 1317 bp fragment was subsequently cloned into the
5 EcoRV site of pBluescript SK (+) (Stratagene, USA) to generate pIJ773 (Fig. 4).

Note that Fig. 2 shows primer B on the right hand side of the knock-out cassette. This is correct for pIJ772, but
10 the knock-out cassette of pIJ773 differs from that of pIJ772 at the right hand side, having instead primer C. See Fig. 4. Both knock-out cassettes have primer A on the left hand side, as shown in Fig. 2.

15 pIJ774 (Fig. 5)

This template helper plasmid was generated to help in creating multiple knock-outs. The 1843 bp *ClaI/BamHI* fragment of the Omega fragment (Blondelet-Rouault et al.,
20 1997) carrying the hygromycin resistance gene was subcloned into the *ClaI/BamHI* sites of the 2929 bp *ClaI/BamHI* restriction fragment of pBluescript SK (+), yielding pBSKhyg.

25 The *oriT* of RK2 was amplified from pOJ436 using primer *oriTrev* (5'-CGCCAGCCTCGCAGAGCAG-3') and *OriTB* (5'-GAGTGGATCCTGCCAGCCTCGCAGAGCAG-3') introducing a new *BamHI* site at one end of the generated PCR product.

30 Amplification was performed in a 50 µl reaction using the Expand High Fidelity PCR System (Roche, Germany), 100 pg template DNA, 200 µM dNTPs, 50 pmol each primer, 5% DMSO and following protocol: a denaturation step at 94°C for 2 min followed by 25 cycles with denaturation at 94°C for 45

s, annealing at 55°C for 45 s and extension at 72°C for 1 min 30 s. A last elongation step was done at 72°C for 5 min.

5 The resulting 120 bp fragment was digested with *Bam*HI and ligated (using standard conditions, Sambrook et al., 1998) to a 4749 bp *Bam*HI/*Hind*III fragment from *Bam*HI/*Hind*III digested pBSKhyg to generate a linear
10 fragment that is blunt at one end and possesses a *Hind*III site at the other end. This linear fragment was then rendered blunt at both ends by using the DNA polymerase activity of the Klenow enzyme and pBSKhyg-*ori*T was then subsequently obtained by ligation of the two ends of the 4864 bp blunt ended fragment. Following this a 1951 bp
15 fragment containing the hygromycin resistance gene and the *ori*T was removed from pBSKhyg-*ori*T by *Xho*I/*Bam*HI digestion and ligated (using standard conditions Sambrook et al., 1998) to the 6934 bp fragment from a *Sma*I/*Bgl*II digest of Supercos1 DNA (Stratagene, USA) to generate a
20 linear fragment that was then blunt ended with Klenow enzyme and religated to generate pIJ774.

Alternative form of pIJ773 (Fig. 8)

25 Fig. 8 shows an alternative form of plasmid pIJ773, containing the ampicillin resistance gene *amp*, in place of the *bla* gene.

This form of pIJ773 was used for the production (see
30 below) of plasmids pIJ778-781.

To allow multiple double or triple mutants without need for the FLP-mediated excision system, different knock-out cassettes in template plasmids for gene disruptions were

- constructed allowing selection in *E.coli* and *Streptomyces*. In addition to the disruption cassettes present in pIJ772 and pIJ773 (containing the apramycin resistance gene *aac(3)IV* with *oriT*), pIJ778 containing
- 5 the spectinomycin-streptomycin resistance gene *aadA* with *oriT* and pIJ780 containing the viomycin resistance gene *vph* with *oriT* were constructed using pIJ773 as target for PCR targeting.
- 10 After a single disruption with an *oriT*-containing cassette, further disruptions can be performed on the same cosmid with *oriT*-free cassettes (pIJ779 contains *aadA* without *oriT* and pIJ780 contains *vph* without *oriT*).
- 15 The PCR products to generate pIJ778-781 were constructed as follows (see Figs. 9-11):

- The spectinomycin-streptomycin resistance gene *aadA* was amplified from the omega fragment from pHP45Ω (Prentki et
- 20 al., 1991; Prentki and Krisch, 1984) using primers pIJ778F (5'-
 GAACTTCATGAGCTCAGCCAATCGACTGGCGAGCGGCATCTTATTTGCCGACTACCT
 TGG -3') and pIJ778R (5'-
 CTATGATCGACTGATGTCATCAGCGGTGGAGTGCAATGTCATGAGGGAAGCGGTGAT
 25 CGC -3') or pIJ778F and pIJ779R (5'-
 AAGTTCCCGCCAGCCTCGCAGAGCAGCATTCCCGTTGAGCGTACAGTCTATGCCTGC
 GGC -3'), respectively. Underlined is the 3' end of the primer homologous to the *aadA* gene of the omega fragment (primer pIJ779R includes the promoter sequence of *aadA*).
- 30 The 5' end of the primer pIJ778F is homologous to the last 40 bp of the *aac(3)IV* gene of pIJ773. The 5' ends of the primers pIJ778R and pIJ779R are homologous to the first 40 bp of the *aac(3)IV* gene or to the first 40 bp of the *oriT* of pIJ773 respectively.

The PCR products to generate pIJ780 and pIJ781 were constructed in the same manner as pIJ778 and pIJ779. The viomycin resistance gene *vph* was amplified from the omega fragment (Blondelet-Rouault *et al.*, 1997) using primer pIJ780F (5'-
5 GAACTTCATGAGCTCAGCCAATCGACTGGCGAGCGGCATCCTACCGGTAGCCGCTGA
GGC-3') and pIJ780R (5'-
CTATGATCGACTGATGTCATCAGCGGTGGAGTGCAATGTCATGAGAATCATTTGAGAC
10 GCA-3') or pIJ780F and pIJ781R (5'-
AAGTTCCCGCCAGCCTCGCAGAGCAGGAATCCCGTTGAGCGATGAAGGCAGGAACCC
AGT-3') respectively. Underlined is the 3' end of the primer homologous to the *vph* gene of the omega fragment (primer pIJ781R includes the promoter sequence of *vph*).
15 The 5' end of the primer pIJ780F is homologous to the last 40 bp of the *aac(3)IV* gene of pIJ773. The 5' ends of the primers pIJ780R and pIJ781R are homologous to the first 40 bp of the *aac(3)IV* gene or to the first 40 bp of the *oriT* of pIJ773 respectively.
20
Amplification was performed in a 50 µl reaction using the Expand High Fidelity PCR System (Roche), 100 pg template DNA, 200 µM dNTPs, 50 pmol each primer, 5% DMSO and the following protocol: a denaturation step at 94°C for 2 min
25 followed by 25 cycles with denaturation at 94°C for 45 s, annealing at 55°C for 45 s and extension at 72°C for 1 min 30 s. A last elongation step was done at 72°C for 5 min. The generated PCR products were purified using the PCR purification kit (Qiagen) according to the manufacturer's
30 instructions.

RESULTS

A preliminary gene disruption experiment was carried out using the *whiI* gene of *S.coelicolor*. This gene is known to be involved in sporulation and its deletion results in a white non-sporulating phenotype (Ainsa et al., 1999).

- 5 In the gene disruption experiments described here, deletion of the *whiI* gene resulted (as expected) in a visible phenotype that lacked sporulation and appeared as white colonies on SFM (Kieser et al., 2000) agar plates. In three knock-out experiments using template plasmids
- 10 pIJ771, pIJ772 and pIJ773 respectively, the gene disruption was done in the *S.coelicolor* cosmid SC1C3 (Redenbach et al., 1996) followed by conjugation of the mutagenised cosmid into *S.coelicolor* (strain M145).

15 **Experiment 1**

- This experiment was performed with the template plasmid pIJ771 which should, following double cross-over in *S.coelicolor*, result in an 84 bp scar sequence that
- 20 possesses stop codons in all 6 reading frames.

Gene disruption in the *S.coelicolor* cosmid SC1C3

- Using a 1860 bp *HindIII* fragment of pIJ771 containing the
- 25 *aac(3)IV* gene and the *oriT* flanked by FRT sites as template a *whiI*-specific knock-out PCR product was generated by using primers *whiIFRTforw* (5'-
CAGTGTGTCCGAATTGCTCGGATTGTTACTCACCAGATCGGTGTAGGCTGGAGCTGC
TTC-3') and *whiIFRTrev* (5'-
30 GAGCGGGGAGCGGAGATCGGGGAATCGGGGATCGGAGCGGCATATGAATATCCTCCT
TAGT -3') (underlined is the 3' end of the primer with
homology to the primer site of the knock-out cassette).

Amplification was performed in a 50 μ l reaction using the Expand High Fidelity PCR System (Roche, Germany), using 100 pg template DNA, 200 μ M dNTPs, 50 pmol each primer, 5% DMSO and following protocol: a denaturation step at 5 94°C for 2 min followed by 10 cycles with denaturation at 94°C for 45 s, annealing at 45°C for 45 s and extension at 72°C for 1 min 30 s followed by 15 cycles with denaturation at 94°C for 45 s, annealing at 50°C for 45 s and extension at 72°C for 1 min 30 s. A last elongation 10 step was done at 72°C for 5 min.

E.coli BW25113 containing the recombination plasmid pKD46 was grown in 10 ml SOB (Hanahan, 1983) with carbenicillin (100 μ g/ml) at 30°C to an OD₆₀₀ of $\approx$ 0.6 and then made 15 electro-competent by washing twice with 10 ml ice-cold 10 % glycerol and resuspending in 100 μ l 10 % glycerol. Cells were electroporated with $\approx$ 100 ng of cosmid SC1C3 DNA. Electroporation was carried out in 0.2 cm ice-cold electroporation chambers using a BioRad GenePulser II set 20 to the following parameters: 200 Ω , 25 μ F and 2.5 kV.

Shocked cells were added to 1 ml LB (Luria-Bertani medium; Sambrook *et al.*, 1998), incubated 1 h at 30°C and then spread onto LB agar to select for carbenicillin^R (100 25 μ g/ml), and kanamycin^R (50 μ g/ml).

Transformants containing the recombination plasmid pKD46 and the *S.coelicolor* cosmid SC1C3 were grown in 10 ml SOB cultures with carbenicillin (100 μ g/ml) kanamycin (50 30 μ g/ml) and L-arabinose (1mM final) in order to induce the expression of the λ RED genes, at 30°C to an OD₆₀₀ of $\approx$ 0.6 and then made electro-competent by concentrating 100-fold

and washing twice with 10 ml ice-cold 10 % glycerol and finally resuspending in 100 µl 10 % glycerol. Cells were electroporated with ≈ 100ng of purified (QIAquick Purification Kit, Qiagen, Hilden, Germany) knock-out cassette PCR product using the same parameters and
5 conditions as during electroporation (see above).

Shocked cells were added to 1 ml LB, incubated 1 h at 37°C and then spread onto LB agar to select for carbenicillin^R
10 (100 µg/ml) apramycin^R (50 µg/ml) and kanamycin^R (50 µg/ml).

Recombinant cosmid DNA (designated SC1C3::*whiI*771) from antibiotic resistant transformants was further
15 characterized by restriction enzyme digestion with *Xho*I/*Sac*I (the knock-out cassette contains one internal *Xho*I and two internal *Sac*I site(s)) and PCR using primer Aprarev (5'-GTGCAATACGAATGGCGA-3') with primer *whiI*2 (5'-CTATTTTCGGAGGGCTCTC-3') and primer oriTforw (5'-
20 GCCAGCCTCGCAGAGCAG-3') with primer *whiI*1 (5'-GAGAGATCGGGACAGGCC-3') to verify left and right novel joints.

The restriction enzyme digests and PCR products generated
25 fragments of the size expected (as determined by agarose gel electrophoresis) and as predicted. The PCR fragments generated were approximately 1487 bp and 715 bp from the Aprarev/*whiI*2 and oriTforw/*whiI*1 primer combinations respectively.

30

On average, ~ 1500 transformants were obtained by transforming the *E.coli* strain containing the recombination plasmid pKD46 and the cosmid SC1C3 with 100

ng of PCR product. Restriction digestion and PCR verification showed that less than 2% of all resulting transformants have the PCR knock-out cassette fragment integrated within the expected location in the cosmid SC1C3. Most of the transformants have the knock-out cassette integrated in the Supercos 1 sequence within the cosmid. The reason for this is the presence of a 226 bp sequence upstream of the *aac(3)IV* gene and a 52 bp sequence between *oriT* and the right FRT site within the knock-out cassette that show 100% homology to the vector Supercos 1 (Fig. 1B), which is the vector used for the construction of the ordered *S.coelicolor* cosmid library used in this experiment (Redenbach et al., 1996).

To avoid this problem of false positives (transformants which are apramycin^R, but show an integration of the *whiI*-specific knock-out cassette within the vector of the cosmid SC1C3), a new template plasmid pIJ772 was constructed as described above lacking the sequence homology to Supercos 1. See Experiment 2.

Conjugation of *S.coelicolor* M145 with mutagenised cosmid SC1C3::*whiI771*

In order to test the functionality of the *oriT* in SC1C3::*whiI771*, a conventional conjugation was carried out following the protocol described by Kieser et al., 2000 p249. The methylation-deficient *E.coli* host ET12567 (MacNeil et al., 1992) carrying the non-transmissible pUZ8002 (Kieser et al., 2000) was grown in 10 ml LB containing kanamycin (25 µg/ml) and chloramphenicol (25 µg/ml) to maintain selection for pUZ8002 and the *dam* mutation, respectively, and made electro-competent as described above. Electro-competent cells of

ET12567/pUZ8002 were transformed with ~ 100 ng of SC1C3::*whiI*771 DNA and selected for apramycin^R (50 µg/ml) on LB agar plates.

- 5 A single colony of ET12567/pUZ8002/SC1C3::*whiI*771 was inoculated into 10 ml LB containing kanamycin (25 µg/ml), chloramphenicol (25 µg/ml) and apramycin (50 µg/ml) and grown overnight at 37°C. The overnight culture was diluted 1:100 in fresh LB with antibiotic selection (see above)
- 10 and grown at 37°C to an OD₆₀₀ of ≈ 0.4 - 0.6. The cells were washed twice with an equal volume of LB to remove antibiotics and resuspended in 1 ml of LB.

- For each conjugation, approximately 10⁸ *Streptomyces*
- 15 spores were added to 500 µl 2 x YT broth, heat shocked at 50°C for 10 min, allowed to cool and then mixed with 500 µl ET12567/pUZ8002/SC1C3::*whiI*771 cells, spun briefly and most of the supernatant was poured off. The pellet was resuspended in the residual liquid and plated out on SFM
- 20 agar + 10 mM MgCl₂ (Kieser et al., 2000) and incubated at 30°C for 16-20 h.

- Plates were overlaid with 1 ml water containing 0.5 mg nalidixic acid (which kills *E.coli*, thereby selecting for
- 25 the *Streptomyces* cells) and 1 mg apramycin and incubated at 30°C for 3 - 6 days. Exconjugants were replica-plated on DNA medium (Kieser et al., 2000) containing kanamycin (50 µg/ml) and apramycin (50 µg/ml) and onto another
- plate containing kanamycin (50 µg/ml) only, to screen for
- 30 double cross-over exconjugants which are kanamycin^S and apramycin^R.

On average, 1000 - 1500 exconjugants were obtained by conjugating pregerminated *S.coelicolor* M145 spores with the *E.coli* strain ET12567/pUZ8002/SC1C3::*whiI*771. The rate of double cross-over was above 25 % as determined by the ratio of the number of colonies which were kanamycin^S and apramycin^R to the number which were kanamycin^R and apramycin^R.

Ten positive double cross-over mutants showing the expected white colony phenotype were verified by PCR using primers *whiI*1 and *whiI*2 following the protocol described by Kieser et al., 2000 p402-403. A fragment of about 2566 bp should be generated as predicted. A fragment of approximately 2566 bp (as determined by agarose gel electrophoresis) was observed in all ten cases.

Experiment 2

This experiment was performed with the template plasmid pIJ772 which has had the region of homology to SC1C3 in the PCR knock-out cassette removed and should, following double cross-over in *S. coelicolor*, result in an 84 bp scar sequence that possesses stop codons in all 6 reading frames.

Gene disruption in the *S.coelicolor* cosmid SC1C3.

Using an internal 1399 bp *EcoRI*/*HindIII* fragment of pIJ772 containing the apramycin resistance and the *oriT* flanked by FRT and primer sites, a *whiI*-specific knock-out cassette was generated by PCR and *E.coli* BW25113 containing the recombination plasmid pKD46 was made electro-competent and transformed with *S.coelicolor*

cosmid SC1C3 following the procedure described in Experiment 1.

E. coli BW25113 containing the recombination plasmid pKD46 and the cosmid SC1C3 was made chemical-competent with TSS (85% LB, 5% DMSO, 10% PEG, 50 mM MgCl₂: pH 6.5) following the protocol of Chung et al., 1989. Competent cells were transformed with ~ 100 ng unpurified knock out cassette PCR product, heat shocked at 42°C for 90 s and added to 1 ml LB, incubated 90 min at 37°C and then spread onto LB agar to select for carbenicillin^R (100 µg/ml), kanamycin^R (50 µg/ml) and apramycin^R (50 µg/ml).

On average, 70 antibiotic resistant transformants were obtained for each transformation of *E. coli* BW25113/pKD46/SC1C3 with ~ 100 ng knock out cassette PCR product.

Recombinant cosmid DNA (designated SC1C3::*whiI772*) for 48 of the resulting 70 transformants was further characterized by restriction enzyme digestion with *XhoI*/*SacI* and PCR using primer Aprarev with primer *whiI2* and primer oriTforw with primer *whiI2*. In recombinant cosmids from all 48 transformants the restriction enzyme digests and PCR products generated fragments of the size expected (as determined by agarose gel electrophoresis) and predicted. The PCR fragments generated were approximately 1242 bp and 600 bp for the Aprarev/*whiI2* and oriTforw/*whiI2* primer combinations respectively.

30

Conjugation of *S.coelicolor* M145 with mutagenised cosmid SC1C3::*whiI772*

To avoid the transformation step of ET12567/pUZ8002 with DNA of SC1C3::*whiI*772 a "triparental" like conjugation was carried out between an *E. coli* strain harbouring SC1C3::*whiI*772, ET12567/pUZ8002 and ET124567/pUB307
5 together with pre-germinated *S.coelicolor* spores following the conjugation procedure described by Kieser et al., 2000 and in Experiment 1.

120 exconjugants were obtained per conjugation in
10 comparison with 1000 - 1500 exconjugants obtained using a conventional conjugation method (including the additional step of transforming ET12567/pUZ8002 with SC1C3::*whiI*772) which represents a sufficient yield of exconjugants, considering a double-cross-over rate of at least 25%. Ten
15 positive double cross-over mutants showing the expected white colony phenotype were verified by PCR using primers *whiI*1 and *whiI*2 following the protocol described by Kieser et al., 2000. These PCR reactions generated a fragment of approximately 2267 bp (as determined by
20 agarose gel electrophoresis) in all ten cases as expected and predicted.

Experiment 3

25 This experiment was performed with the template plasmid pIJ773 which has had the region of homology to SC1C3 in the PCR knock out cassette removed and should, following double cross-over in *S. coelicolor*, result in an 81 bp scar sequence that possesses stop codons in all 3 forward
30 and one reverse reading frames.

Gene disruption in the *S.coelicolor* cosmid SC1C3

Using an internal 1396 bp *EcoRI/HindIII* fragment of pIJ773 containing the apramycin resistance and the *oriT* flanked by FRT and primer sites, a *whiI*-specific knock-out cassette was generated by PCR using primers

5 *whiIFRTforw* and *whiI773* (5'-
GAGCGGGGAGCGGAGATCGGGGAATCGGGGATCGGAGCGGTTCCGGGGATCCGTCGA
CC-3'; underlined is the 3' end of the primer that
possesses homology to the primer site of the knock-out
cassette pIJ773). Conditions of the PCR are described in
10 Experiment 1.

E. coli BW25113 containing the recombination plasmid pKD46 was made electro-competent and transformed with *S. coelicolor* cosmid SC1C3 following the procedure

15 described in Experiment 1.

E. coli BW25113 containing the recombination plasmid pKD46 and the cosmid SC1C3 were made chemical-competent and transformed with the *whiI* knock-out cassette PCR product

20 as described in Experiment 2.

Recombinant cosmid DNA (designated SC1C3::*whiI773*) from 48 of the resulting transformants was further characterized by restriction enzyme digestion with

25 *XhoI/SacI* and PCR using primer Aprarev with primer *whiI2* and primer *oriTforw* with primer *whiI2*. In all 48 selected transformants the restriction enzyme digests and PCR products generated fragments of the size expected (as determined by agarose gel electrophoresis) and predicted.

30 The PCR fragments generated were approximately 1242 bp and 586 bp for the Aprarev/*whiI2* and *oriTforw/whiI2* primer combinations respectively.

Conjugation of *S.coelicolor* M145 with mutagenised cosmid
SC1C3::*whiI773*

The conjugation procedure followed the description in
5 Experiment 2. 90 exconjugants were obtained with a double
cross-over rate of at least 25%. Ten positive double
cross-over mutants showing the expected phenotype were
verified by PCR using primers *whiI1* and *whiI2* following
the protocol described by Kieser et al., 2000. These PCR
10 generated a fragment of approximately 2253 bp fragment
(as determined by agarose gel electrophoresis) in all ten
cases as expected and predicted.

**Experiment 4: Elimination of the knock-out cassette by
15 FLP-mediated recombination**

BT340 [CGSC7629] is an ampicillin and chloramphenicol
resistant plasmid that shows temperature-sensitive
replication (only replicates at 30°C) and allows the
20 thermal induction of FLP synthesis at 42°C (Cherepanov and
Wackernagel, 1995).

E. coli strains harbouring SC1C3::*whiI771*, SC1C3::*whiI772*
or SC1C3::*whiI773* cosmids and resistant to all three
25 apramycin, kanamycin and carbenicillin antibiotics were
transformed with the BT340 plasmid, and apramycin^R and
chloramphenicol^R transformants were selected at 30°C.

For each recombinant cosmid strain (i.e. SC1C3::*whiI771*,
30 SC1C3::*whiI772* or SC1C3::*whiI773* derivatives), 50
colonies were streaked on a single LB agar plate
(masterplate) with no antibiotics and were grown non-
selectively at 43°C for 6 h to allow the expression and
synthesis of the FLP recombinase and the loss of the

plasmid BT340 simultaneously. The masterplate was then replica-plated on LB agar containing apramycin (50 µg/ml) to test for loss of the knock-out cassette.

- 5 After 6 h of non-selective growth at 43°C more than 50% of the colonies were apramycin^s indicating that recombinant cosmids had lost the apramycin resistance gene and the *oriT* simultaneously resulting in the predicted 84 bp or 81 bp "scar". Using primers *whiI1* and *whiI2*, the
- 10 remaining "scar" of five apramycin^s derivatives for each of SC1C3::*whiI771*, SC1C3::*whiI772* and SC1C3::*whiI773* was amplified by PCR and sequenced (Fig. 6).

- For each recombinant SC1C3::*whiI* cosmid the five
- 15 sequences obtained were all identical and confirmed that the FLP mediated excision of the knock out cassette had occurred as expected and resulted in the predicted scar sequence. After excision of the knock-out cassette from SC1C3::*whiI771* and SC1C3::*whiI772* a 84 bp "scar" sequence
- 20 was left in place of the disrupted gene(s) with stop codons in all six reading frames. In contrast, the remaining "scar" from SC1C3::*whiI773* (81 bp) has stop codons in all three forward but in only one reverse frame.

- 25
- It is believed that this procedure may also work in actinomycete cells, to remove the origin sequence and marker sequence following an integration step into the chromosomal gene, thereby to leave a single FRT site (and
- 30 optionally any intervening sequence, such as the primer sites A and B or C) as a scar sequence.

Experiment 5: Development to generate multiple knock-outs

The helper plasmid pIJ774 was constructed to allow multiple knock-outs in *S.coelicolor* by repeated use of the knock-out cassette containing apramycin resistance and *oriT*.

5

After FLP-mediated excision of the apramycin resistance and the *oriT* in the mutagenised cosmid (Fig. 7, at A), the 3270 bp *HpaI/NcoI* fragment of pIJ774 (hygromycin resistance with *oriT*), whose ends correspond to 889 bp and 431 bp sequences in Supercos 1, is introduced by λ RED mediated targeting following the same protocol described in Experiment 2 (Fig. 7, at A). (This procedure is comparable to the disruption of the cosmid gene in earlier experiments.) In this way, the Supercos 1 kanamycin resistance marker is replaced by the hygromycin resistance marker fused with an *oriT*.

10

15

The use of a different selectable marker in the helper plasmid allows confirmation that the new *oriT* sequence has been introduced into the helper plasmid.

20

After conjugation with a *S.coelicolor* mutant containing the apramycin resistance gene/*oriT* knock out cassette inserted into the chromosome (generated as described in Experiments 1, 2 and 3), exconjugants can be selected on hygromycin (50 μ g/ml; Fig. 7, at B) and screened by replica-plating on DNA-medium with and without apramycin (50 μ g/ml) and hygromycin (50 μ g/ml; Fig. 7, at C) for the loss of both resistance markers.

25

30

This *S.coelicolor* mutant (containing one "scar" within the first disrupted gene; Fig. 7, at C) can then be conjugated with another mutagenised cosmid containing an apramycin resistance cassette derived from pIJ771, pIJ772

or pIJ773 generating a double mutant (containing one "scar" within the first disrupted gene and the knock-out cassette at a second location; Fig. 7, at D).

- 5 This procedure can be repeated to generate multiple knock-outs in *S.coelicolor* without necessarily using more than two different resistance markers.

**Experiment 6: Use of recombination plasmid pIJ790 instead
10 of pKD46, along with template plasmid pIJ773 and optionally pIJ778-781**

E.coli BW25113 containing the recombination plasmid pIJ790 and the template plasmid pIJ773 was grown in 10 ml
15 SOB (Hanahan, 1983) with carbenicillin (100 µg/ml) and L-arabinose (final concentration of 10mM) at 30°C to an OD₆₀₀ of ~ 0.6 and then made electro-competent by washing twice with 10 ml ice-cold 10 % glycerol and resuspended in 100 µl 10 % glycerol. Cells were electroporated with ~ 100 ng
20 of the purified PCR product, respectively and electroporated in a 0.2 cm ice-cold electroporation chambers using a BioRad GenePulser II set to the following parameters: 200 Ω, 25 µF and 2,5 kV. Shocked cells were added to 1 ml LB (Luria-Bertani medium;
25 Sambrook et al., 1998), incubated 1 h at 37°C and then spread onto LB agar plates to select for carbenicillin^R (100 µg/ml), spectinomycin^R (50 µg/ml) and streptomycin^R (50 µg/ml) for transformants where the pIJ778 and pIJ779 PCR product was used and DNA agar plates to select for
30 carbenicillin^R (100 µg/ml) and viomycin^R (30 µg/ml) for transformants where the pIJ780 and pIJ781 PCR product was used.

The insert of each template plasmid was confirmed by sequencing with T3 and T7 primers in a standard sequencing reaction.

- 5 In more than 25 further knock-out experiments, pIJ790 was used successfully instead of pKD46 for gene disruption experiments.

Summary of Disruption Experiment Results

10

In the above experiments, and continuations thereof, the following gene disruptions have so far successfully been performed.

15

Cosmid SC1C3, gene *whiI*, template plasmid pIJ771, recombination plasmid pKD46, primers used:

*whiI*FRTforw

CAGTGTGTCCGAATTGCTCGGATTGTTACTCACCAGATCGGTGTAGGCTGGAGCTGCTTC

20

*whiI*FRTrev

GAGCGGGGAGCGGAGATCGGGGAATCGGGGATCGGAGCGGCATATGAATATCCTCCTTAGT

25

Cosmid SC1C3, gene *whiI*, template plasmid pIJ772, recombination plasmid pKD46, primers used:

*whiI*FRTforw

CAGTGTGTCCGAATTGCTCGGATTGTTACTCACCAGATCGGTGTAGGCTGGAGCTGCTTC

30

*whiI*FRTrev

GAGCGGGGAGCGGAGATCGGGGAATCGGGGATCGGAGCGGCATATGAATATCCTCCTTAGT

Cosmid SC1C3, gene *whiI*, template plasmid pIJ773, recombination plasmid pKD46, primers used:

whiIFRTforw

CAGTGTGTCCGAATTGCTCGGATTGTTACTCACCAGATCGGTGTAGGCTGGAGCTGC
TTC

whiI773

5 GAGCGGGGAGCGGAGATCGGGGAATCGGGGATCGGAGCGGTTCCGGGGATCCGTCGA
CC

Cosmid SC1C3, gene *whiI*, template plasmid pIJ778,
recombination plasmid pIJ790, primers used:

10 *whiIFRTforw*

CAGTGTGTCCGAATTGCTCGGATTGTTACTCACCAGATCGGTGTAGGCTGGAGCTGC
TTC

whiI773

GAGCGGGGAGCGGAGATCGGGGAATCGGGGATCGGAGCGGTTCCGGGGATCCGTCGA
15 CC

Cosmid SC1C3, gene *whiI*, template plasmid pIJ779,
recombination plasmid pIJ790, primers used:

whiIFRTforw

20 CAGTGTGTCCGAATTGCTCGGATTGTTACTCACCAGATCGGTGTAGGCTGGAGCTGC
TTC

whiI773

GAGCGGGGAGCGGAGATCGGGGAATCGGGGATCGGAGCGGTTCCGGGGATCCGTCGA
CC

25

Cosmid SCH24, gene 18, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

H24.18 forw

GGGCGCCCGTCTCCGCCAAGGCCACGTACCCTGTTTTGTTATTCCGGGGATCCGTCG

30 ACC

H24.18 rev

TCATGCGTTGTCCGGGTGCTGAGCCCACCATTCTTCAGTTGTAGGCTGGAGCTGCT
TC

Cosmid SC3C3, gene *gsp*, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

3C3.*gsp* forw

GGCGTGGGACAAAAGACGACAAGCATCGGAGAAACCACTAATTCCGGGGATCCGTCG

5 ACC

3C3.*gsp* rev

GTGGCGCTACGGGACGTCATCTGTCTGGGGTCGTCCTTCTTTGTAGGCTGGAGCTGCT
TC

10 Cosmid SC7G11, gene *wblE*, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

7G11.*wblE* forw

AAGCATCCCAGCATGTAATACCAAGGAGAGGTAGCAGCCATTCCGGGGATCCGTCGA
CC

15 7G11.*wblE* rev

ACGCGCCGCCAAGCTGAGGCTGTCAGTAGGGGTGGGTGTGTAGGCTGGAGCTGCTC

Cosmid SC3B6, gene 14, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

20 3B6.14 forw

GTCAAGCGGCCGGAACGAACAGACGAGGAAGACCACCCCATTCGGGGATCCGTCGA
CC

3B6.14 rev

ACGTCAGGGGGCGGTCGATCGTTCCTGCGGTGTTTCATGCTGTAGGCTGGAGCTGCTT

25 C

Cosmid SC1B2, gene 28c, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

1B2.28c forw

30 CCTAGCCCGACCGGTGCGGCAGCACATCGGAGACCCGCCATTCCGGGGATCCGTCGA
CC

1B2.28c rev

AGGTGGCACCGAAGGCGAGGACGGGGTGGGTTCATGCTGCTGTAGGCTGGAGCTGCTT

C

Cosmid SC5F8, gene 16c, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

5F8.16c forw

5 TGGACGAGGCGCGAACGATACCGATCGCAGGTGACTTTGATTCCGGGGATCCGTCGA
CC

5F8.16c rev

GTGCCCCGCGCCGGGCTGCCTCCGGCGTGGGCGGCGAGGGTGTAGGCTGGAGCTGCTT
C

10

Cosmid SC4B10, gene 07, template plasmid pIJ773;
recombination plasmid pIJ790, primers used:

4B10.07 forw

CCCCGAACCGGTATCCGTTTCTCATACAGGAGACAATTCATTCCGGGGATCCGTCGA

15 CC

4B10.07 rev

CCAGATCCCCTCGACCGGCCCCGACATCGACGGCCCAGGCTGTAGGCTGGAGCTGCTT
C

20 Cosmid SCP1/17, gene 95, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

SCP117.95 forw

CGGAGCTGTCTCCTGCACTGGACCTTCTGGGAGGCCCTGATTCCGGGGATCCGTCGA
CC

25 SCP117.95 rev

GGGGCGGGCCCGAACTTGCCCGCCGGTCCCGGACAGCGGTGTAGGCTGGAGCTGCTT
C

Cosmid SCP1/32, gene 115, template plasmid pIJ773,

30 recombination plasmid pIJ790, primers used:

SCP132.115 forw

CGAGTTGCACGCAATGCTCCCGGAGATCGCGTCAGTACTATTCCGGGGATCCGTCGA
CC

SCP132.115 rev

CTTCAACGGACCCCGCCGGTCACGTGCAGGGCCAGCGCTGTAGGCTGGAGCTGCTT
C

- 5 Cosmid SCH17, gene 13c, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

H17.13c forw

CCGTGGGGGAGCCTCGATTCTGGGAGAGGACGGCGCCGGTATTCCGGGGATCCGTCGA
CC

- 10 H17.13c rev

GCCGGGCGCCTGATCCGGGCCCCTGGTCCGGCCTCGGGGTGTAGGCTGGAGCTGCTT
C

- 15 Cosmid SC9B1, gene 20, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

9B1.20 forw

CGGCGGATGCGGTCTGAAGAGCCCTGGGTAGGGCCGGGCCATTCCGGGGATCCGTCGA
CC

9B1.20 rev

- 20 CGAGCCACGAAAGAGTGAGACTGAACGTCCGTCAGCGCGTGTAGGCTGGAGCTGCTT
C

- 25 Cosmid SC4G10, gene 13c, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

4G10.13c forw

CGCGCCGCAGAACGGAGGGTCAGGAGCCTCATGACCGTCATTCCGGGGATCCGTCGA
CC

4G10.13c rev

- 30 GTCTTCGGCATGGTGACGCGTCCTTTCCTCGGCCCACGCTGTAGGCTGGAGCTGCTT
C

- Cosmid SC4G10, gene glgPII, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

4G10.glgPII forw

ATCGTCCGTACCCCCGGACCGTCCGGAAGGTGACCGGCCATTCCGGGGATCCGTCGA
CC

4G10.glgPII rev

ACCGACTGCCCCGGACAGCCGGTCGGACACCTGGCAGGCTGTAGGCTGGAGCTGCTT

5 C

Cosmid SC7H1, gene 32, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

7H1.32 forw

10 TCTAGCCTCGGTTCGCAGACCCAGGAGGAACCCCATGACGATTCCGGGGATCCGTCGA
CC

7H1.32 rev

CCGGTCGTAGTAGCGCAGGGTGCGGATGCTCAGCCCGGTTGTAGGCTGGAGCTGCTT

C

15

Cosmid SCE39, gene 25, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

E39.25 forw

GACGGACTGACGCGAAGAGCCAGGAGAACAACGGTGGACATTCCGGGGATCCGTCGA

20 CC

E39.25 rev

GACCAGGCCGGTCGTCTCGTAGTGCCGGATGGTGCGCAGTGTAGGCTGGAGCTGCTT

C

25 Cosmid SC6A11, gene 13, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

6A11.13 forw

TTGACGTCCGCGTCAAGGCTTACGTTCGTCCCCAGTCGAATTCCGGGGATCCGTCGA

CC

30 6A11.13 rev

GCCCCGCCGCGCGGGCAGCAGACCCCGCGACTCGTAGTATGTAGGCTGGAGCTGCTT

C

Cosmid SC5F1, gene 39, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

5F1.39 forw

CGCTAGACCCTGACACTGACGTCAGAGGCAAGTATGTGTATTCCGGGGATCCGTCTGA

5 CC

5F1.39 rev

CTCGTAGTAGCGCACCGACCGGACGCTGACTCCGGCCCCGTGTAGGCTGGAGCTGCTT

C

10 Cosmid SC5H1, gene sigL, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

5H1.sigL forw

AGCCGTGACTTTCGTGGGAGAGGCAATGCAGACCGCCGTGATTCCGGGGATCCGTCT

ACC

15 5H1.sigL rev

CTCAGGCGCAGCCGAGTTCGCCCAGCATGCCCTCGCGCAGTGTAGGCTGGAGCTGCT

TC

Cosmid SC5F8, gene sigM, template plasmid pIJ773,

20 recombination plasmid pIJ790, primers used:

5F8.sigM forw

GCGGATCCGCATGCTCATAGAAACGCCCACCATCCGTCCCATTCCGGGGATCCGTCTG

ACC

5F8.sigM rev

25 CTCCCTGTCACCTGCTGTGTCCCCTGTGATCCTTGTGTCCTGTAGGCTGGAGCTGCT

TC

Cosmid SCF55, gene sigJ, template plasmid pIJ773,

recombination plasmid pIJ790, primers used:

30 F55.sigJ forw

CTTCGAGCAGCTCACCGTGCTGGAGGAGGGCACGCCCCGAGATTCCGGGGATCCGTCTG

ACC

F55.sigJ rev

ACCCGAGCCCCGCTCAGGTGGTGCTGAGCATGCCTTCCCGTGTAGGCTGGAGCTGCT
TC

- 5 Cosmid SCF55, gene dps, template plasmid pIJ773,
recombination plasmid pIJ790, primers used:

F55.dps forw

CGAAACGCATGAGGAGCTGAGGTCCGCCATGACGCACGACATTCCGGGGATCCGTCG
ACC

- 10 F55.dps rev

CGGCGCTCTCGAGATGCGCCCGCACGAACCACTGGAAGTGTGTAGGCTGGAGCTGCT
TC

Cosmid SC5C7, gene sigK, template plasmid pIJ773,

- 15 recombination plasmid pIJ790, primers used:

5C7.sigK forw

GCGACGCACCATGCCGATCCAGCCAGCGTGAAGCACCCGATTCCGGGGATCCGTCG
ACC

5C7.sigK rev

- 20 TTCACGCCTGCGGACCCGGATCCCGCAGGACCTGTTCCCGTGTAGGCTGGAGCTGCT
TC

Cosmid SCD13, gene sigN, template plasmid pIJ773,

recombination plasmid pIJ790, primers used:

- 25 D13.sigN forw

CATGTCCGCAGAACAGGGCAGCTCGAAGGTGCTCGCGCTCATTCCGGGGATCCGTCG
ACC

D13.sigN rev

- 30 CGGCGCGGGCTCAGTCGGAGATGAGACCCTCGCGCAGCTGTGTAGGCTGGAGCTGCT
TC

Cosmid SC7C7, gene dps, template plasmid pIJ773,

recombination plasmid pIJ790, primers used:

7C7.dps forw

ACCGACGGAGGCGTGGATGTACGTCGTGAAGAGCCCGTTGATTCCGGGGATCCGTCG
ACC

7C7.dps rev

CCGTCCGCGCCGAGACCATCCCGCGTCATCCGTTTTTCGGCTGTAGGCTGGAGCTGCT

5 TC

Cosmid SC9B1, gene 20, resulting in an N-terminal protein deletion, template plasmid pIJ773, recombination plasmid pIJ790, primers used:

10 9B1.20NF

AGCCCTGGGTAGGGCCGGGCCATGACGCAACAGCCCTTCATTCCGGGGATCCGTCGA
CC

9B1.20NR

GAGCAGTGCTCCACGTCCGCCGCGGAGGTGCCGGGCCCTGTAGGCTGGAGCTGCTT

15 C

Cosmid SC9B1, gene 20, resulting in a C-terminal protein deletion, template plasmid pIJ773, recombination plasmid pIJ790, primers used:

20 9B1.20CF

GTGCCCTTCCAGAAGGTCCGGCCCGTCCGTCATCCCCGACATTCCGGGGATCCGTCGA
CC

9B1.20CR

CAGTGAACGTCCGTCAGCGCGTCAGTGCGTCAGTGCGGGTGTAGGCTGGAGCTGCTT

25 C

All disruptions in the *S.coelicolor* cosmids have been successfully been proved by restriction- and PCR analysis. Disruptions in *S.coelicolor* have also been
30 verified by PCR and southern blot analysis.

The disruption of the gene number 20 on cosmid SC9B1 codes for a possible cyclase, which was predicted to be involved in the production of the volatile secondary metabolite geosmin (*trans*-1,10-dimethyl-*trans*-9-decalol)

in *S.coelicolor*. A common tool for compound identification is mass spectrometry (MS). After separation of the analytes by gas chromatography (GS) the mass spectrometer converts the sample into gaseous ions, normally by the exposure to a beam of electrons under vacuum. After separation and detection the resulting fragmentation pattern of different positive ions with different mass-to-charge ratios (m/z) enables the identification of compounds by comparison to libraries (Silverstein et al., 1981). Such an GS-MS was performed with the *S.coelicolor* SC9B1.20 mutant and confirmed the prediction of the involvement in geosmin production. The mutant in comparison to the wildtype did not produce any geosmin.

15

Experiment 7: Gene Disruption by Insertion ("Knock-In" Experiment)

The nucleic acid constructs ("disruption cassettes") of the invention also allow the insertion of new DNA sequences into any position within a target nucleic acid (e.g. *S.coelicolor* cosmid), with or without deleting original sequences during replacement with the disruption cassette ("knock-in"). As an example, shown in Fig. 12, the promotor of the gene *glgE* in cosmid 6A11 was replaced by the thiostreptone inducible promotor *tipAp* (Kataoka et al., 1996).

Shown in Fig. 12A, in a 3-step PCR, the inducible promotor was amplified together with the disruption cassette pIJ773 to allow targeting of upstream regions (shown as 1 and 2 in Fig. 12) by λ RED mediated recombination.

The first PCR was performed in a final volume of 50 μ l with 50 pmol primer P1 (5'-GCTCCGCGGAGGTGGCTGACGAGTGGTGCGTGGCGGGCATTTCGCTCCCTTCTCTGACG-3') and 50 pmol primer P2 (5'-

- 5 GGTCGACGGATCCCCGGAATCACCAATAAAAAACGCCCGG-3'), 200 μ M dNTPs, 5% DMSO, 1 x buffer, 2.5 U High Fidelity Taq polymerase (Roche) and 50 ng pIJ6021, which contains the inducible promotor *tipAp* (Takano et al., 1995). The conditions were: a denaturation step at 94°C for 2 min
10 followed by 10 cycles with denaturation at 94°C for 45 s, annealing at 50°C for 45 s and extension at 72°C for 30 s followed by another 10 cycles with denaturation at 94°C for 45 s, annealing at 55°C for 45 s and extension at 72°C for 30 s. A last elongation step was done at 72°C for 5 min.

15

The generated first PCR product (~ 250 bp) is flanked by two different extensions, 40 bp homologous to upstream regions of the *S.coelicolor* sequence of *glgE* (2) and 20 bp homologous to the priming sequence of the template
20 plasmid pIJ773.

- In a second PCR, pIJ773 and the first PCR product are added together to allow annealing and extension of the complete cassette, containing the promotor *tipAp* with the
25 apramycin-*oriT* cassette.

- The second PCR was performed in a final volume of 50 μ l with 50 ng pIJ773 template (1384 bp *EcoRI/HindIII* fragment), 1 μ l first PCR product (purified with Qiagen
30 PCR purification kit), 200 μ M dNTPs, 5% DMSO, 1 x buffer and 2.5 U High Fidelity Taq polymerase (Roche). The conditions were: a denaturation step at 94°C for 2 min

followed by 5 cycles with denaturation at 94°C for 15 s, annealing at 45°C for 30 s and extension at 72°C for 2 min. A last elongation step was done at 72°C for 5 min.

- 5 The complete "knock-in" cassette was generated in a third PCR with primer P1 and primer P3 (5'- GAACTGACGCGCACGGCGTGTTCGATATGGCCGGATCTTTCTGTAGGCTGGAGCTGCT YCG-3'), which contains a 40 nt extension homologous to chromosomal *S.coelicolor* DNA (1) in a final volume of 50
- 10 µl with 50 pmol primer P1 and 50 pmol primer P3, 200µM dNTPs, 5% DMSO, 1 x buffer, 2.5 U High Fidelity Taq polymerase (Roche), 10µl second PCR product (unpurified), 2µl first PCR product and 50 ng pIJ773 (1384 bp *EcoRI/HindIII* fragment). The conditions were: a
- 15 denaturation step at 94°C for 2 min followed by 10 cycles with denaturation at 94°C for 45 s, annealing at 50°C for 45 s and extension at 72°C for 1 min 30 s followed by another 10 cycles with denaturation at 94°C for 45 s, annealing at 55°C for 45 s and extension at 72°C for 1 min
- 20 30 s . A last elongation step was done at 72°C for 5 min.

- Shown in Fig. 12B, the third PCR product was purified (Qiagen PCR purification kit) and used for λ RED mediated PCR-targeting of the *S.coelicolor* cosmid 6A11. After
- 25 conjugal transfer of the mutagenised cosmid into *S.coelicolor* M145, the marked (apramycin^R) mutant was verified by PCR and Southern blot analysis.

- Shown in Fig. 12 C, to replace the marked insertion of
- 30 the promotor *tipAp* by the unmarked version, a FLP-mediated excision of the disruption cassette was performed as described in the section "FLP-mediated excision of the disruption cassette". The only divergence

between the wanted insertion of the promotor *tipAp* upstream of *glgE* and the resulting mutation after FLP-mediated excision of the disruption cassette is a FRT "scar" sequence 81 bp upstream of the promotor *tipAp*.

Annex 1: Sequence annotation for pIJ773 (Fig. 8B)

LOCUS pIJ773 4336 bp DNA CIRCULAR SYN 17-OCT-2001

DEFINITION Ligation of Apra-oriT disruption cassette into the EcoRV site of pBluescript SK(+)

ACCESSION pIJ773

REFERENCE 1 (bases 1 to 4336)

AUTHORS Self

JOURNAL Unpublished.

FEATURES	Location/Qualifiers
CDS	699..717 /marker="Primer site reverse" /product="TGT AGG CTG GAG CTG CTT C"
CDS	complement (718..751) /region="FRT" /product="Natural FRT site"
CDS	complement (764..1549) /gene="aac(3) IV" /product="aminoglycoside acetyltransferase inactivating apramycin"
CDS	1890..1999 /region="oriT" /product="origin of transfer from plasmid RP4 (also designated RP1/RK2) "
CDS	2008..2041 /region="FRT" /product="Natural FRT site"
CDS	complement (2050..2069) /marker="Primer site forward" /product="ATT CCG GGG ATC CGT CGA CC"
CDS	complement (3345..4205) /gene="amp" /product="b-lactamase"
BASE COUNT	1048 a 1121 c 1143 g 1024 t

Annex 2: Sequence annotation for pIJ778 (Fig. 10B)

LOCUS pIJ778 4376 bp DNA CIRCULAR SYN 17-OCT-2001

DEFINITION Ligation of Spec-Strep-oriT disruption cassette
into the EcoRV site of pBluescript SK(+)

ACCESSION pIJ778

REFERENCE 1 (bases 1 to 4376)

AUTHORS Self

JOURNAL Unpublished.

FEATURES Location/Qualifiers

CDS 699..717

/marker="Primer site reverse"

/product="TGT AGG CTG GAG CTG CTT C"

CDS complement (718..751)

/region="FRT"

/product="Natural FRT site"

CDS complement (798..1589)

/gene="aadA"

/product="spectinomycin-streptomycin

adenyltransferase gene from R100.1"

CDS 1930..2039

/region="oriT"

/product="origin of transfer from plasmid

RP4 (also designated RP1/RK2) "

CDS 2048..2081

/region="FRT"

/product="Natural FRT site"

CDS complement (2090..2109)

/marker="Primer site forward"

/product="ATT CCG GGG ATC CGT CGA CC"

CDS complement (3385..4245)

/gene="amp"

/product="b-lactamase"

BASE COUNT 1065 a 1120 c 1108 g 1083 t

Annex 3: Sequence annotation for pIJ780 (Fig. 11B)

LOCUS pIJ780 4448 bp DNA CIRCULAR SYN 17-OCT-2001
 DEFINITION Ligation of Vio-oriT disruption cassette into the
 EcoRV site of pBluescript SK(+)
 ACCESSION pIJ780
 REFERENCE 1 (bases 1 to 4448)
 AUTHORS Self
 JOURNAL Unpublished.

FEATURES Location/Qualifiers
 CDS 699..717
 /marker="Primer site reverse"
 /product="TGT AGG CTG GAG CTG CTT C"
 CDS complement (718..751)
 /region="FRT"
 /product="Natural FRT site"
 CDS complement (798..1661)
 /gene="vph"
 /product="viomycin phosphotransferase
 gene from *S.vinaceus*"
 CDS 2002..2111
 /region="oriT"
 /product="origin of transfer from plasmid
 RP4 (also designated RP1/RK2)"
 CDS 2120..2153
 /region="FRT"
 /product="Natural FRT site"
 CDS complement (2162..2181)
 /marker="Primer site forward"
 /product="ATT CCG GGG ATC CGT CGA CC"
 CDS complement (3457..4317)
 /gene="amp"
 /product="b-lactamase"

BASE COUNT 1002 a 1223 c 1211 g 1012 t

References :

- Ainsa, J.A., Parry, H.D. and Chater, K.F. (1999) *Mol. Microbiol.*, 34, 607-619.
- Bennett, P.M., Grinsted, J. and Richmond, M.H. (1977) *Mol.Gen.Genet.*, 154, 205-211.
- Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Rao, R.N. and Schoner, B.E. (1992) *Gene*, 116, 43-49.
- Blondelet-Rouault, M.H., Weiser, J., Lebrihi, A., Branny, P. and Pernodet, J.L. (1997) *Gene*, 190, 315-317.
- Cherepanov, P.P. and Wackernagel, W. (1995) *Gene*, 158, 9-14.
- Chung, C.T., Niemela, S.L. and Miller, R.H. et al., (1989) *Proc.Natl.Acad.Sci*, 86, 2172-2175.
- Datsenko, K.A. and Wanner, B.L. (2000) *Proc. Natl. Acad. Sci. USA* 97, 6640-6645.
- Flett, F., Mersinias, V. and Smith, C.P. (1997) *FEMS Microbiol.Lett.*, 155, 223-229.
- Hanahan, D. (1983) *J.Mol.Biol.*, 166, 557-580.
- Hoang, T.T., Karkhoff-Schweizer, R.R., Kutchma, A.J., Schweizer, H.P. (1998) *Gene* 212, 77-86.
- Kataoka M, Tatsuta T, Suzuki I, Seki T, and Yoshida T (1996) *J.Bacteriol.*, 178, 5540-5542
- Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F. and Hopwood, D.A. (2000) *Practical Streptomyces Genetics*; John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK.
- Kieser T and Melton RE (1988) *Gene*, 65, 83-91
- MacNeil, D.J., Gewain, K.M., Ruby, C.L., Dezeny, G., Gibbons, P.H. and MacNeil, T. (1992) *Gene*, 111, 61-68.
- Matsushima, P., Broughton, M.C., Turner, J.R. and Baltz, R.H. (1994) *Gene*, 146, 39-45.
- Pansegrau, W., Lanka, E., Barth, P.T., Figurski, D.H., Guiney, D.G., Haas, D., Helinski, D.R., Stanisich, V.A. and Thomas, C.M. (1994) *J.Mol.Biol.*, 239, 623-663.

- Prentki P, Binda A and Epstein A (1991) *Gene*, 103, 17-23
- Prentki P and Krisch HM (1984) *Gene*, 29, 303-313
- Redenbach, M., Kieser, H.M., Denapaite, D., Eichner, A., Cullum, J., Kinashi, H. and Hopwood, D.A. (1996) *Mol. Microbiol.*, 21, 77-96.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Silverstein RM, Bassler, GC and Morrill, TC (1981) *Spectrometric Identification of Organic Compounds*. New York, John Wiley & Sons, Inc.
- Takano E, White J, Thompson CJ and Bibb MJ (1995) *Gene*, 166, 133-137

Claims:

1. A method for generating a targeted nucleic acid disruption in an actinomycete, the method comprising:

- (1) providing a first nucleic acid construct comprising:
 - (a) a nucleic acid sequence (the "marker sequence") encoding a marker;
 - (b) a nucleic acid sequence (the "origin sequence") comprising an origin of transfer;
 - (c) a pair of nucleic acid sequences (the "recombining sequences") which are capable of undergoing recombination with each other, thereby to eliminate nucleic acid lying between them and leave only a single recombining sequence; and
 - (d) a pair of nucleic acid sequences (the "targeting sequences") capable of undergoing homologous recombination with a nucleic acid of interest of said actinomycete,

wherein parts (a) and (b) lie between the pair (c) of recombining sequences and parts (a), (b) and (c) lie between the pair of targeting sequences;

- (2) recombining the first nucleic acid construct with a first plasmid bearing at least part of the actinomycete target nucleic acid of interest to form a "recombined plasmid";

- (3) transferring the recombined plasmid by conjugation from a first bacterial cell into an actinomycete bacterial cell;

- (4) integrating nucleic acid of the first nucleic acid construct from the transferred plasmid into the target actinomycete nucleic acid by homologous recombination; and

- (5) eliminating the origin sequence and marker sequence from the integrated nucleic acid, to leave an insertion in the actinomycete nucleic acid, the insertion having a single

recombining sequence of the original pair of recombining sequences.

2. The method of claim 1, wherein the first nucleic acid construct further comprises:

(e) a pair of nucleic acid sequences (the "intervening sequences") which flank the recombining sequences and which are flanked by the targeting sequences.

3. The method of claim 1, wherein the first nucleic acid construct further comprises:

(e) an intervening sequence between the recombining sequence and the targeting sequence at one end only of the construct.

4. The method of any preceding claim, wherein the insertion comprises a stop codon in at least one reading frame in a forward or backward direction.

5. The method of claim 4, wherein the insertion comprises a stop codon in each of the three forward or in each of the three backward reading frames.

6. The method of claim 5, wherein the insertion comprises a stop codon in each of the six forward and backward reading frames.

7. The method of any preceding claim, wherein the nucleic acid sequence of interest is a gene.

8. The method of any preceding claim, wherein the targeting sequences are capable of undergoing homologous recombination with the upstream, non-coding region of a gene of interest or the coding region of a gene of interest.

9. The method of any preceding claim, wherein the marker is an antibiotic resistance marker.

10. The method of any preceding claim, wherein the recombining sequences are FRT (FLP recognition target) sequences, capable of being recombined by FLP recombinase activity *in trans*.

11. The method of claim 10, wherein each FRT sequence comprises a palindrome having the palindromic half site GAAGTTCCTAT.

12. The method of claim 11, wherein each FRT sequence comprises the double stranded core sequence:

5'-GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC-3'

3'-CTTCAAGGATATGAAAGATCTCTTATCCTTGAAG-5'

13. The method of any preceding claim, wherein the origin sequence is an *oriT* sequence which can be mobilised by the transfer plasmids pUZ8002 and pUB307.

14. The method of claim 13, wherein the *oriT* sequence has the nucleic acid sequence shown in Fig. 1B, residues 1450-1558.

15. The method of any preceding claim, wherein the nucleic acid construct is linear.

16. The method of any preceding claim, wherein step (1) comprises providing the nucleic acid construct by homology extension PCR from a plasmid containing parts (a), (b), (c) and preferably (e), the PCR primers (or their complementary strands) providing parts (d) and optionally (e).

17. The method of claim 16, wherein the generation of parts (d) and optionally (e) occurs in multiple PCR steps.
18. The method of any preceding claim, wherein step (2) comprises recombining the nucleic acid construct with nucleic acid from an actinomycete cosmid or BAC library.
19. The method of any preceding claim, wherein the first bacterial cell is methylation deficient.
20. The method of any preceding claim, wherein the first bacterial cell is an *E.coli* cell.
21. The method of any preceding claim, wherein step (2) occurs in a bacterial cell having λ RED (*gam*, *bet*, *exo*) function.
22. The method of claim 21, wherein the bacterial cell is an *E. coli* cell.
23. The method of any preceding claim, wherein step (2) also occurs in the first bacterial cell.
24. The method of any preceding claim, wherein step (3) includes the provision *in trans* of transfer function.
25. The method of claim 24, wherein the recombined plasmid is transformed into a donor strain carrying a non-transmissible transfer plasmid, followed by incubation with the actinomycete cell.
26. The method of claim 24, wherein the first bacterial cell, a donor strain carrying a non-transmissible transfer plasmid,

a donor strain carrying a self-transmissible transfer plasmid and the actinomycete cell are all incubated together.

27. The method of claim 24 or claim 25, wherein the donor strain carrying a non-transmissible transfer plasmid is ET12567/pUZ8002.

28. The method of claim 26 or claim 27 wherein the donor strain carrying a self-transmissible transfer plasmid is ET12567/pUB307.

29. The method of any one of claims 25 to 28, wherein the actinomycete cell is a pre-germinated actinomycete spore.

30. The method of any preceding claim, further comprising the step of confirming the elimination of the marker sequence and origin sequence by replica plating of exconjugants, elimination being reflected by lack of marker function.

31. The method of any preceding claim wherein step (5) follows steps (3) and (4).

32. The method of claim 31, wherein step (5) occurs by FLP recombinase-mediated recombination.

33. The method of claim 32, wherein FLP recombinase activity is provided *in trans* using a plasmid that shows inducible replication and induction of FLP synthesis under conditions in which replication will not occur, thereby allowing simultaneous induction of FLP synthesis and loss of the plasmid.

34. The method of claim 33, wherein the plasmid is temperature-inducible.

35. The method of claim 31, wherein step (5) comprises homologous recombination between the actinomycete nucleic acid of interest into which the nucleic acid construct has been integrated and a helper plasmid, which helper plasmid contains:

- a helper nucleic acid construct that:
 - comprises targeting sequences capable of undergoing homologous recombination with the same actinomycete nucleic acid of interest as the nucleic acid construct,
 - comprises a single recombining sequence between the targeting sequences, and
 - lacks an origin sequence between the targeting sequences, and
- a second origin sequence outside the nucleic acid construct.

36. The method of claim 35, wherein the helper nucleic acid construct further comprises an intervening sequence or a pair of intervening sequences.

37. The method of claim 35 or claim 36, wherein the helper nucleic acid construct lacks a marker sequence between the targeting sequences.

38. The method of any one of claims 35 to 37, wherein the helper plasmid further comprises a second marker sequence.

39. The method of any one of claims 35 to 38, wherein the helper plasmid is constructed by a method comprising steps (1), (2) and (5) as defined in claim 1.

40. The method of claim 39, wherein the construction method further comprises the subsequent step of introducing the

second origin sequence and second marker sequence by homologous recombination between the helper plasmid and a helper nucleic acid construct, the helper nucleic acid construct comprising the second origin sequence and second marker sequence in between two sequences (the "helper construct targeting sequences") which share homology with sequences of the helper plasmid.

41. The method claim 39 or 40, wherein the recombination step (2) as defined in claim 1 and the recombination step as defined in claim 39 involve recombination with identical plasmids.

42. The method of any preceding claim wherein the first plasmid and/or the helper plasmid are circular.

43. The method of any preceding claim which is repeated to generate a disruption in another target nucleic acid sequence of the actinomycete.

44. The method of any one of claims 1 to 42, further comprising a step between steps (2) and (3) of generating an additional disruption, the step comprising recombining an additional disruption construct with the recombined plasmid, wherein

the additional disruption construct comprises the same component parts as the first nucleic acid construct, with the exception that it lacks an origin sequence;

the marker sequence of the additional disruption construct is different from that of the first nucleic acid construct; and

the targeting sequences of the additional disruption construct are different from those of the first nucleic acid construct.

45. The method of claim 44, wherein the recombining sequences of the additional disruption construct are the same as those of the first nucleic acid construct.

46. A method for producing a library of actinomycete host cells having respective different gene disruptions, the method comprising:

repeating the method of any preceding claim to generate different gene disruptions in different respective host cells.

47. A library of actinomycete host cells having different respective gene disruptions as produced or as producible according to the method of claim 46.

48. The library of claim 47, wherein different host cells respectively have insertions in different chromosomal genes.

49. The library of claim 48, wherein the insertions each comprise a single FRT site, and optionally an intervening sequence or pair of intervening sequences.

50. A method comprising screening the library of any one of claims 47 to 49 for phenotypic characteristics or changes resulting from the gene disruptions.

51. The method of claim 50, wherein the screening is followed by the step of identifying the disrupted gene.

52. The method of any preceding claim, wherein the actinomycete cell is a cell of the genus *Streptomyces*.

53. The method of claim 52, wherein the species is *S.coelicolor* or *S.ambofaciens*.

54. A first nucleic acid construct as defined in any one of claims 1 to 15, with the optional exception of lacking targeting sequences.

55. A helper nucleic acid construct as defined in any one of claims 35 to 37.

56. An additional disruption construct as defined in claim 44 or claim 45.

57. A helper plasmid as defined in any one of claims 35 to 38.

58. A vector comprising the nucleic acid construct of any one of claims 54 to 56, flanked by restriction sites for excising the nucleic acid construct from the remainder of the vector.

59. A cell comprising a vector or plasmid as defined in claim 57 or claim 58.

60. A kit comprising a first nucleic acid construct as defined in claim 54; and a helper nucleic acid construct as defined in claim 55 or an additional disruption construct as defined in claim 56.

1/24

Fig. 1A

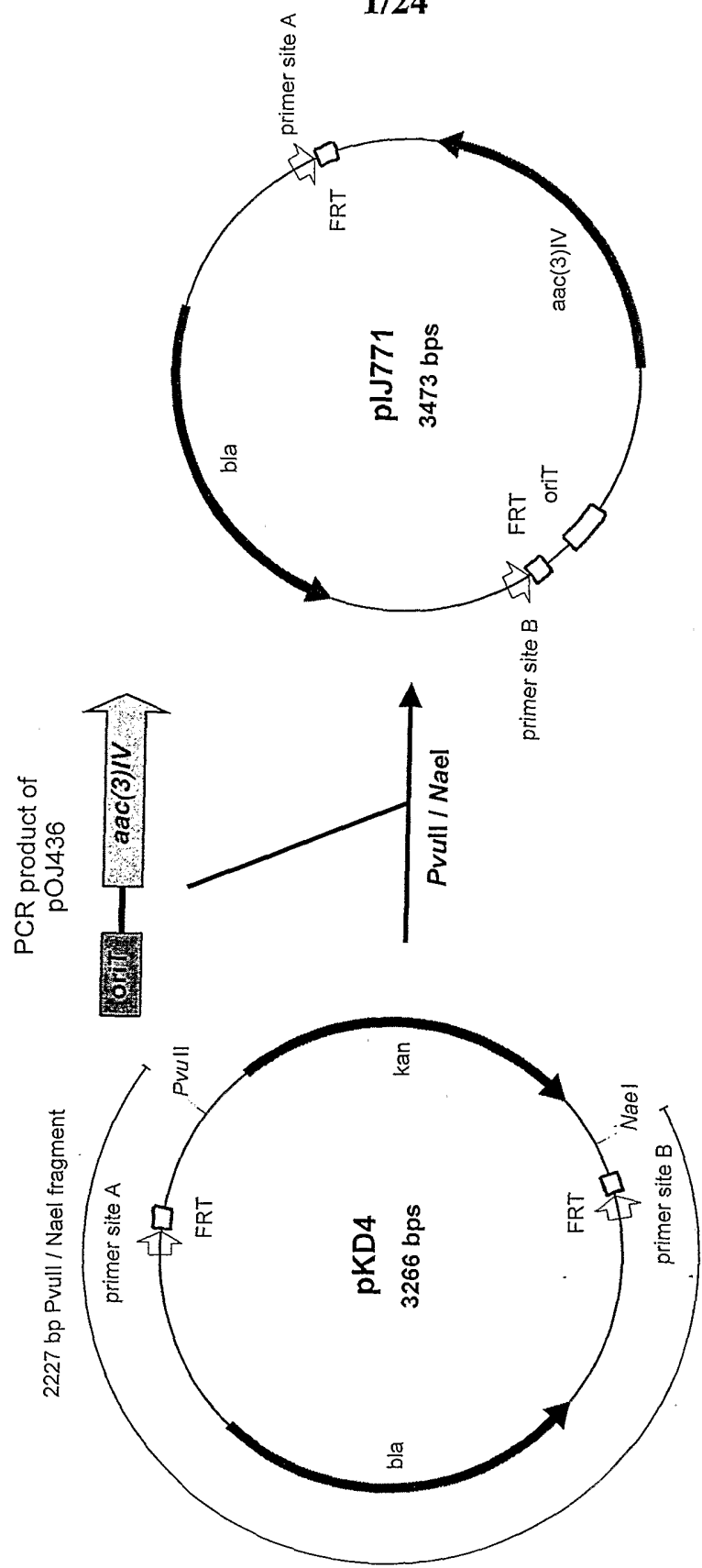
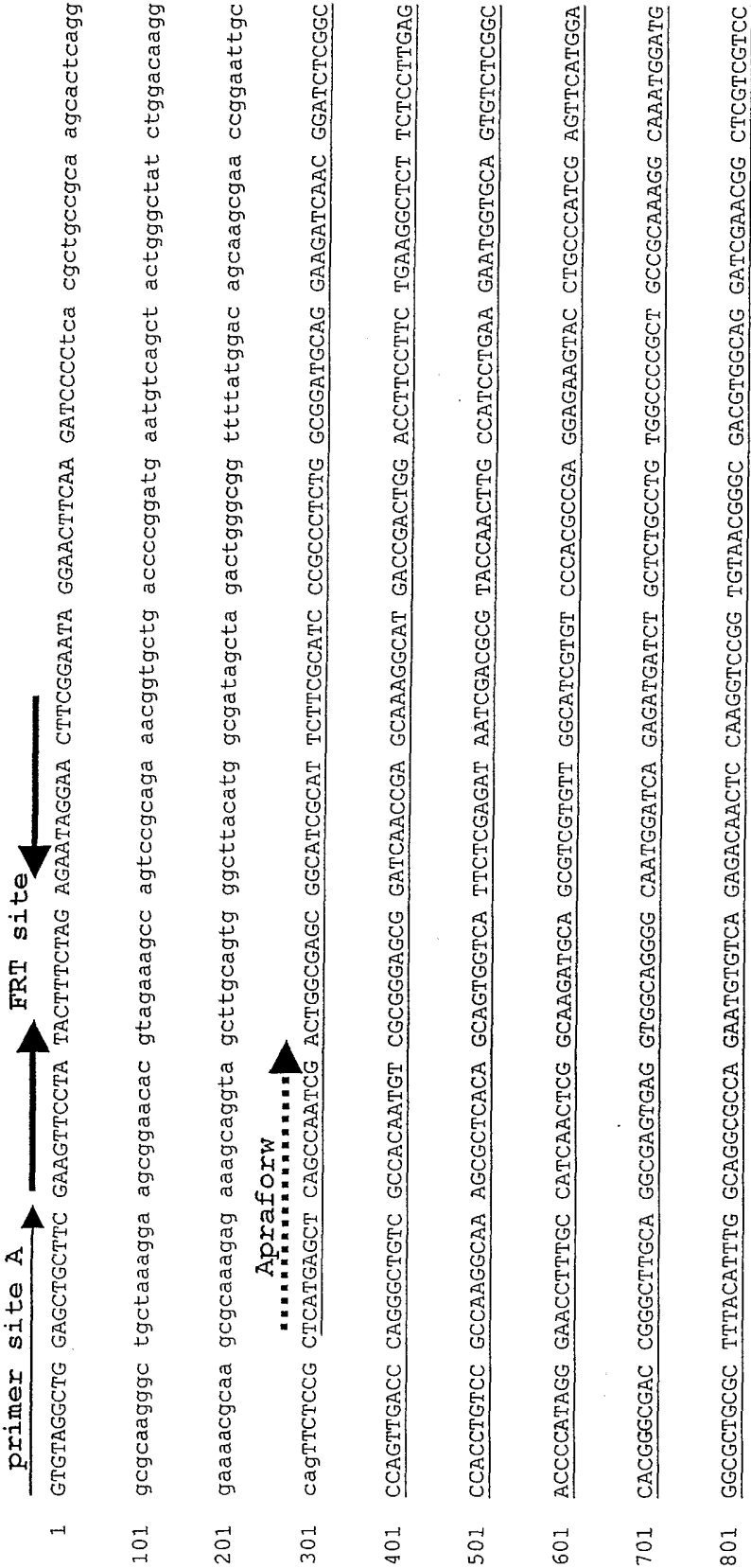


Fig. 1B



3/24

Fig. 1B (cont.)

```

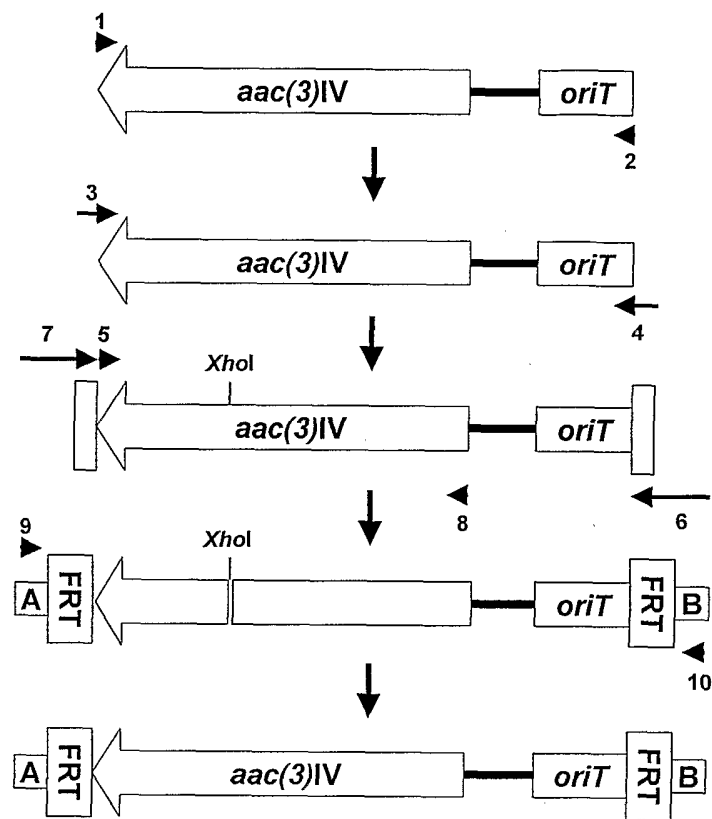
901  AGACCTGACC ACGAGGGGCAT GACGAGCGTC COTCCCGGAC CCAGCGCAGC ACGCAGGGCC TCGATCAGTC CAAGTGGCCC ATCTTCGAGG GCGCGGACGC
1001  TACGGAAGGA GCTGTGGACC AGCAGCACAC CGCCGGGGGT AACCCCAAGG TTGAGAAGCT GACCGATGAG CTCGGGCTTT CGCCATTGCT ATTGCACGAC
1101  ATTGCACTCC ACCGCTGATG ACATCAGTGG ATCATAGCAC GATCAACGGC ACTGTTGCAA ATAGTCGGTG GTGATAAACT TATCATCCCC TTTTGCTGAT
1201  GGAGCTGCAC ATGAACCCAT TCAAAGGCCG GCATTTTCAG CGTGACATCA TTCTGTGGC CGTACGCTGG TACTGCAAAT ACGGCATCAG TTACCGTGAG
1301  CTGCAGGTGG ACGGATCTTT TCCGCTGCAT AACCCCTGCTT CGGGGTCATT ATAGCGATTT TTTCGGTATA TCCATCCTTT TTTCGACGAT ATACAGGATT
1401  TTGCCAAAGG GTTCGTGTAG ACTTTCCTTG GTGTATCCAA CGGCGTCAGC CGGCAGGAT AGGTGAAGTA GGCCACCCG CGAGCGGGTG TTCCTTCTTC
1501  ACTGTCCTT ATTCCGACCT GCGGTGCTC AACGGGAATC CTGCTCTGCG AGGCTGGCGG ctggatgac ctccagcgc gggatctcat gctggagttc
1601  ttcgcccacc ccAGCTTCAA AAGCGCTCTG AAGTTCCTAT ACTTCTCTAGA GAATAGGAAC TTCGGAATAG GAACTAAGGA GGATATTCAT ATG

```

FRT site primer site B

Fig. 2

4/24



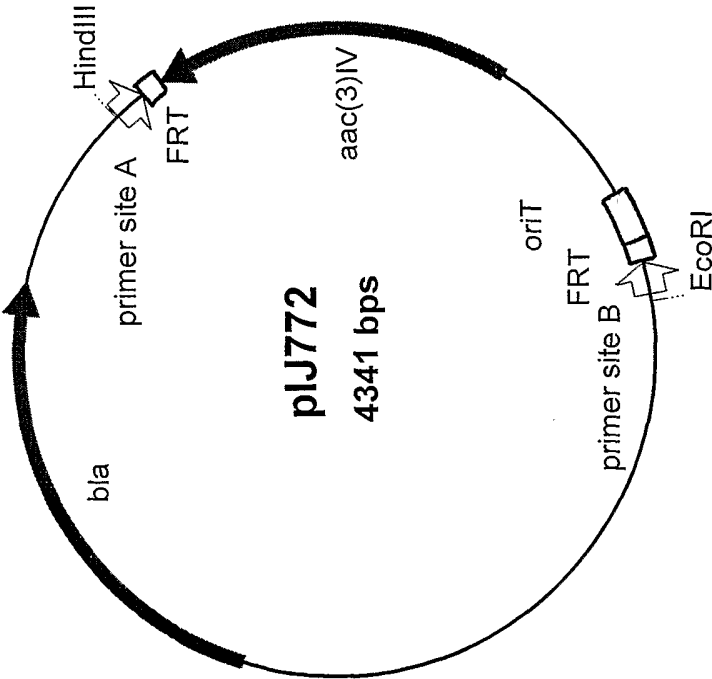
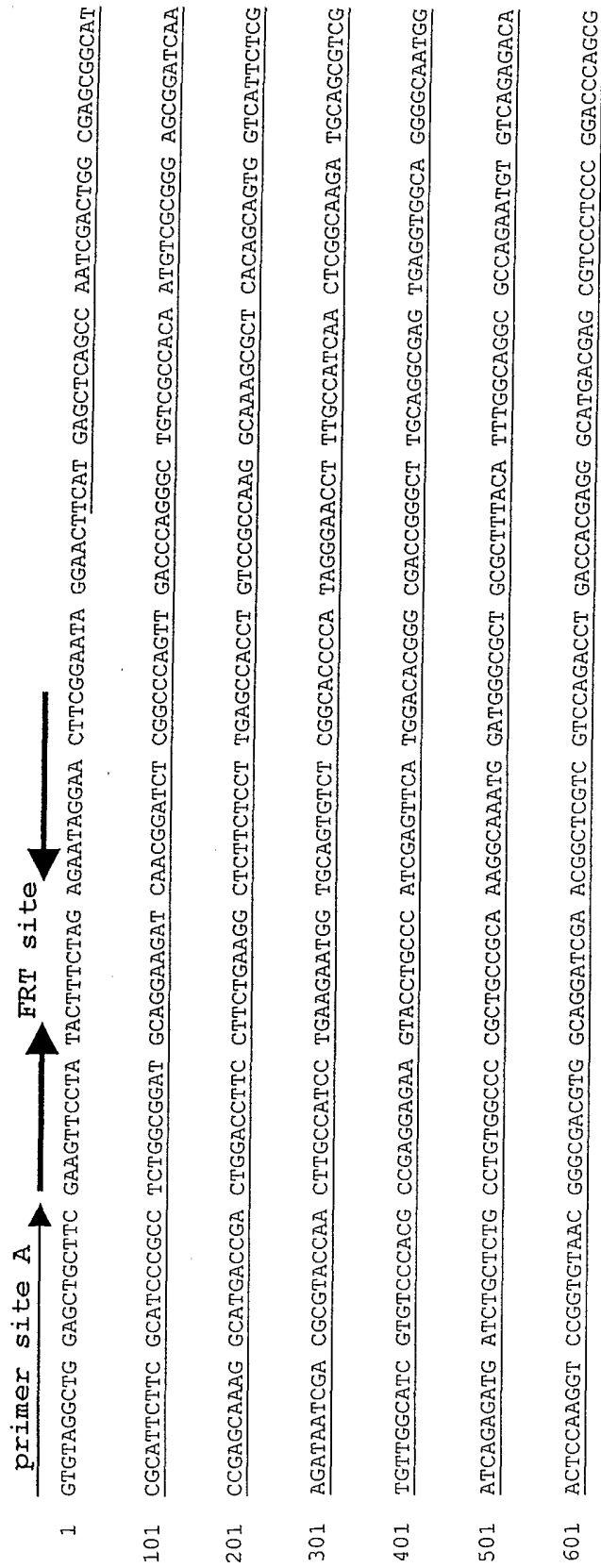


Fig. 3A

6/24

Fig. 3B



7/24

Fig. 3B (cont.)

701 CAGCACGCAG GGCCTCGATC AGTCCAAGTG GCCCATCTTC GAGGGGCCGG ACGCTACGGA AGGAGCTGTG GACCAGCAGC ACACGCCCGG GGGTAACCCC

801 AAGGTTGAGA AGCTGACCGA TGAGCTCGGC TTTTCGCCAT TCGTATTGCA CGACATTGCA CTCCACCGCT GATGACATCA GTCGATCATA GCACGATCAA

901 CGGCACCTGTT GCAAAATAGTC GGTGGTGATA AACTATATCAT CCCCTTTTGC TGATGGAGCT GCACATGAAC CCATTCAAAG GCCGGCATT TCCAGCGTGAC

1001 ATCATTCTGT GGGCCGTAAG CTGGTACTGC AAATACGGCA TCAGTTACCG TGAGCTGCAG GTCGACGGAT CTTTTCGGCT GCATAACCCT GCTTCGGGGT

1101 CATTATACCG ATTTTTCGG TATATCCATC CTTTTTCGCA CGATATACAG GATTTTGCCA AAGGGTTGCT GTAGACTTTC CTTGGTGTAT CCAACGGCGT

1201 CAGCCGGGCA GGATAGGTGA AGTAGGCCCA CCCGGGAGCG GGTGTTCCCTT CTCACCTGTC CCTTATTCGC ACCTGGCGGT GCTCAACGGG AATCCTGCTC

1301 TCGAGGCTG GCGGAAGTTC CTATACTTTC TAGAGAATAG GAACTTCGGA ATAGGAAC TAAGAGGATAT TCATATG

→ FRT site

← primer site B

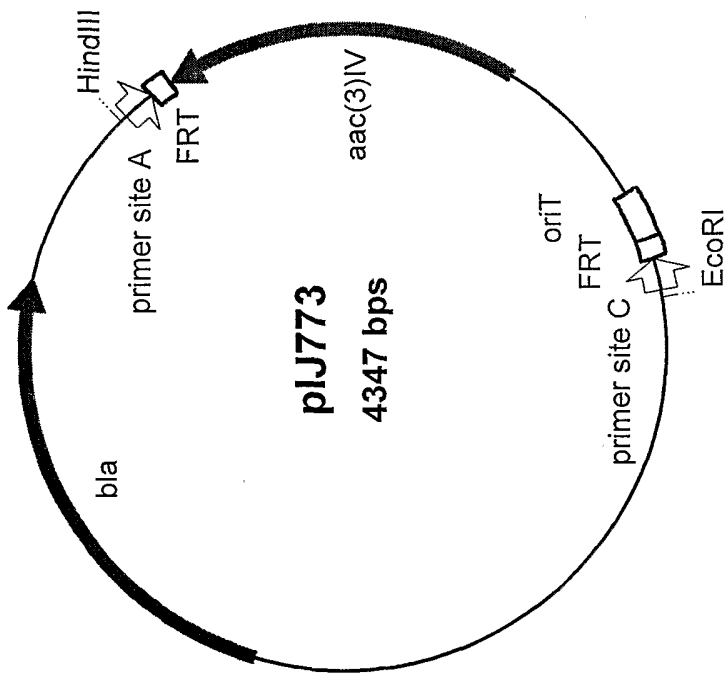
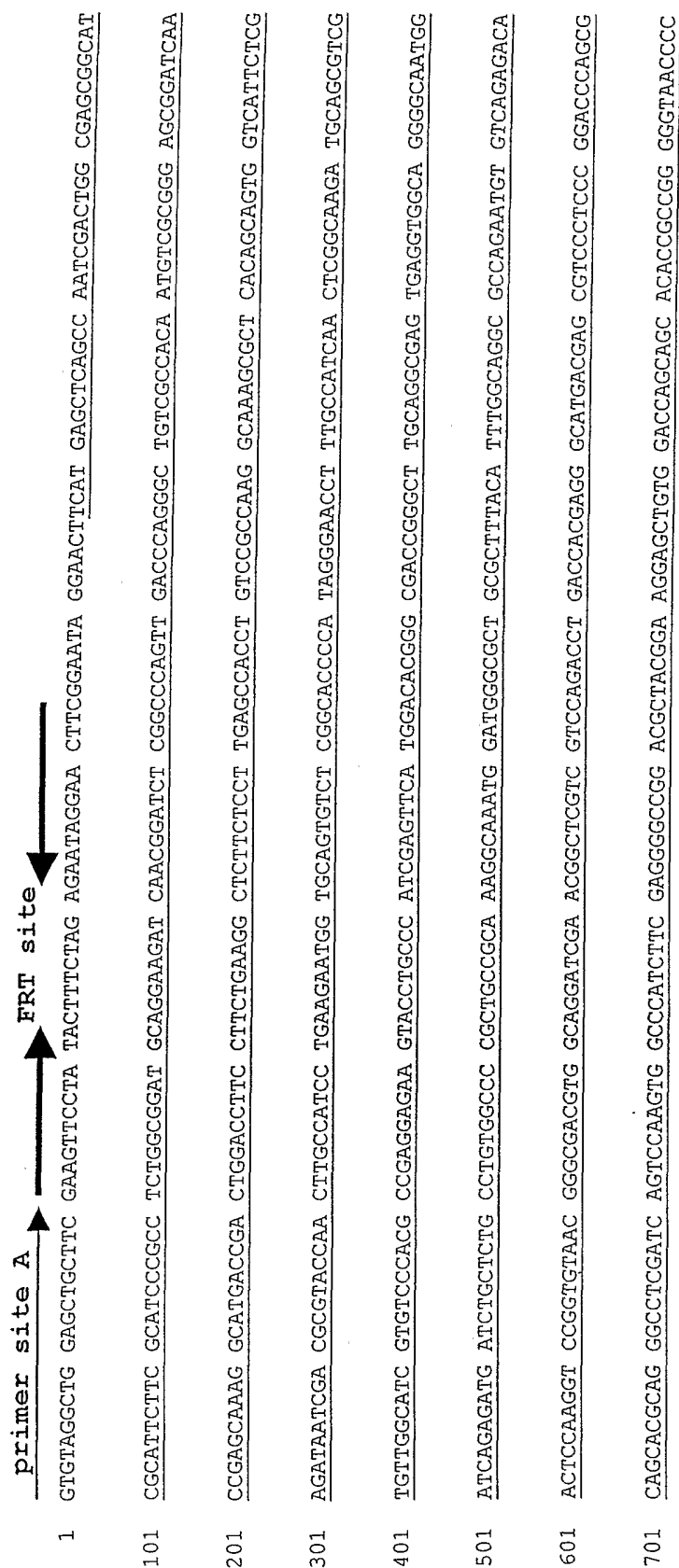


Fig. 4A

9/24

Fig. 48



10/24

Fig. 4B (cont.)

801 AAGTTGACA AGCTGACCGA TGAGCTCGGC TTTTCGCCAT TCGTATTGCA CGACATTGCA CTCCACCGCT GATGACATCA GTCGATCATA GCACGATCAA

901 CGGCACCTGTT GCAAATAGTC GGTGGTGATA AACTTATCAT CCCCTTTTGC TGATGGAGCT GCACATGAAC CCATTCAAAG GCCGGCAITTT TCAGCGTGAC

001 ATCATTCTGT GGGCCGTAGG CTGGTACTGC AAATACGGCA TCAGTTACCG TGAGCTGCAG GTCGACGGAT CTTTTCCGCT GCATAACCCCT GCTTCGGGGT

101 CATTATACCG ATTTTTCGG TATATCCATC CTTTTTCGCA CGATATACAG GATTTTGCCA AAGGTTTCGT GTAGACTTTC CTTGGTGTAT CCAACGGCGT

201 CAGCCGGGCA GGATAGGTGA AGTAGGCCA CCCGCCAGCG GGTGTTCCCTT CTTCACTGTC CCTTATTCGC ACCTGGCGGT GCTCAACGGG AATCCTGCTC

301 TGCGAGGCTG GCGGGAACTT CGAAGTTCCT ATACTTTCTA GAGATAGGA ACTTCGAACT GCAGGTCGAC GGATCCCCGG AAT

FRT site primer site C

11/24

Fig. 5

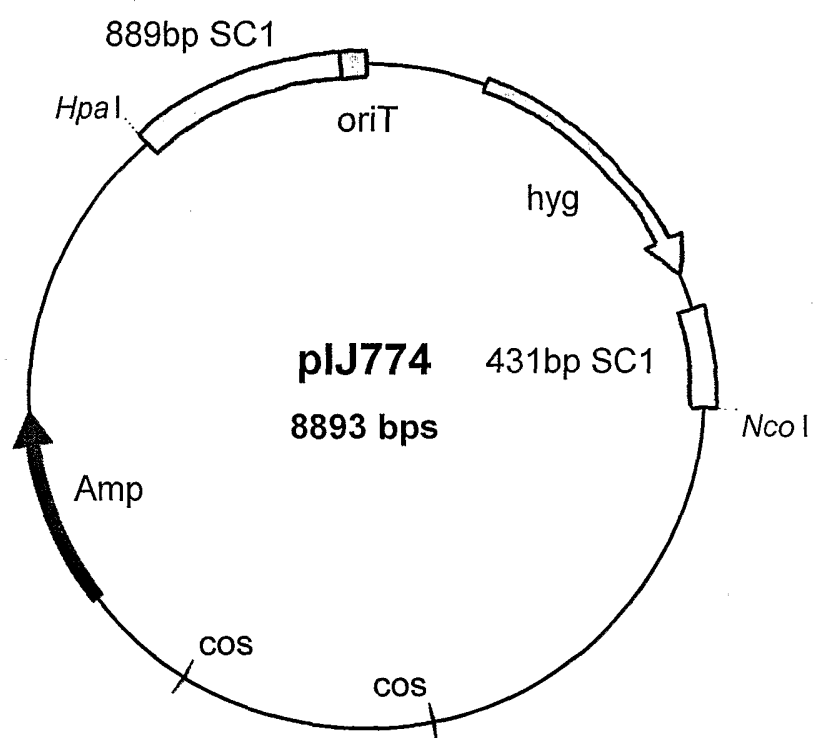


Fig. 6

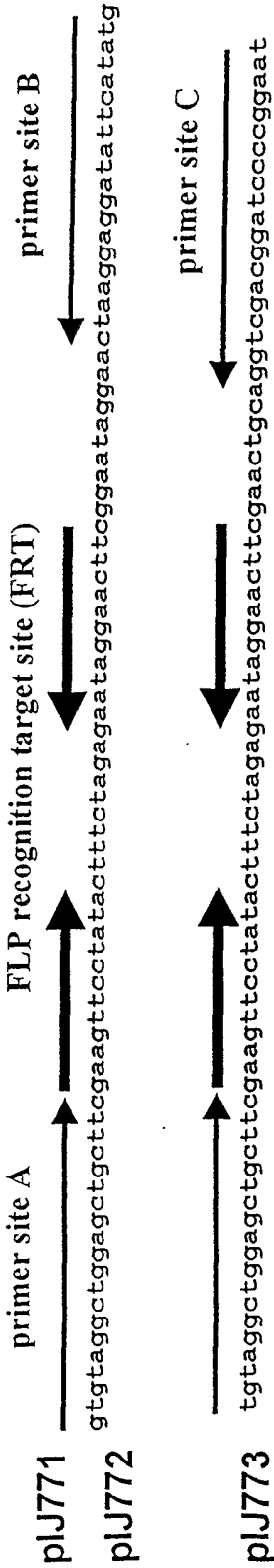
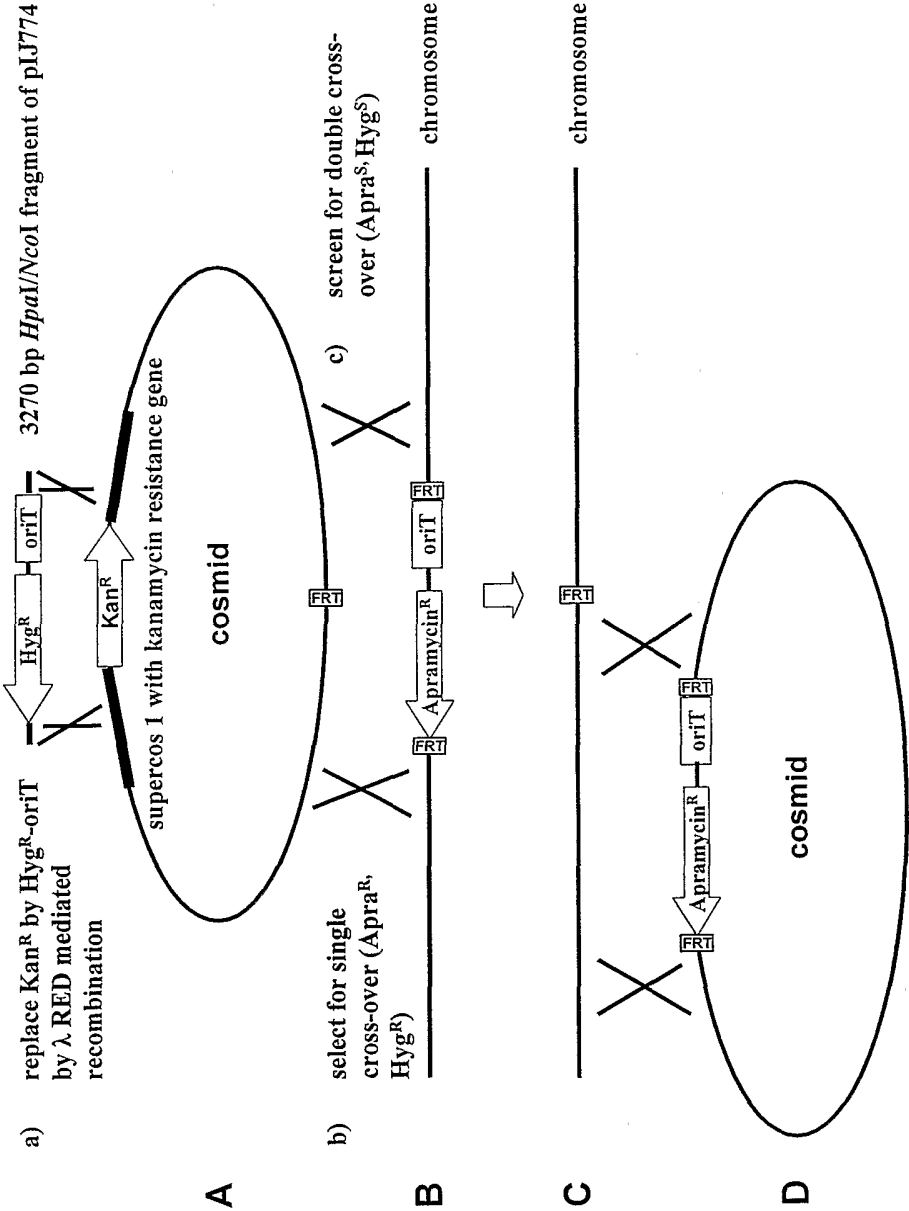


Fig. 7



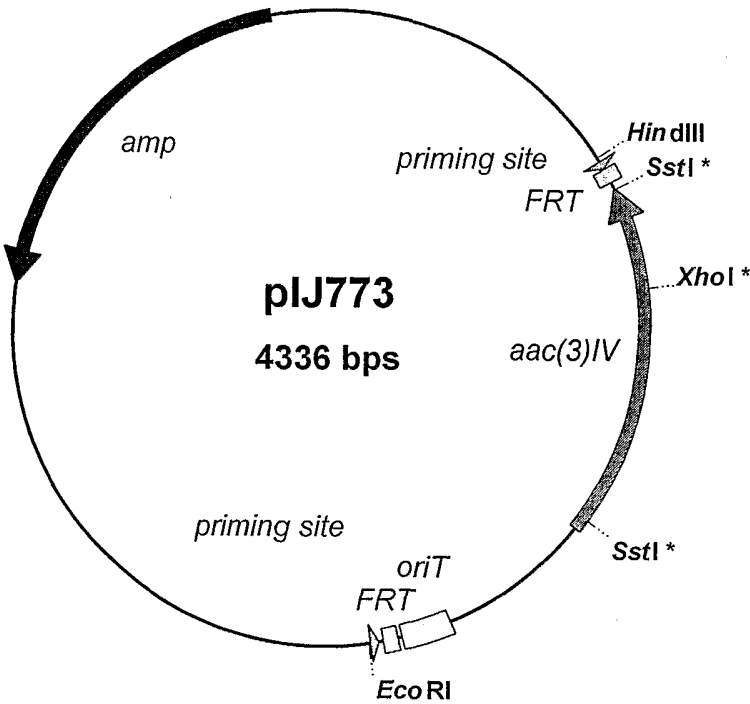


Fig. 8A

15/24

Fig. 8B

ORIGIN

```

1  GGAAATTGTA AACGTTAATA TTTTGTAAAT ATTCGCGTTA AATTTTGTGT AAATCAGCTC
61  ATTTTTTAAAC CAATAGGCCG AAATCGGCAA AATCCCTTAT AAATCAAAAG AATAGACCGA
121 GATAGGGTTG AGTGTGTGTC CAGTTTGGAA CAAGAGTCCA CTATTAAAGA ACGTGGACTC
181 CAACGTCAAA GGGCGAAAAA CCGTCTATCA GGGCGATGGC CCACTACGTG AACCATCACC
241 CTAATCAAGT TTTTGGGGT CGAGGTGCCG TAAAGCACTA AATCGGAACC CTAAAGGGAG
301 CCCCCGATTT AGAGCTTGAC GGGGAAAGCC GGCGAACGTG GCGAGAAAGG AAGGGAAGAA
361 AGCGAAAGGA GCGGGCGCTA GGGCGCTGGC AAGTGTAGCG GTCACGCTGC GCGTAACCAAC
421 CACACCCGCC GCGCTTAATG CGCCGCTACA GGGCGCGTCG CGCCATTCGC CATTACGGCT
481 GCGCAACTGT TGGGAAGGGC GATCGGTGCG GGCTCTTCG CTATTACGCC AGCTGGCGAA
541 AGGGGGATGT GCTGCAAGGC GATTAGTTG GGTAAACGCC GGGTTTTCCT AGTCACGACG
601 TTGTAAACG ACGGCCAGTG AATTGTAATA CGACTCACTA TAGGGCGAAT TGGGTACCGG
661 GCCCCCCTC GAGGTGACG GTATCGATAA GCTTGATGTG TAGGCTGGAG CTGCTTCGAA
721 GTTCCTATAC TTTCTAGAGA ATAGGAACCT CGGAATAGGA ACTTCATGAG CTCAGCCAAT
781 CGACTGGCGA GCGGCATCGC ATTCTTCGCA TCCCGCCCTC TGGCGGATGC AGGAAGATCA
841 ACGGATCTCG GCCCAGTTGA CCCAGGCGTG TCGCCACAAT GTCGCGGGAG CGGATCAACC
901 GAGCAAAGGC ATGACCGACT GGACCTTCCT TCTGAAGGCT CTTCTCCTTG AGCCACCTGT
961 CCGCCAAGGC AAAGCGCTCA CAGCAGTGGT CATCTCGAG ATAATCGACG CGTACCAACT
1021 TGCCATCCTG AAGAATGGTG CAGTGTCTCG GCACCCATA GGAACCTTT GCCATCAACT
1081 CGGCAAGATG CAGCGTCGTG TTGGCATCGT GTCCACGCC GAGGAGAAGT ACCTGCCCAT
1141 CGAGTTCATG GACACGGGCG ACCGGGCTTG CAGGCGAGTG AGGTGGCAGG GGCAATGGAT
1201 CAGAGATGAT CTGCTCTGCC TGTGGCCCCG CTGCCGAAA GGCAATGGA TGGGCGCTGC
1261 GCTTTACATT TGGCAGGCGC CAGAATGTGT CAGAGACAAC TCCAAGTCC GGTGTAACGG
1321 GCGACGTGGC AGGATCGAAC GGCTCGTCGT CCAGACCTGA CCACGAGGGC ATGACGAGCG
1381 TCCCTCCCGG ACCCAGCGCA GCACGCAGGG CCTCGATCAG TCCAAGTGGC CCATCTTCGA
1441 GGGGCCGGAC GCTACGGAAG GAGCTGTGGA CCAGCAGCAC ACCGCCGGG GTAAACCCAA
1501 GGTTGAGAAG CTGACCGATG AGCTCGGCTT TCCGCCATTC GTATTGCACG ACATTGCACT
1561 CCACCGCTGA TGACATCAGT CGATCATAGC ACGATCAACG GCACTGTTGC AAATAGTCGG
1621 TGGTGATAAA CTTATCATCC CTTTTGTGTC ATGGAGCTGC ACATGAACCC ATTCAAAGGC
1681 CGGCATTTTC AGCGTGACAT CATTCTGTGG GCCGTACGCT GGTACTGCAA ATACGGCATC
1741 AGTTACCGTG AGCTGCATTT TCCGCTGCAT AACCTGCTT CGGGGTCATT ATAGCGATTT
1801 TTTCCGTATA TCCATCCTTT TTCGCACGAT ATACAGGATT TTGCCAAAGG GTTCGTGTAG
1861 ACTTTCCTTG GTGTATCCAA CGGCGTCAGC CGGGCAGGAT AGGTGAAGTA GGCCACCCG
1921 CGAGCGGGTG TTCCTTCTTC ACTGTCCCTT ATTCGCACCT GGCGGTGCTC AACGGGAATC
1981 CTGCTCTGCG AGGCTGGCGG GAACTTCGAA GTTCTATAC TTTCTAGAGA ATAGGAACTT
2041 CGAACTGCAG GTCGACGGAT CCCCAGGAATA TCGAATTCCT GCAGCCCGGG GGATCCACTA
2101 GTTCTAGAGC GGCCGCCACC GCGGTGGAGC TCCAGCTTTT GTTCCCTTTA GTGAGGGTTA
2161 ATTCCGAGCT TGGCGTAATC ATGGTCATAG CTGTTTCCTG TGTGAAATTG TTATCCGCTC
2221 ACAATTCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA AAGCCTGGGG TGCCTAATGA
2281 TCGAGCTAAC TCACATTAAT TCGTTGCGC TCACTGCCCG CTTTCCAGTC GGGAAACCTG
2341 TCGTGCCAGC TGCATTAATG AATCGGCCAA CGCGCGGGGA GAGCGGGTTT GCGTATTGGG
2401 CGCTCTTCG CTTCCTCGCT CACTGACTCG CTGCGCTCGG TCGTTCCGGCT GCGGCGAGCG
2461 GTATCAGCTC ACTCAAAGGC GGTAAATACG TTATCCACAG AATCAGGGGA TAACGCAGGA
2521 AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG
2581 GCGTTTTTCC ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG
2641 AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC
2701 GTGCGCTCTC CTGTTCCGAC CCGCCGCTT ACCGGATACC TGTCCGCCTT TCTCCCTTCG
2761 GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCGGT GTAGGTGCTT
2821 CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTGAGC CCGACCGCTG CGCCTTATCC
2881 GGTAACATAT GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC
2941 ACTGCTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG
3001 TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA
3061 GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC CGCTGGTAGC
3121 GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA AAAAAGGATC TCAAGAAGAT
3181 CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGACG AAAACTCACG TTAAGGGATT
3241 TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC TTTTAAATTA AAAATGAAGT
3301 TTTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTACCA ATGCTTAATC
3361 AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTCGTTTCAT CCATAGTTGC CTGACTCCCC
3421 GTCGTGTAGA TAACTACGAT ACGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA

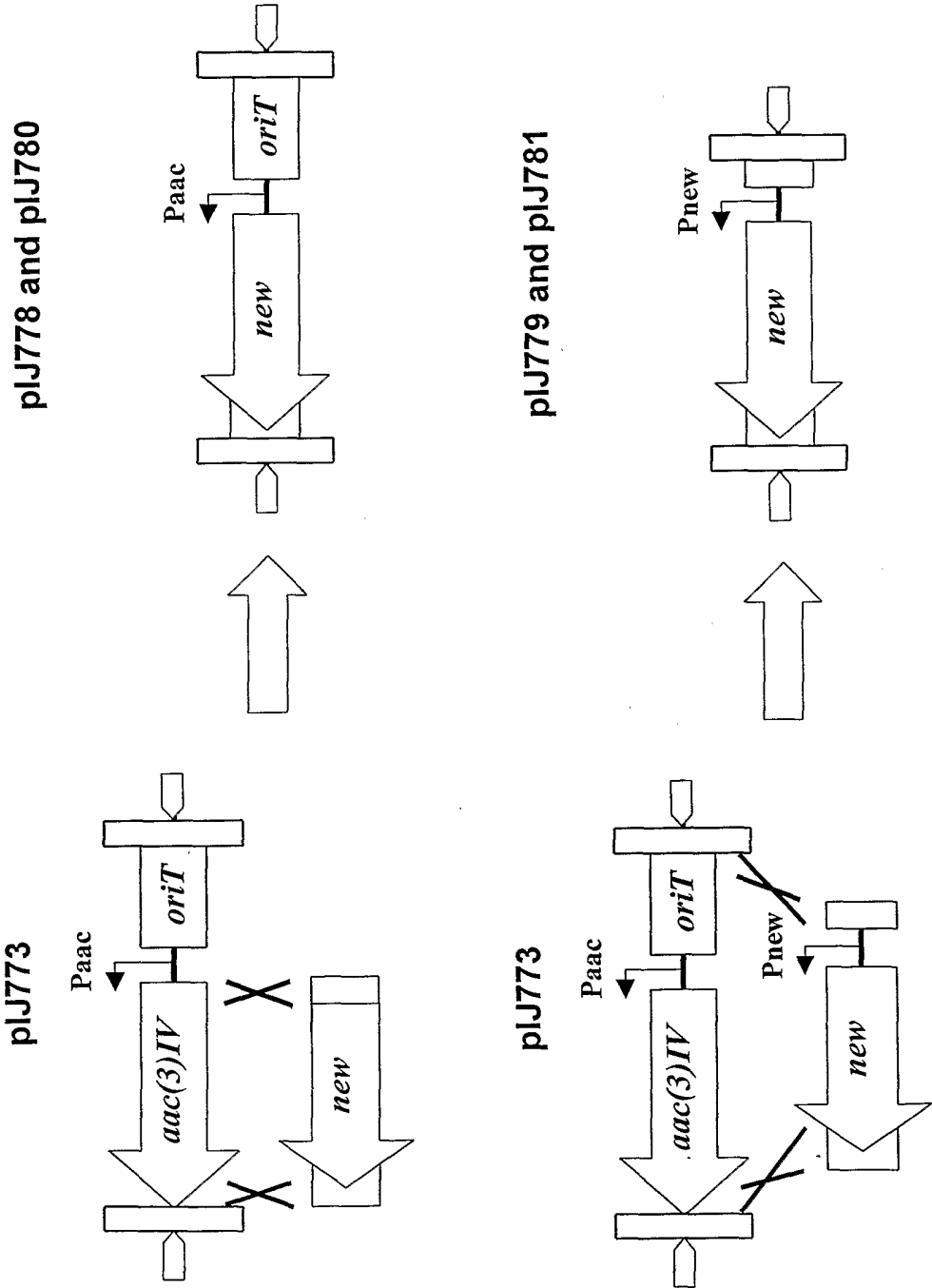
```

16/24

```
3481 CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA TAAACCAGCC AGCCGGAAGG
3541 GCCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA TCCAGTCTAT TAATTGTTGC
3601 CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT AATAGTTTGC GCAACGTTGT TGCCATTGCT
3661 ACAGGCATCG TGGTGTACG CTCGTCGTTT GGTATGGCTT CATTCAGCTC CGGTCCCCAA
3721 CGATCAAGGC GAGTTACATG ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT
3781 CCTCCGATCG TTGTCAGAAG TAAGTTGGCC GCAGTGTTAT CACTCATGGT TATGGCAGCA
3841 CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC
3901 TCAACCAAGT CATTCTGAGA ATAGTGATG CGGCGACCGA GTTGCTCTTG CCCGGCGTCA
3961 ATACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAAG TGCTCATCAT TGGAAAACGT
4021 TCTTCGGGGC GAAACTCTC AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC
4081 ACTCGTGAC CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGCGTTTC TGGGTGAGCA
4141 AAAACAGGAA GGCAAAATGC CGCAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA
4201 CTCATACTCT TCCTTTTCA ATATTATTGA AGCATTATC AGGGTTATTG TCTCATGAGC
4261 GGATACATAT TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTCCGCG CACATTTCCT
4321 CGAAAAGTGC CACCTG
```

Fig. 8B (cont.)

Fig. 9



18/24

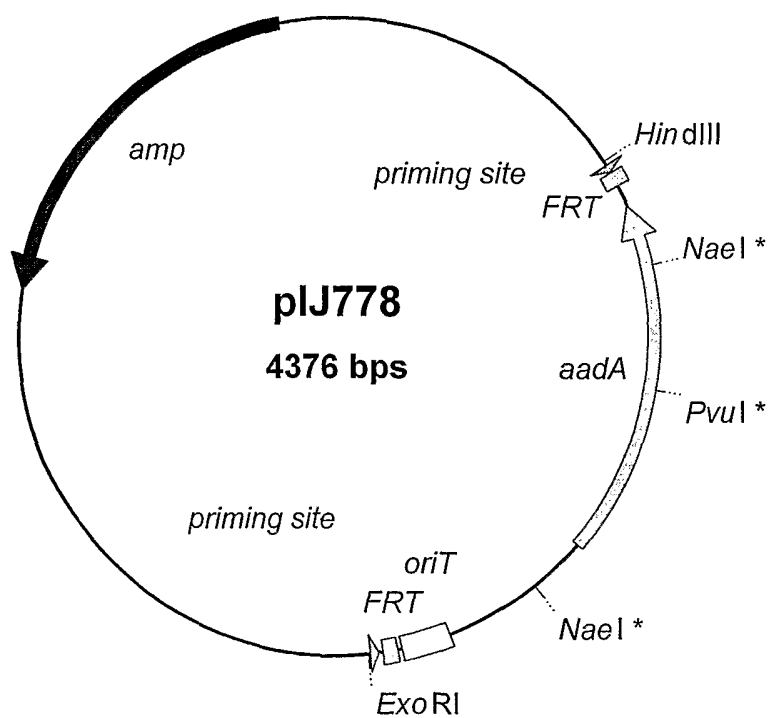


Fig. 10A

19/24

Fig. 10B

ORIGIN

```

1  GGAAATTGTA AACGTTAATA TTTTGTAAAT ATTCGCGTTA AATTTTTGTT AAATCAGCTC
61  ATTTTTTAAC CAATAGGCCG AAATCGGCAG AATCCCTTAT AAATCAAAAG AATAGACCGA
121 GATAGGGTTG AGTGTGTGTC CAGTTTGGAA CAAGAGTCCA CTATTAAAGA ACGTGGACTC
181 CAACGTCAAA GGGCGAAAAA CCGTCTATCA GGGCGATGGC CCACTACGTG AACCATCACC
241 CTAATCAAGT TTTTGGGGT CGAGGTGCCG TAAAGCACTA AATCGGAACC CTAAAGGGAG
301 CCCCCGATTT AGAGCTTGAC GGGGAAAGCC GGCGAACGTG GCGAGAAAGG AAGGGAAGAA
361 AGCGAAAGGA GCGGGCGCTA GGGCGCTGGC AAGTGTAGCG GTCACGCTGC GCGTAACCAC
421 CACACCCGCC GCGCTTAATG CGCCGCTACA GGGCGCGTCG CGCCATTGCG CATTAGGCT
481 GCGCAACTGT TGGGAAGGGC GATCGGTGCG GGCCTCTTCG CTATTACGCC AGCTGGCGAA
541 AGGGGGATGT GCTGCAAGGC GATTAAGTTG GGTAACGCCA GGGTTTTCCC AGTCACGACG
601 TTGTAAAACG ACGGCCAGTG AATTGTAATA CGACTCACTA TAGGGCGAAT TGGGTACCGG
661 GCCCCCCTC GAGGTCGACG GTATCGATAA GCTTGATGTG TAGGCTGGAG CTGCTTCGAA
721 GTTCTATAC TTTCTAGAGA ATAGGAACCT CGGAATAGGA ACTTCATGAG CTCAGCCAAT
781 CGATGGCGA GCGGCATCTA TTTGCCGATA ACCTTGGTGA TCTGCGCTTT CCGTAGTGG
841 ACAAAATTCTT CCAACTGATC TCGCGCGCAG GCCAAGCGAT CTTCTTCTTG TCCAAGATAA
901 GCCTGTCTAG CTTCAAGTAT GACGGGCTGA TACTGGGCCG GCAGGCGCTC CATTGCCAG
961 TCGGCAGCGA CATCCTTCGG CGCGATTTTG CCGGTTACTG CGCTGTACCA AATGCGGGAC
1021 AACGTAAGCA CTACATTTG CTCATCGCCA GCCCAGTCGG GCGCGAGGTT CCATAGCGTT
1081 AAGGTTTCAT TTAGCGCCTC AAATAGATCC TGTTCAAGAA CCGGATCAAA GAGTTCCTCC
1141 GCCGCTGGAC CTACCAAGGC AACGCTATGT TCTCTTGCTT TTGTCAGCAA GATAGCCAGA
1201 TCAATGTCTG TCGTGGCTGG CTCGAAGATA CCTGCAAGAA TGTCAATTGC CTGCCATTCT
1261 CCAAATTGCA GTTCGCGCTT AGCTGGATAA CGCCACGGAA TGATGTCGTC GTGCACAACA
1321 ATGGTGAAGT CTACAGCGCG GAGAATCTCG CTCTCTCCAG GGAAGCCGA AGTTTCCAAA
1381 AGGTGCTTGA TCAAAGCTCG CCGGTTGTT TCATCAAGCC TTACGGTCAC CGTAACCAGC
1441 AAATCAATAT CACTGTGTGG CTTAGGCCG CCATCCACTG CCGAGCCGTA CAAATGTACG
1501 GCCAGCAACG TCGGTTGCG AGTGGCGCTC ATGACGCCAA CTACCTCTGA TAGTTGAGTC
1561 GATACTTCGG CGATCACC GCATCACCCTT ACATTGCACT CCACCGCTGA TGACATCAGT
1621 CGATCATAGC ACGATCAACG GCACTGTTGC AAATAGTCGG TGGTGATAAA CTTATCATCC
1681 CCTTTTGCTG ATGGAGCTGC ACATGAACCC ATTCAAAGGC CGGCATTTTC AGCGTGACAT
1741 CATTCTGTGG GCCGTACGCT GGTACTGCAA ATACGGCATC AGTTACCGTG AGCTGCATTT
1801 TCCGCTGCAT AACCTGCTT CGGGGTCATT ATAGCGATTT TTTCGGTATA TCCATCCTTT
1861 TTCGCACGAT ATACAGGATT TTGCCAAAGG GTTCGTGTAG ACTTTCCTTG GTGTATCCAA
1921 CGGCGTCAGC CGGGCAGGAT AGGTGAAGTA GGCCACCCG CGAGCGGGTG TTCCTTCTTC
1981 ACTGTCCCTT ATTCGCACCT GCGGTGCTC AACGGGAATC CTGCTCTGCG AGGCTGGCGG
2041 GAATCTCGAA GTTCTATACA TTTCTAGAGA ATAGGAACCT CGAACTGCG AGTCAGGAT
2101 CCCCAGGAATA TCGAATTCTT GCAGCCCGGG GGATCCACTA GTTCTAGAGC GGCCGCCACC
2161 GCGGTGGAGC TCCAGCTTTT GTTCCCTTTA GTGAGGGTTA ATTCGAGCT TGGCGTAATC
2221 ATGGTCATAG CTGTTTCTTG TGTGAAATTG TTATCCGCTC ACAATTCCAC ACAACATACG
2281 AGCCGGAAGC ATAAAGTGTA AAGCCTGGGG TGCCTAATGA GTGAGCTAAC TCACATTAAT
2341 TCGGTTGCGC TCACTGCCCG CTTTCCAGTC GGGAAACCTG TCGTGCCAGC TGCATTAATG
2401 AATCGGCCAA CGCGCGGGGA GAGGCGGTTT GCGTATTGGG CGCTCTTCCG CTTCTCGCT
2461 CACTGACTCG CTGCGCTCGG TCGTTCCGCT GCGGCGAGCG GTATCAGCTC ACTCAAAGGC
2521 GGTAATACGG TTATCCACAG AATCAGGGGA TAACGCAGGA AAGAACATGT GAGCAAAAGG
2581 CCAGCAAAAG GCCAGGAACC GTAAAAGGC GCGGTTGCTG GCGTTTTTCC ATAGGCTCCG
2641 CCCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG AGGTGGCGAA ACCCGACAGG
2701 ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC GTGCGCTCTC CTGTTCCGAC
2761 CCTGCCGCTT ACCGGATACC TGTCCGCCTT TCTCCCTTCG GGAAGCGTGG CGCTTTCTCA
2821 TAGTCAACGC TGTAGGTATC TCAGTTCCGT GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT
2881 GCACGAACCC CCGTTTACG CCGACCGCTG CGCCTTATCC GGTAACATC GTCTTGAGTC
2941 CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC ACTGGTAACA GGATTAGCAG
3001 AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG TGGCCTAACT ACGGCTACAC
3061 TAGAAGGACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA GTTACCTTCG GAAAAAGAGT
3121 TGGTAGCTCT TGATCCGGCA AACAAACCAC CGCTGGTAGC GGTGGTTTTT TTGTTTGCAA
3181 GCAGCAGATT ACGCGCAGAA AAAAAGGATC TCAAGAAGAT CCTTTGATCT TTTCTACGGG

```

fig. 108 (cont.)

20/24

```
3241 GTCTGACGCT CAGTGAACG AAAACTCACG TTAAGGGATT TTGGTCATGA GATTATCAAA
3301 AAGGATCTTC ACCTAGATCC TTTTAAATTA AAAATGAAGT TTTAAATCAA TCTAAAGTAT
3361 ATATGAGTAA ACTTGGTCTG ACAGTTACCA ATGCTTAATC AGTGAGGCAC CTATCTCAGC
3421 GATCTGTCTA TTTCGTTCAT CCATAGTTGC CTGACTCCCC GTCGTGTAGA TAACTACGAT
3481 ACGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA CCGCGAGACC CACGCTCACC
3541 GGCTCCAGAT TTATCAGCAA TAAACCAGCC AGCCGGAAGG GCCGAGCGCA GAAGTGGTCC
3601 TGCAACTTTA TCCGCCTCCA TCCAGTCTAT TAATTGTTGC CGGGAAGCTA GAGTAAGTAG
3661 TTCGCCAGTT AATAGTTTGC GCAACGTTGT TGCCATTGCT ACAGGCATCG TGGTGTACG
3721 CTCGTCGTTT GGTATGGCTT CATTAGCTC CGGTTCCCAA CGATCAAGGC GAGTTACATG
3781 ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT CCTCCGATCG TTGTCAGAAG
3841 TAAGTTGGCC GCAGTGTAT CACTCATGGT TATGGCAGCA CTGCATAATT CTCTTACTGT
3901 CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC TCAACCAAGT CATTCTGAGA
3961 ATAGTGATG CGGCGACCGA GTTGCTCTTG CCCGGCGTCA ATACGGGATA ATACCGCGCC
4021 ACATAGCAGA ACTTTAAAAG TGCTCATCAT TGGAAAACGT TCTTCGGGGC GAAACTCTC
4081 AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC ACTCGTGCAC CCAACTGATC
4141 TTCAGCATCT TTTACTTTCA CCAGCGTTTC TGGGTGAGCA AAAACAGGAA GGCAAAATGC
4201 CGCAAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA CTCATACTCT TCCTTTTCA
4261 ATATTATTGA AGCATTATC AGGGTTATTG TCTCATGAGC GGATACATAT TTGAATGTAT
4321 TTAGAAAAAT AAACAAATAG GGGTCCGCG CACATTTCCC CGAAAAGTGC CACCTG
```

//

21/24

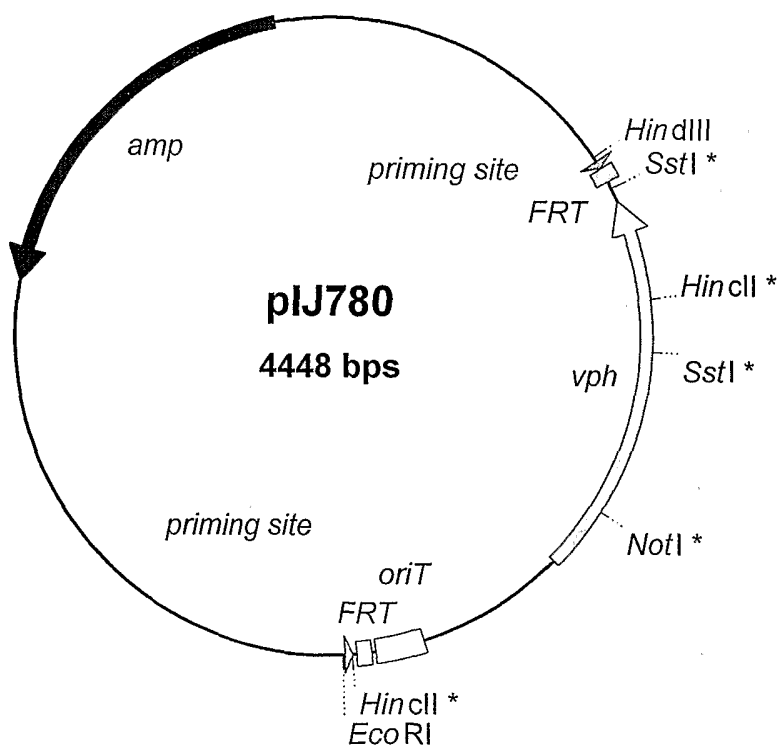


Fig. 11A

22/24

Fig. 11B

ORIGIN

```
1  GGAAATTGTA AACGTTAATA TTTTGTTAAA ATTTCGCGTTA AATTTTGTGT AAATCAGCTC
61  ATTTTTTAAAC CAATAGGCCG AAATCGGCAA AATCCCTTAT AAATCAAAAG AATAGACCGA
121 GATAGGGTTG AGTGTGTGTT CAGTTTGGA CAAGAGTCCA CTATTAAAGA ACGTGGACTC
181 CAACGTCAAA GGGCGAAAAA CCGTCTATCA GGGCGATGGC CCACTACGTG AACCATCACC
241 CTAATCAAGT TTTTGGGGT CGAGGTGCCG TAAAGCACTA AATCGGAACC CTAAAGGGAG
301 CCCCCGATTT AGAGCTTGAC GGGGAAAGCC GGCGAACGTG GCGAGAAAGG AAGGGAAGAA
361 AGCGAAAGGA GCGGGCGCTA GGGCGCTGGC AAGTGTAGCG GTCACGCTGC GCGTAACCAC
421 CACACCCGCC GCGCTTAATG GCGCGCTACA GGGCGCGTCG CGCCATTGCG CATTCAGGCT
481 GCGCAACTGT TGGGAAGGGC GATCGGTGCG GGCTCTTTCG CTATTACGCC AGCTGGCGAA
541 AGGGGGATGT GCTGCAAGGC GATTAAGTTG GGTAACGCCA GGGTTTTCCC AGTCACGACG
601 TTGTAAAACG ACGGCCAGTG AATTGTAATA CGACTCACTA TAGGGCGAAT TGGGTACCGG
661 GCCCCCCTC GAGGTGCGAC GTATCGATAA GCTTGATGTG TAGGCTGGAG CTGCTTCGAA
721 GTTCTATAC TTTCTAGAGA ATAGGAACTT CGGAATAGGA ACTTCATGAG CTCAGCCAAT
781 CGACTGGCGA GCGGCATCTA CCGGTAGCCG CTGAGGCCGT CGGCGAGTTC CTCTCGTCCG
841 CCGTCGCGCT GCGCGTAGAG GGCCTGCTGG AGTGCGAAGG TGCTCGGAT CGCCGAGATC
901 CGCTCGGCCG TTCCGTTGTC GGCCAGCCG CCGAGCGCGA GCACTCGGCC CAGCAGTTCC
961 TCGCCGTAGC TCGCCCCGAT GGCGGCCAGG TCCTCAGCCG GGTGCGCGAT GCCGACCTCG
1021 TCCCAGTCGA CGACGCCGCT CATGCGCGCG ACTCCGTCCA CCGTCTCCA CAGGACGTTT
1081 TCGCCGCCGA GGTCACCGTG GACCACCGCG GAGGTGAGAT GGGGCAGGGC GTGAGCGCG
1141 GCGAGCTCGC GCTCGGCACG CTCCCGGCCG CCGTCGGACA TCAGCGGGAA CAGTTCGGTA
1201 CGCACCCCCG TGGCGAACTC CTGCCACTCG TTCGCGGGAG CCTCCGGCAG CGCGGCGCGC
1261 ACCTTCTCCT CGTCGCCCCG CGCCGCGAGC CCGGACAGCA GGGTCGCGTA CTGTGCGGCG
1321 ACGGCCTCCG CCACCTCCGG GCTGGTGAGC ACATCGTCCT CCAACGGTGC TCCGGGAATG
1381 CGGCTCAGCA CCAGGTACGG CGGCTCGTCC GTGCCCTGGG CGCCGCCCTC GGACAGCGGC
1441 TGCGGCGTGC GAAACCCGAG GTCGATCCCG GCAAGAGCGC GCAGGACGTC CGCCCTGCCG
1501 GGCAGACGGT CGGCGGCCCG CCGGGTGCGG GCGAAGCAGA CCACCCGGTG CGATCCGATC
1561 ACCACATGGT GGAAGTGGCC CTCGTGGACG GCGAGTCCGC CCACGGTGTG CCCGGGCAGG
1621 GACTCGCTCA GCAGATCGCG GTGCGTCTCA ATGATTCTCA TGACATTGCA TCCACCCTG
1681 GATGACATCA GTCGATCATA GCACGATCAA CGGCACTGTT GCAAATAGTC GTGTGATGATA
1741 AACTTATCAT CCCCTTTTGC TGATGGAGCT GCACATGAAC CCATTCAAAG GCCGGCATTT
1801 TCAGCGTGAC ATCATTCTGT GGGCCGTACG CTGGTACTGC AAATACGGCA TCAGTTACCG
1861 TGAGCTGCAT TTTCCGCTGC ATAACCCTGC TTCGGGGTCA TTATAGCGAT TTTTTCGGTA
1921 TATCCATCCT TTTTCGCACG ATATACAGGA TTTTGCCAAA GGGTTTCGTG AGACTTTCCT
1981 TGGTGTTATC AACGGCGTCA GCCGGGCAGG ATAGGTGAAG TAGGCCACC CGCGAGCGGG
2041 TGTTCTTCT TCACTGTCCC TTATTCGCAC CTGGCGGTGC TCAACGGGAA TCCTGCTCTG
2101 CGAGGCTGCG GGGAACTTCG AAGTTCCTAT ACTTCTAGA GAATAGGAAC TTCGAACTGC
2161 AGGTGACGCG ATCCCGGAA TATCGAATTC CTGCAGCCCG GGGGATCCAC TAGTCTAGA
2221 GCGGCCGCCA CCGCGGTGGA GCTCCAGCTT TTGTTCCCTT TAGTGAGGGT TAATTCCGAG
2281 CTTGGCGTAA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT TGTTATCCGC TCACAATTCC
2341 ACACAACATA CGAGCCGGAA GCATAAAGTG TAAAGCCTGG GGTGCCTAAT GAGTGAGCTA
2401 ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTTCAG TCGGGAAACC GTGCGTGCCA
2461 GCTGCATTAA TGAATCGGCC AACGCGCGG GAGAGGCGGT TTGCGTATTG GCGCTCTTC
2521 CGCTTCTCTG CTA CTGACT GCTGCGCT GGTGCTTCG CTGCGGCGAG CCGTATCAGC
2581 TCACTCAAAG GCGGTAATAC GGTTATCCAC AGAATCAGGG GATAACGCAG GAAAGAACAT
2641 GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA CCGTAAAAAG GCCGCGTTGC TGGCGTTTTT
2701 CCATAGGCTC CGCCCCCTG ACGAGCATCA CAAAAATCGA CGCTCAAGTC AGAGGTGGCG
2761 AAACCCGACA GGA CTATAA GATACCAGG GTTTCCTCC GGAAGCTCCC TCGTGCGCTC
2821 TCCTGTTCCG ACCCTGCCGC TTACCGGATA CCTGTCCGCC TTTCTCCCTT CGGGAAGCGT
2881 GCGGCTTCT CATAGCTCAC GCTGTAGGTA TCTCAGTTCG GTGTAGGTCG TTCGCTCCAA
2941 GCTGGGCTGT GTGCACGAAC CCCCCGTCA GCCCGACCGC TGCGCCTTAT CCGGTAACATA
3001 TCGTCTTGAG TCCAACCCGG TAAGACACGA CTTATCGCCA CTGGCAGCAG CCACTGGTAA
3061 CAGGATTAGC AGAGCGAGGT ATGTAGGCGG TGCTACAGAG TTCTTGAAGT GGTGGCCTAA
3121 CTACGGCTAC ACTAGAAGGA CAGTATTTGG TATCTGCGCT CTGCTGAAGC CAGTTACCTT
3181 CGGAAAAAGA GTTGGTAGCT CTTGATCCGG CAAACAAACC ACCGCTGGTA GCGGTGTTTT
```

23/24

Fig. 11B (cont.)

```
3241 TTTTGTTC AAGCAGCAGA TTACGCGCAG AAAAAAAGGA TCTCAAGAAG ATCCTTTGAT
3301 CTTTCTACG GGGTCTGACG CTCAGTGGAA CGAAACTCA CGTTAAGGGA TTTTGGTCAT
3361 GAGATTATCA AAAAGGATCT TCACCTAGAT CCTTTTAAAT TAAAAATGAA GTTTTAAATC
3421 AATCTAAAGT ATATATGAGT AAACCTGGTC TGACAGTTAC CAATGCTTAA TCAGTGAGGC
3481 ACCTATCTCA GCGATCTGTC TATTCGTTC ATCCATAGTT GCCTGACTCC CCGTCGTGTA
3541 GATAACTACG ATACGGGAGG GCTTACCATC TGGCCCAGT GCTGCAATGA TACCGCGAGA
3601 CCCACGCTCA CCGGCTCCAG ATTTATCAGC AATAAACCCAG CCAGCCGGAA GGGCCGAGCG
3661 CAGAAGTGGT CCTGCAACTT TATCCGCCTC CATCCAGTCT ATTAATTGTT GCCGGGAAGC
3721 TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT GTTGCCATTG CTACAGGCAT
3781 CGTGGTGTCA CGCTCGTCGT TTGGTATGGC TTCATTGAGC TCCGGTTCCC AACGATCAAG
3841 GCGAGTTACA TGATCCCCCA TGTTGTGCAA AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT
3901 CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT ATCACTCATG GTTATGGCAG CACTGCATAA
3961 TTCTCTTACT GTCATGCCAT CCGTAAGATG CTTTCTGTG ACTGGTGAGT ACTCAACCAA
4021 GTCATTCTGA GAATAGTGTA TCGGCGCACC GAGTTGCTCT TGCCCGGCGT CAATACGGGA
4081 TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC ATTGGAAC GTTCTTCGGG
4141 GCGAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT TCGATGTAAC CCACTCGTGC
4201 ACCCAACTGA TCTTCAGCAT CTTTACTTT CACCAGCGTT TCTGGGTGAG CAAAAACAGG
4261 AAGGCAAAAT GCCGCAAAA AGGGAATAAG GGCGACACGG AAATGTTGAA TACTCATACT
4321 CTTCTTTTT CAATATTATT GAAGCATTTA TCAGGGTTAT TGTCTCATGA GCGGATACAT
4381 ATTTGAATGT ATTTAGAAAA ATAAACAAAT AGGGGTTCG CGCACATTTC CCCGAAAAGT
4441 GCCACCTG
```

//

24/24

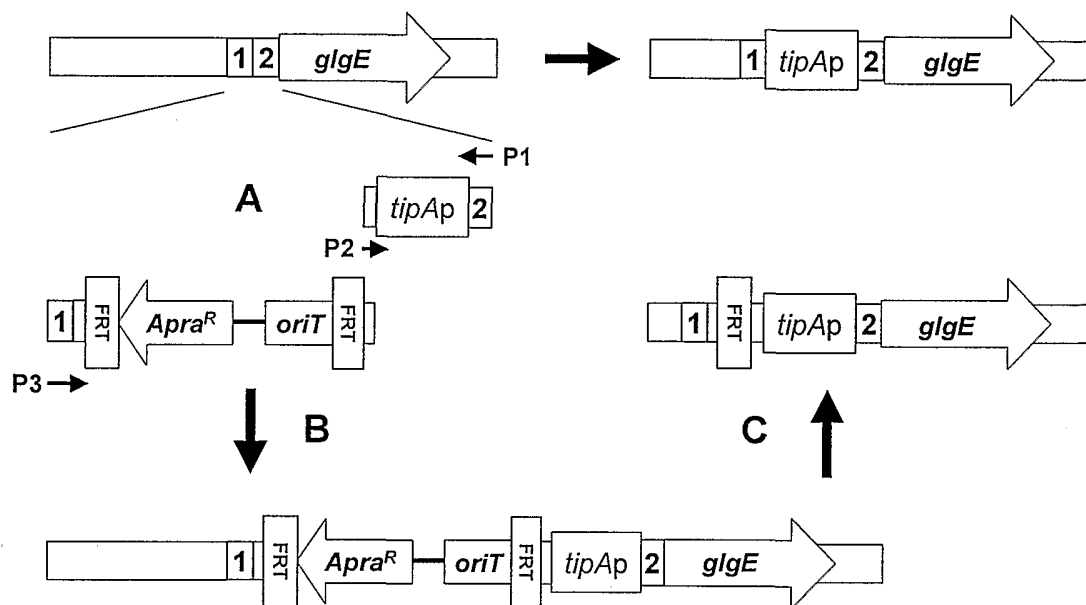


Fig. 12

INTERNATIONAL SEARCH REPORT

International Application No.

PC 1, vB 02/02798

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/10 C12N15/65 C12N15/76 C12N15/90 C12N1/21
//C12R1:465

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)

IPC 7 C12N

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 practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATSENKO KIRILL A ET AL: "One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 97, no. 12, 6 June 2000 (2000-06-06), pages 6640-6645, XP002210218 June 6, 2000 ISSN: 0027-8424 p. 6640; 6644/6645; Figures 1 and 3 cited by the applicant --- -/--	1-60



Further documents are listed in the continuation of box C.



Patent family members are listed in 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 document 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

21 August 2002

Date of mailing of the international search report

04/09/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Kurz, B

INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 02/02798

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GEHRING, A. ET AL.: "Genomewide insertional mutagenesis in Streptomyces coelicolor reveals additional genes involved in morphological differentiation" PNAS, vol. 97, no. 17, 15 August 2000 (2000-08-15), pages 9642-9647, XP002210219 page 9642 -page 9643 ---	47,48, 50-53
X	WILD J ET AL: "Targeting and retrofitting pre-existing libraries of transposon insertions with FRT and oriV elements for in-vivo generation of large quantities of any genomic fragment" GENE, ELSEVIER BIOMEDICAL PRESS. AMSTERDAM, NL, vol. 223, no. 1-2, 26 November 1998 (1998-11-26), pages 55-66, XP004153577 ISSN: 0378-1119 abstract; p. 56 ---	47,48, 50-53
A	WO 01 31000 A (OCHIAI MISA ;SUNTORY LTD (JP); ASHIKARI TOSHIHIKO (JP)) 3 May 2001 (2001-05-03) abstract	11,12
P,A	-& EP 1 142 998 A (SUNTORY LIMITED) 10 October 2001 (2001-10-10) abstract ---	11,12
A	KITANI SHIGERU ET AL: "Conjugal transfer of plasmid DNA from Escherichia coli to Streptomyces lavendulae FRI-5." JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY, vol. 10, no. 4, August 2000 (2000-08), pages 535-538, XP001097428 ISSN: 1017-7825 the whole document ---	13
A	FLETT FIONA ET AL: "High efficiency intergeneric conjugal transfer of plasmid DNA from Escherichia coli to methyl DNA-restricting streptomycetes." FEMS MICROBIOLOGY LETTERS, vol. 155, no. 2, 15 October 1997 (1997-10-15), pages 223-229, XP002210221 ISSN: 0378-1097 the whole document -----	28

INTERNATIONAL SEARCH REPORT

formation on patent family members

Int: onal Application No

PCT/GB 02/02798

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0131000 A	03-05-2001	JP 2001120276 A	08-05-2001
		AU 7958500 A	08-05-2001
		EP 1142998 A1	10-10-2001
		WO 0131000 A1	03-05-2001