



GPS[®]-1

Genome Priming System

A tool for generating sequence data by placing a transposon (Transprimer™) containing universal priming sites into target DNA at random locations *in vitro*

I n s t r u c t i o n M a n u a l

Catalog #E7100S
Store at -20°C



Version 2.5
10/03

Table of Contents:

Kit Components 2

Customer Supplied Materials 3

Introduction 4

Strategy 8

GPS®-1 Reaction Protocol 11

Protocol Tips 13

Appendices

 Appendix I: Kit Component Composition 14

 Appendix II: Target DNA Requirements 15

 Appendix III: Theory of the Transposition Reaction 16

 Functions of the Transposase Proteins

 Tn7 Target Immunity and its Consequences

 Strand Transfer

 Appendix IV: Properties of Reaction Products 18

 Appendix V: Recipes 19

 Appendix VI: Frequently Asked Questions (FAQ) 20

 Appendix VII: Additional Information 30

References 35

Kit Components Sold Separately/Companion Products 36

Notice to Buyer/User: While the Transprimer donors pGPS1.1 and pGPS2.1 may be used for all described purposes, both research and commercial, these donors may not be resold in their original or modified form without the written permission of New England Biolabs, Inc.

Kit Components:

Each kit contains sufficient reagents for 10 reactions.

■ 10X GPS Buffer – (Reagent 1)	0.5 ml
■ pGPS1.1 (Transprimer-1 Donor Plasmid) – (Reagent 2a)	15 µl (20 µg/ml)
■ pGPS2.1 (Transprimer-2 Donor Plasmid) – (Reagent 2b)	15 µl (20 µg/ml)
■ TnsABC* Transposase (10 reactions) – (Reagent 3)	10 µl
■ Start Solution – (Reagent 4)	200 µl
■ Control Target Plasmid (LITMUS 28) – (Reagent 5)	10 µl (80 µg/ml)

Sequencing Primers

■ Primer S for the Left (Spe I) End of the Transprimer – (Reagent 6)	100 µl (3.2 pmol/µl)
■ Primer N for the Right (Not I) End of the Transprimer – (Reagent 7)	100 µl (3.2 pmol/µl)
■ Instruction Manual	

Note: pGPS1.1 and pGPS2.1 replace pGPS1 and pGPS2 found in previous versions of the GPS-1 Kit. The Transprimers in these donor plasmids are identical. i.e., pGPS1.1 carries Transprimer-1 and pGPS2.1 carries Transprimer-2. For more information, see Appendix VI: Frequently Asked Questions.

Customer Supplied Materials

Target DNA, 0.08 µg per reaction, (> 5 µg/ml)

Other equipment and materials needed are standard laboratory supplies. Particular considerations apply to:

- Water bath or heat block, 75°C, for killing the reaction. **Note: not 65°C.**

- Transformable cells:

Any standard kanamycin- or chloramphenicol-sensitive *E. coli* strain can be used. DNA preparations are generally of better quality when the host carries a mutation in the major nonspecific nuclease, *endA*. When the target plasmid carries repetitive DNA, a recombination-deficient (*recA*) host is preferred. Commercially available competent or electrocompetent cells may be used.

Chemically competent *E. coli* cells will normally give enough product transformants to complete a sequencing project on ordinary plasmid targets. The chemical transformation procedure introduces a strong bias against large molecules (over 20 kb). Consequently, for large targets such as cosmids, electroporation is recommended even when large numbers of colonies are not needed.

Electrocompetent *E. coli* cells give about 500-fold more transformants than do chemically competent cells. This method is also recommended for large targets, because it does not bias recovery to small (deletion) derivatives that may have accumulated in the population (7,8).

- Media:

Use Luria Bertani Broth without antibiotic for outgrowth after transformation, or as recommended.* For selective plating, use Luria Bertani Agar with antibiotic or equivalent.

Recipes are found in the Appendix V, page 19; these are standard media. Selective plates must include kanamycin (20 µg/ml final concentration) to select for the Transprimer-1 element, or chloramphenicol (15 µg/ml) to select for the Transprimer-2 element.

For growing cultures for plasmid preps, use Luria Bertani Broth* with 20 µg/ml kanamycin or 15 µg/ml chloramphenicol.

*When using commercially-supplied transformable cells or plasmid prep kits, follow the manufacturer's recommended procedure.

Introduction:

The GPS®-1 Genome Priming System provides a simple *in vitro* method for generating a population of DNA sequencing templates with randomly interspersed primer-binding sites. This system is useful for sequencing DNAs that require more than one sequencing reaction to determine the entire sequence. GPS-1 is a faster alternative to primer walking, random subcloning and nested deletion methods; it can also be used for mapping projects.

GPS-1 is a Tn7 transposon-based *in vitro* system which uses TnsABC* Transposase to insert a transposon (Transprimer™) randomly into the DNA target (1,2,3) (Figure 1). There is minimal site preference for Transprimer insertion and, due to target immunity, only one insertion occurs per target DNA molecule (4). Therefore, the *in vitro* reaction produces a population of target DNA molecules each containing the Transprimer element at a different position. The GPS-1 System includes two Transprimer transposons—one encodes kanamycin resistance and the other encodes chloramphenicol resistance. This allows choice of antibiotic for selection of product DNA molecules. Since the transposon donor plasmid is unable to replicate (*ori*⁻) in ordinary lab strains of *E. coli* (5,6), only products (target DNA molecules containing Transprimer insertions) can be recovered. Unique priming sites on both ends of the Transprimer element, together with supplied primers (Primer N and Primer S), allow DNA sequence to be obtained from both strands of the target DNA at the position of the insertion. If gaps in the sequence remain after this initial random phase of sequencing, unique restriction sites within the Transprimer element enable mapping of insertions to specifically target any missing sequence.

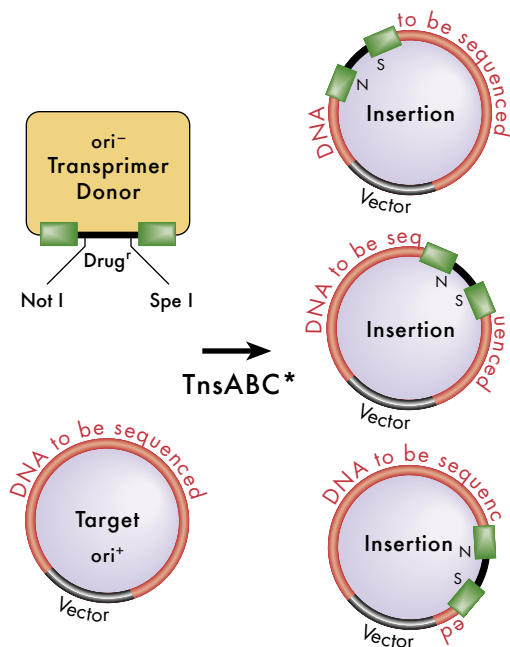
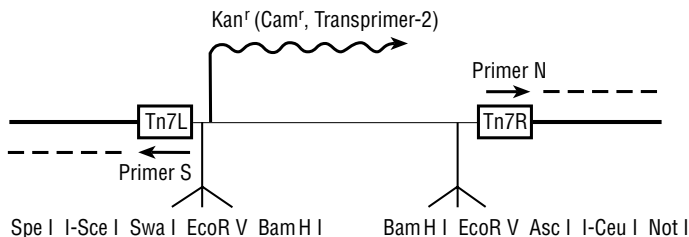


Figure 1: Overview of GPS-1 Genome Priming System



Left Transprimer end (Tn7L) and Primer S (Spe I end)

< TGTGGGCGGACAAAATAGTTGGGAAC TGGGAGGGGTGGAATGGAGTTTTTAAGGATTATTTAGGGAAGA..
 ACACCCGCCTGTTTTATCAACCTTGACCC TCCCACTTTACCTCAAAAATTCCTAATAATCCCTTCT..
 -TCCCACTTTACCTCAAAAATTCCTAATA-5'...Primer S

Right Transprimer end (Tn7R) and Primer N (Not I end)

Primer N... 5' -ACTTTATTGTCATAGTTTAGATCTATTTTG-
 ..TCTAGTTTAAGACTTTATTGTCATAGTTTAGATCTATTTTGTTTCAGTTTAAGACTTTATTGTCCGCCACA
 ..AGATCAAATCTGAAATAACAGTATCAATCTAGATAAAACAAGTCAAATCTGAAATAACAGGCGGGTGT >

Figure 2: Transprimer Schematic: Transprimer-1 and locations of sequencing primers (Primer N and Primer S). Transprimer-1 expresses Kan^r and Transprimer-2 expresses Cam^r .

Advantages of the GPS Method:

- Can be used with plasmids, cDNAs, cosmids and large targets such as BACs.
- Transprimer is inserted randomly into the target DNA. There are no “hot spots” for insertion (3).
- Choice of drug resistance: Kan^r or Cam^r Transprimers included.
- The entire protocol takes less than 90 minutes (Figure 3, page 12).
- Two parallel sequencing reactions can be performed on one insertion template.
- Unique rare-cutting restriction enzyme sites (Not I, Pme I, I-Ceu I, Swa I, Spe I) within the Transprimer facilitate insert mapping. This is particularly useful when sequencing through repetitive DNA.
- High efficiency: 10^3 – 10^6 insertion products are recovered per reaction depending on whether a chemical transformation method or electroporation is used.
- Faster than primer walking, random subcloning and nested deletion methods.

Strategy:

This kit enables the generation of new sequence information from double strand templates using transposon insertion to introduce primer-binding sites at random, but mappable, locations.

Two basic strategies can be used to sequence large fragments of DNA—"random" and "directed". The GPS kit can be used for both approaches. With a random approach, the DNA to be sequenced is subjected to Transprimer insertion reactions, templates are prepared from transformants and sequenced immediately without further analysis. This allows the rapid acquisition of new sequence in the early stages and saves on primer synthesis because universal primers are used. Sequence comparison then allows assembly of contiguous sequence (contigs) by overlapping the sequence obtained from different insertions. The redundancy of sequence obtained from these different priming sites can be very helpful in resolving ambiguities found in any individual run. These ambiguities may persist in repeated runs from the same priming site but may disappear when sequence is obtained from a different priming site.

However, due to the intrinsically uneven distribution of stochastic events, final gap closure by a random approach is laborious. Consequently, more directed approaches to sequencing through remaining gaps are usually used for finishing. A commonly used tactic is to switch to a directed approach when the average level of sequence redundancy reaches 3–5 fold.

The mapping capabilities of the Transprimer elements allow a semidirected approach. The random insertions can be mapped either using rare sites in the Transprimer or by fingerprinting using frequent cutters not found in the Transprimer. Insertions suitably spaced for sequence determination can then be chosen for sequencing. The insertions at contig ends, derived from a random sequencing phase, can also be used to provide convenient sites for subcloning of gap-specific DNA.

Obtaining Transprimer insertions is extremely simple and fast. The following steps may help organize your approach.

1. Choose a donor. For your convenience, donors for both Kan^r (pGPS1.1) and Cam^r (pGPS2.1) Transprimers are included.
2. Choose a transformation method. Chemical transformation and electroporation will both give adequate numbers of recovered colonies for sequencing. Electroporation is preferable for large targets and in applications requiring more than 1,000 independent insertions or when more convenient.
3. Determine how many insertions you wish to sequence. For plasmid or cosmid sequencing, this will be tens to hundreds, depending on the size of your target and the level of sequence redundancy you want. See Appendix VI (page 20) for guidance on choosing an approximate number.
4. Decide on a suitable plasmid preparation procedure that will give sequencing grade template for double-stranded sequencing.
5. Decide whether you want to map the insertions before choosing which to sequence. This will depend on three factors:
 - i. The size of the fragment of interest compared with the dispensable portion of the vector. See Appendix VII (page 30) for the sizes of some selectable markers and replication origins, from which insertions will be excluded. The larger the vector:insert ratio, the more likely the need to map insertions.
 - ii. The relative cost to you of sequencing extraneous templates versus restriction mapping.
 - iii. The number and length of repetitive sequences likely to be present. If repetitive sequences are numerous or similar in length to the average sequencing run, then locating insertions by mapping can help constrain the sequence assembly process.

6. Decide what enzymes to use for restriction mapping, if required. Appendix VII (page 30) includes a list of unique sites in the Transprimer and sites not found in it.
7. Remember that the sequencing primers provided (Figure 2, page 6) are not 5' labeled. They can be radiolabeled or used with labeled nucleotides.
8. The reaction process duplicates a five-base pair sequence at the site of insertion (see Figure 4). We have not seen insertions without this duplication. This is useful to recall when assembling the sequence.

GPS-1 Reaction Protocol

- Mix the following reagents (per 20 μ l reaction):

2 μ l	10X GPS Buffer (Reagent 1)
1 μ l	pGPS1.1 or pGPS2.1 Donor DNA (0.02 μ g) (Reagent 2a or 2b)
variable vol	0.08 μ g Target DNA
variable vol	dH ₂ O
<hr/>	
18 μ l	Total Volume

Mix well by pipetting up and down a few times.

- Add 1 μ l TnsABC* Transposase (Reagent 3) to each tube. Mix again.
- Incubate 10 minutes at 37°C (30°C for BAC targets). This is the assembly reaction.
- Add 1 μ l Start Solution (Reagent 4) to each tube. Mix well by pipetting up and down a few times.
- Incubate 1 hour 37°C (30°C for BAC targets). This is the strand transfer reaction.
- Heat inactivate at 75°C for 10 minutes. **Note: 65°C is not adequate.**
- Transform

For chemical transformation with subcloning efficiency cells (10^7 per microgram of pUC), transform 1 μ l and 10 μ l of undiluted reaction.

For electroporation ($> 10^9$ per microgram of pUC), dilute 10-fold in dH₂O and transform 1 μ l and 10 μ l.

To outgrow, dilute the transformation mixture into 1 ml LB or as directed by the manufacturer, and incubate for **1 hour at 37°C** with aeration. This period without selection is necessary for expression of drug resistance.

8. Plate the outgrown transformations on selective plates. Include plates with drug combinations appropriate to select for the target plasmid alone (e.g., ampicillin) and for the target with a Transprimer insertion (e.g., kanamycin plus ampicillin). For the target with insertion (e.g., Kan + Amp), plate 10 μ l and 100 μ l. One of these plates should have around 100 colonies. For the target drug alone, plates at this dilution should show confluent growth; if you wish to obtain a number for the competence of the cells under these conditions, dilute the outgrowth 1,000-fold in 0.85% saline and plate the same volumes.
9. Purify plasmid templates and sequence, using Primer N and Primer S (Reagents 6 and 7).

Notes: Usually you will need a few tens to a few hundred insertions. More information can be found in Appendix VI: Frequently Asked Questions.

We recommend performing a positive control (add supplied Control Target Plasmid [LITMUS 28, Amp^r, 2,823 bp], donor plasmid and TnsABC) and a negative control (add experimental target and donor DNA, no TnsABC*).*

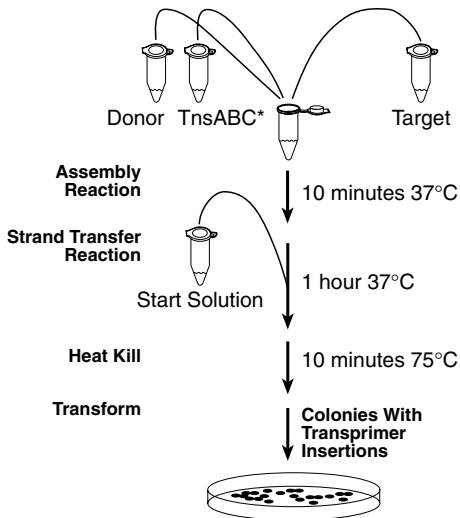


Figure 3: Reaction Overview

■ Protocol Tips

Amount of Cosmid and BAC Targets: A recommended mass of target DNA of 0.08 µg per reaction works well for plasmid targets. For cosmids and BACs, a molar ratio of around 2:1 (donor to target) works well. Increasing the ratio to 4:1 decreases the efficiency slightly.

Donor:Target Ratio: A recommended donor:target mass ratio of 1:4 (0.08 µg target per 20 µl reaction) is optimal. Small deviations produce only small changes in the number of recovered products. However, saturating amounts of donor will inhibit the reaction and may lead to accumulation of double insertions.

Order of Addition: Water, target DNA, GPS Buffer (Reagent 1), and Donor Plasmid (Reagent 2) should be added first, followed by TnsABC* Transposase (Reagent 3). Start solution (Reagent 4) should be added only after the assembly step (step 3).

Assembly Reaction: If this step is omitted, the proportion of complicated products will be increased.

Time of Incubation: The reaction is linear at 37°C for at least one hour. Extremely long incubation times may lead to accumulation of additional double insertions.

Temperature of Incubation: The reaction proceeds, but more slowly, at room temperature and at 30°C. For reactions with BACs, 30°C is recommended.

Heat killing: Heating at 75°C for 10 minutes effectively disrupts the reaction complexes. Few or no transformants will be obtained if this step is omitted. Heating at 65°C for 20 minutes **is not** adequate. Phenol/chloroform extraction followed by alcohol precipitation is also effective.

Scaling the Procedure: Increase or reduce the final volume and the volume of all components by the same percentage; the relative concentrations of the two DNA species and the proteins are very important, as are the buffer conditions.

Choice of Donor: If your target expresses Cam^r, use Transprimer-1 (Kan^r); if it expresses Kan^r (Neo^r), use Transprimer-2 (Cam^r). Otherwise, the choice is yours.

Appendix I: Kit Component Composition

Stock Concentration

■ 10X GPS Buffer (Reagent 1)

Tris-HCl

250 mM pH 8.0

DTT

20 mM

ATP

20 mM

■ pGPS1.1 (Transprimer-1 Donor plasmid) (Reagent 2a)

20 µg/ml

For a restriction map, see Figure 5, Appendix VII. Sequence is available at <www.neb.com>. pGPS1.1 does not replicate in ordinary lab strains of *E. coli*.

■ pGPS2.1 (Transprimer-2 Donor plasmid) (Reagent 2b)

20 µg/ml

For a restriction map, see Figure 6, Appendix VII. Sequence is available at <www.neb.com>. pGPS2.1 does not replicate in ordinary lab strains of *E. coli*.

■ TnsABC* Transposase (Reagent 3):

TnsA

7 µg/ml

TnsB

10 µg/ml

TnsC*

20 µg/ml

Supplied as a mix in 50% glycerol storage buffer.

Store at -20°C. Sufficient protein for 10 separate reactions is provided. At the time of use, keep on ice until ready to add to the reaction, and keep on ice until returned to the freezer.

■ Start Solution (Reagent 4): magnesium acetate

300 mM

■ Control Target (Reagent 5): LITMUS 28

80 µg/ml

■ Sequencing Primers

3.2 pmol/µl in dH₂O

Primer N

Anneals to the top strand of Transprimer as in Figure 2

Primer S

Anneals to the bottom strand of Transprimer as in Figure 2

Appendix II: Target DNA Requirements

Target DNA attributes:

- A. Plasmid targets for sequencing should be in circular form to facilitate recovery. Linear DNA is an efficient substrate, but a repair and ligation step is required before transformation (Figure 4, page 17).
- B. The target should be sensitive either to kanamycin (use Transprimer-1) or to chloramphenicol (use Transprimer-2). Note that the Neo marker on cosmids pWE15 and Supercos2 is expressed in *E. coli*. Kanamycin resistance is mediated by neomycin phosphotransferase (designated variously as Kan^r, Kn, Neo, *npt*). Chloramphenicol resistance is mediated by chloramphenicol transacetylase (Cam^r, Cm^r, *cat*).
- C. Large plasmids, such as cosmids and BACs are usable targets. We have successfully obtained sequence from cosmids and BACs. See Appendix VI for more information.

Target DNA concentration:

Target DNA must be at least 5 µg/ml in a no-salt buffer such as 1X TE. Concentration can be estimated by comparison of agarose gel band intensity with a DNA of known concentration or any other method.

Appendix III: Theory of the Transposition Reaction

■ Functions of the transposase proteins

TnsB specifically recognizes the Transprimer ends, which are derived from transposon Tn7. TnsC binds to target DNA and interacts with TnsB. TnsA binds to DNA:TnsB. When all three proteins are suitably bound in a three-protein, two-DNA complex, TnsA and TnsB together carry out the strand transfer reaction. For a full review of this topic, see reference 1. The reaction used here is enabled by a mutation in TnsC (TnsC*) that bypasses the usual requirement for TnsD or TnsE (2).

■ Tn7 Target immunity and its consequences

The TnsABC* Transposase will not efficiently use a target molecule already containing Tn7 ends; as a result, double insertions are minimized. *In vivo*, target immunity operates at distances of at least 190 kb. At present, the model for this immunity invokes an activity of TnsC, mediated by its ATPase activity and its specific recognition of DNA:TnsB complexes. TnsC* binds nonspecifically to DNA and surveys it. If TnsB is encountered (i.e., if a Tn7 end is already present), the incipient strand transfer complex is disrupted by an ATP-mediated activity of TnsC. For more extensive discussion of this property, see reference 4.

■ Strand Transfer (Figure 4, page 17)

Transposition proceeds via two DNA scissions and two transesterification reactions. For simplicity of representation, this is presented as six cleavages and two ligations. Donor DNA is cleaved three bases 5' to the transposon end in one strand and precisely at the transposon 3' end in the other strand (Figure 4); this occurs on each side of the transposon. A five-base staggered cleavage is made in the target. The 3' hydroxyls of the transposon are ligated to the 5' phosphates of target. The net result is excision of the transposon from the donor, together with three-base 5' extensions on each end and addition of the transposon to the target at the opening. *In vivo*, the three-base flap of donor origin is resected, and the resulting gap filled in and ligated. A five-base duplication of target sequence results from this process. The covalent linkages labeled with a star are those involved in or produced by transesterification reactions between the 3' transposon end and the 5' end of the target.

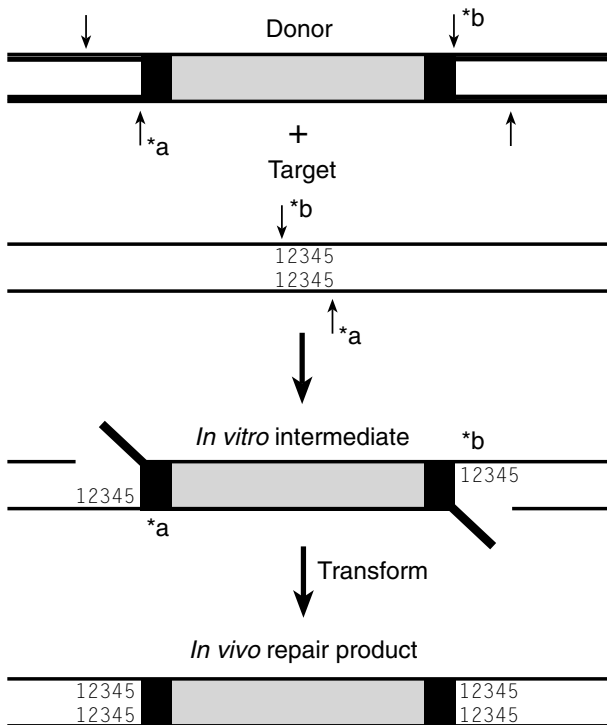


Figure 4: Strand Transfer Reaction. Note that 5 bp of target sequence will be duplicated at the insertion site: the same 5 bp will appear to the left of Tn7L and to the right of Tn7R (Appendix III, page 16).

Appendix IV: Properties of Reaction Products

Random insertion:

In trials at New England Biolabs, 63 sequenced insertions were found at 62 separate sites. The distribution of insertions in 100 bp intervals was statistically indistinguishable from a random (Poisson) distribution (3). Physical detection methods suggest that insertion can be obtained at every internucleotide linkage. This means that useful sequence information can be obtained from almost all products recovered, even without first locating the insertions by restriction digest.

Simple insertion:

Single simple insertions are the principal products of the TnsABC* transposition reaction. Of 74 insertions examined by restriction digest, 73 were single simple insertions.

Other products, which are produced at a significant level by some other transposition systems, give unreadable sequence. These more complicated products include double insertions and cointegrates between the donor and target replicons.

Few double insertions are seen with TnsABC* Transposase presumably because of target immunity (see page 16) (4). Cointegrates are rare in reactions that include the assembly step (Step 3) in the GPS-1 Protocol. However, a high fraction of products were cointegrates when the assembly step (Step 3) was omitted (i.e., when Mg^{++} is included in the starting buffer or is added before the TnsABC* Transposase).

Target site duplication:

The reaction process duplicates a five-base pair sequence at the site of insertion (see Figure 4). We have not seen insertions without this duplication. This is useful to recall when assembling the sequence.

Appendix V: Recipes

■ **Luria Bertani Broth, per liter** (for outgrowth following transformation)

Tryptone (Difco)	10 g
Yeast Extract (Difco)	5 g
NaCl	5 g
NaOH (1 N)	2 ml
autoclave	

■ **Luria Bertani Agar with drug, per liter** (for plating transformation outgrowth cultures)

Tryptone (Difco)	10 g
Yeast Extract (Difco)	5 g
NaCl (Baker)	5 g
NaOH (1 N)	2 ml
Agar (Difco)	15 g
autoclave	

Add (depending on choice of donor plasmid) after autoclaving and cooling to 55°C, per liter:

Kanamycin	20 mg
or	
Chloramphenicol	15 mg

■ **0.85% saline, per liter** (for diluting cells after transformation)

NaCl	8.5 g
autoclave	

■ **1X TE, per liter** (for diluting DNA)

1 M Tris-HCl pH 8.0	10 ml
0.5 M EDTA pH 8.0	2 ml

Appendix VI: Frequently Asked Questions

■ *Where will the insertions be?*

Insertions will not be found in the drug resistance gene of the target if it has been selected for, nor in the origin of replication of your target plasmid. They will be randomly located within all the rest of your plasmid in all of the unselected DNA.

■ *How do I find the molecules with insertions?*

After transformation and a period of outgrowth to allow expression of drug resistance, plate on media with kanamycin (20 µg/ml) or chloramphenicol (15 µg/ml) depending on which Donor Plasmid (pGPS1.1 or pGPS2.1) was used.

■ *Do I have to get rid of the donor molecules?*

No. The donor plasmid will not replicate in ordinary lab strains of *E. coli*. The origin of replication is from the transmissible plasmid R6K, and it depends on a replication initiation protein of plasmid origin (the Π protein, product of the *pir* gene) not normally present in laboratory strains (5,6).

■ *What happens if I heat kill at 65°C instead of 75°C?*

The number of colonies will be reduced by at least 100-fold.

■ *Can I use T4 DNA Ligase Buffer instead of GPS Buffer? Its composition looks very similar.*

Ligase Buffer should not be used. Since Mg^{++} is present from the beginning of the reaction, the assembly step (page 11) does not occur. We also find a high fraction of cointegrate molecules (replicon fusions) when Ligase Buffer is used.

- *Can I use a different Mg^{++} salt instead of magnesium acetate for the Start Solution?*
Yes, $MgSO_4$ and $MgCl_2$ have been tested and are acceptable.

- *Should I use electroporation or chemical transformation?*

Use chemical transformation if it is convenient for you and your target is small (< 20 kb). Do not use this procedure with large plasmids since chemical transformation selects against large molecules, and the colonies obtained with such targets will be enriched for those carrying deletions. For smaller targets, the procedure should yield plenty of insertions for sequencing purposes since over 1,000 transformants can be obtained from one reaction.

Use electroporation if it is more convenient, your target is large (> 20 kb) or you need more than 1,000 insertions for your application (see next question to determine this).

- *How many transformants should I look at?*

For the random sequencing phase, usually you will need a few tens to a few hundred insertions. To get a better feeling for this, fill out the table below (round to the nearest hundred base pairs). To answer question C, see page 34, where the sizes of some commonly used drug resistance genes and replication origins are given. For a discussion of random sequencing approaches, see reference 9.

- A. Average sequence run length obtained by your double stranded sequencing protocol: _____
- B. Number of base pairs in your target plasmid, including insert: _____
- C. Number of base pairs in your target plasmid subject to selection (the origin of replication and the drug resistance marker in your target that you plan to select for): _____

- D. Number of unselected base pairs in your target plasmid: $B - C =$ _____
- E. Number of base pairs in your target plasmid that you want to sequence: _____
- F. To cover all unselected sequence in your plasmid to a depth of 5-fold (9), you will need:

$$N_1 = \frac{5 \times D}{2 \times A} \text{ insertions}$$

- G. To cover all of the sequence of interest to you, you will need:

$$N_2 = \frac{5 \times E}{2 \times A} \text{ insertions}$$

If $N_1 \approx N_2$ or sequence runs are much cheaper than restriction mapping procedures, collect N_1 insertions and proceed directly to sequencing.

If $N_1 \gg N_2$ or sequence runs are more expensive than restriction mapping procedures, collect N_1 insertions and map them using the rare restriction sites in the Transprimer until N_2 of them have been located within or close to the sequence of interest.

These numbers assume a 5-fold depth of coverage; that is, each base pair within the sequenced segment is sequenced from an average of five different priming sites. You can choose a different depth of coverage by changing the number 5 in the equations in F and G.

Example 1: Short fragment, short run

- A. Average run length 300 bp
- B. 1.5 kb fragment in LITMUS (2.8 kb) = 4,300
- C. Selected sequence (*bla* + origin) = 1,600 bp
- D. Unselected sequence = 2,700
- E. 1,500 bp

$$N_1 = 5 \times 2,700 / 2 \times 300 = 22$$

$$N_2 = 5 \times 1,500 / 2 \times 300 = 12$$

$N_1 > N_2$: *map first*

Example 2: Short fragment, long run

- A. Average run length 500
- B. 1.5 kb fragment in LITMUS = 4,300
- C. Selected sequence (*bla* + origin) = 1,600 bp
- D. Unselected sequence = 2,700
- E. 1,500 bp

$$N_1 = 5 \times 2,700 / 2 \times 500 = 13$$

$$N_2 = 5 \times 1,500 / 2 \times 500 = 7$$

$N_1 > N_2$: *map first*

Example 3: Large fragment, long run

- A. Average run length 500 bp
- B. 40 kb fragment of interest in a cosmid of 8.1 kb = 48,100
- C. Selected sequence (*bla* + origin) = 1,600 bp
- D. Unselected sequence = 48,100 – 1,600 = 46,500.
- E. Sequence of interest = 40,000

$$N_1 = 5 \times 46,500 / 2 \times 500 = 232$$

$$N_2 = 5 \times 40,000 / 2 \times 500 = 200$$

$N_1 \approx N_2$: *sequence directly*

■ *What fraction of the target molecules will get insertions during the reaction?*

This depends in part on the size of the target molecule since 0.08 µg of a 3 kb plasmid has ten times as many molecules as the same mass of a 30 kb plasmid. Using the recommended procedure with a 2.8 kb target plasmid, we calculate that 2.3% of target molecules gained a Transprimer insertion. For a 30 kb plasmid, the same number of insertions distributed over a smaller number of molecules would give a larger fraction of the target with insertions.

■ *Will all of my reaction products be simple insertions?*

Multiple transformants are seen occasionally (~5% of transformants when chemically-competent cells are used). That is, a transformant receives both an unaltered target plasmid and a target plasmid with a Transprimer insertion. In sequencing applications, transformants with an extraneous unaltered target molecule will not be a problem because the primer will not anneal. The signal would be low however because the concentration of DNA with the priming site will be only about half of the total DNA concentration. Accurate quantitation can be achieved by gel analysis following linearization with one of the rare restriction enzymes in the Transprimer. For additional information about reaction products, visit <www.neb.com>.

■ *What E. coli strain should I use for transformation of the GPS reaction mix?*

Any standard *E. coli* strain can be used as long as it does not contain the same antibiotic-resistance as the Transprimer in the GPS reaction being transformed. A general recommendation is to use whichever strain you would normally use for making minipreps.

■ *What conditions are recommended for sequencing reactions using Primer N and Primer S?*

Standard sequencing reaction conditions work well using these primers. For example, a 50°C annealing temperature is used with these primers with an ABI Sequencer.

■ *Can GPS be used with BACs?*

Yes. Recommendations for use of GPS with BACs are similar to those for cosmid targets. A donor:target molar ratio of 2:1 is effective. As an example, for a standard reaction using 0.02 µg donor, we recommend using 0.1 µg of a 50 kb cosmid and 0.3 µg of a 150 kb BAC. Electroporation should always be used with such large molecules since using chemically-competent cells will tend to favor deletions. For BAC targets, we have often found that performing the GPS reaction at 30°C increases its efficiency. It is useful to spread a small amount of the transformation mix onto a plate selecting only for the BAC target DNA. This ensures that the BAC prep is good and that transformation was efficient.

■ *Can I easily sequence cosmids and BACs containing Transprimer insertions?*

Yes. We have obtained excellent sequencing results by following these general recommendations:

Cosmids

per 20 μ l sequencing reaction:

500 ng–1 μ g DNA

4 μ l Primer N or Primer S at 3.2 pmol/ μ l.

Thermal cycling:

5 minutes at 96°C

25–50 cycles of:

30 seconds at 96°C

15 seconds at 50°C

3 minutes at 60°C

BACs

Protocol from Boysen, C., Simon, M. & Hood, L. (1997) *Biotechniques* 23, 978–982.

Per 40 μ l sequencing reaction:

1–2 μ g DNA

50 pmol Primer N or Primer S

Double amount of other components

Thermal Cycling:

5 minutes at 96°C

25–50 cycles of:

10 seconds at 96°C

5 seconds at 50°C

4 minutes at 60°C

- *What are the differences between pGPS1.1 and pGPS2.1 and their predecessors pGPS1 and pGPS2?*

The only difference between the new donors and the original donors is in the vector backbone. The Transprimers are identical. i.e., pGPS1.1 carries Transprimer-1 and pGPS2.1 carries Transprimer-2. pGPS1.1 is 4814 bp, pGPS2.1 is 4490 bp. The increase in donor plasmid backbone size further decreases the already very low incidence of unusual GPS reaction products. The GPS reaction protocols have not changed. However, the increase in donor sizes should be taken into account if large targets such as cosmids or BACs are being used, where a 2:1 donor:target molar ratio is recommended.

- *Can I store unused GPS reaction mix?*

Yes. After the heat-kill step the reaction mix can be stored at -20°C and can be transformed later.

- *Can I use linear DNA as target DNA?*

Yes. Linear DNA is an efficient substrate for the GPS reaction. However, the reaction products must be ligated to DNA containing an origin of replication (e.g., vector DNA) in order to form a circular plasmid capable of forming colonies after transformation into *E. coli*.

- *Can I reduce the number of colonies formed by Transprimer insertions into the vector backbone of the target DNA?*

Yes. Selecting for markers, such as antibiotic-resistance, in the vector backbone of the target will prevent colony formation by Transprimer insertions within these markers. Also, the GPS reaction can be performed using linear target DNA such as a PCR product or an insert fragment. Products of this GPS reaction can then be ligated to vector DNA, transformed and plated on media selecting for the antibiotic resistance encoded by the Transprimer.

- *Where are the Primer N and Primer S sequences located within pGPS1.1 and pGPS2.1?*

Primer N: pGPS1.1 bp 3123–3152
 pGPS2.1 bp 3123–3152

Primer S: pGPS1.1 bp 4732–4761
 pGPS2.1 bp 4408–4437

- *What are the melting temperatures for Primer N and Primer S?*

Primer S: 82°C
Primer N: 74°C

- *Where can I find the entire sequences of the Transprimers and the donor plasmids?*

These are available at <www.neb.com>.

- *Does Transprimer-1 encode neomycin-resistance?*

Yes. The *npt* gene in Transprimer-1 encodes neomycin phosphotransferase, which confers resistance to neomycin and kanamycin.

- *Do the kanamycin- and chloramphenicol-resistance genes of pGPS1.1 and pGPS2.1 express in gram-positive organisms?*

No. The drug resistances in the GPS donor plasmids are not designed for expression in *B. subtilis* or other gram-positive bacteria.

- *Are donor plasmids available with Transprimers containing markers other than kanamycin-resistance or chloramphenicol-resistance?*

NEB does not currently sell donors with other markers. However, the donor plasmid pGPS3 (NEB #N7130S), a component of the GPS®-M Mutagenesis System, is designed to be easily manipulated to contain the marker of your choice. More information is available at <www.neb.com>.

- *Can GPS be used with very GC-rich, highly-repetitive target DNA?*

The GPS system may discriminate against regions with very high G + C (> 65%). In this situation we recommend digesting minipreps before sequencing to determine whether Transprimer insertions are within the target's vector backbone DNA or the DNA of interest. If the numbers of insertions within the vector are disproportionately high, it is possible to do the GPS reaction on linear DNA of interest and then ligate this to the vector, transform and plate on media selecting for the antibiotic-resistance of the Transprimer. This ensures that all Transprimer insertions will be within the DNA of interest and not the target's vector backbone. This technique can be applied to all types of target DNA.

- *Can I map Transprimer insertions using PCR?*

Yes. DNA from colonies can be analyzed using PCR. Since the Transprimer can insert in either orientation, PCR reactions can be performed using one primer complementary to the target DNA, and *both* GPS primers. Alternatively, two PCR reactions can be set up for each DNA, one with Primer N and one with Primer S.

Appendix VII: Additional Information

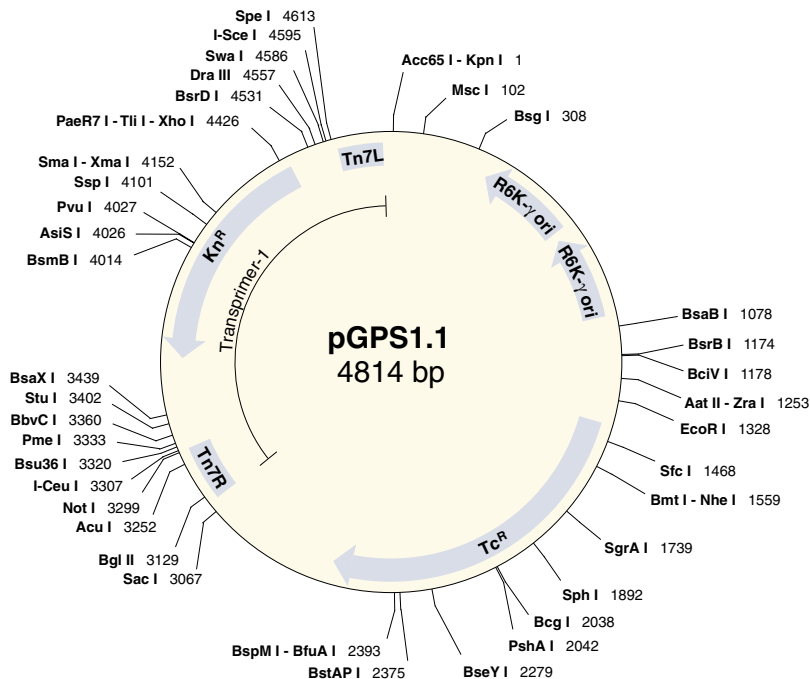
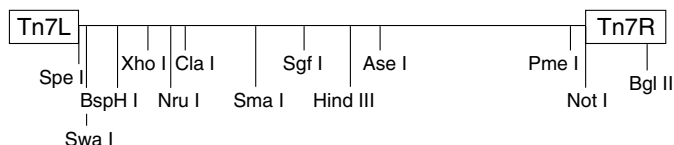


Figure 5: pGPS1.1

■ **Map of Transprimer-1 (Kan^r): 1699 bp**

(Unique palindromic restriction sites)



■ **Sites in the Transprimer-1 element:**

Enzymes that cut once:

Ase I, Ban II, BbvC I, Bgl II, BsaW I, BsmB I, BsmF I, Bsp1286 I, BspD I, BspH I, BsrD I, BsrF I, BssS I, Bsu36 I, Cla I, Dra III, Eco57 I, EcoN I, Hind III, I-Ceu I, I-Sce I, MspA1 I, Not I, Nru I, PaeR7 I, PflIM I, Pme I, Pvu I, Rsa I, Sgf I, Sma I, Sml I, Spe I, Ssp I, Stu I, Swa I, Xho I, Xma I

Enzymes that do not cut the Transprimer-1 element:

Aat II, Acc65 I, Acl I, Afl II, Afl III, Age I, Ahd I, AlwN I, Apa I, ApaL I, Asc I, Avr II, Ban I, Bbs I, Bbv I, Bcg I, Bci V I, Bcl I, Bgl I, Blp I, Bpm I, Bsa I, BsaA I, BsaB I, BsaH I, BseR I, Bsg I, BsiHKA I, BsiW I, BspE I, BspLU11 I, BspM I, BsrB I, BsrG I, BssH II, BstAP I, BstB I, BstE II, BstX I, BstZ17 I, Drd II, Eco47 III, EcoO109 I, EcoR I, Fau I, Fse I, Fsp I, Hae II, Hpa I, Kas I, Kpn I, Mfe I, Mlu I, Msc I, Msl I, Nae I, Nar I, Nco I, Nde I, NgoA IV, NgoM IV, Nhe I, Nsp I, Pac I, PflF I, Pml I, PpuM I, PshA I, Pst I, Pvu II, Rsr II, Sac I, Sac II, SanD I, Sap I, Sbf I, Sca I, SexA I, Sfc I, Sfi I, Sfo I, SgrA I, SnaB I, Sph I, Srf I, Sty I, Tat I, Tse I, Tth111 I, Xba I, Xcm I, Xmn I

Enzymes that cut once in pGPS1.1 but do not cut in Transprimer-1 element:

Aat II, Acc65 I, Bcg I, Bci V I, BsaB I, Bsg I, BspM I, BsrB I, BstAP I, EcoR I, Kpn I, Msc I, Nhe I, PshA I, Sac I, Sfc I, SgrA I, Sph I

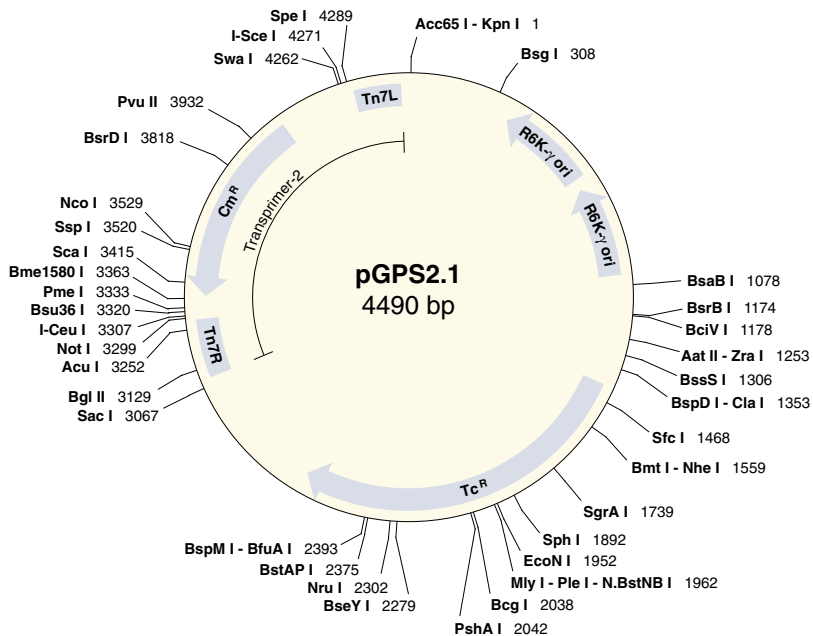
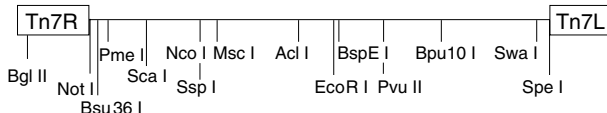


Figure 6: *pGPS2.1*

■ **Map of Transprimer-2 (Cam'):** 1375 bp

(Unique palindromic restriction sites)



■ **Sites in Transprimer-2 element:**

Enzymes that cut once:

Acl I, Apo I, Ava II, Ban I, Bgl II, Bpm I, Bpu 10 I, BsaA I, BsaW I, BseS I, BsiE I, BsmF I, Bsp1286 I, BspE I, BsrD I, BstDS I, BstF5 I, BstU I, Bsu36 I, Btg I, Eag I, Eco57 I, EcoR I, Fok I, Hinf I, HpyCH4 V, I-Ceu I, I-Sce I, Msc I, Nco I, Not I, Pme I, Pvu II, Sca I, Spe I, Ssp I, Sty I, Swa I, Tat I, Tfi I, Tsp R I

Enzymes that do not cut the Transprimer element:

Aat II, Acc65 I, Afl II, Afl III, Age I, Ahd I, AlwN I, Apa I, ApaL I, Asc I, Ase I, Ava I, Avr II, Ban II, Bbs I, Bbv I, BbvC I, Bcg I, BciV I, Bcl I, Bgl I, Blp I, Bsa I, BsaB I, BsaH I, BsaX I, BseR I, Bsg I, BsiHKA I, BsiW I, BsoB I, BspD I, BspH I, BspLU11 I, BspM I, BsrB I, BsrF I, BsrG I, BssH II, BssS I, BstAP I, BstB I, BstE II, BstX I, BstZ17 I, Cla I, Dra III, Drd II, Eco47 III, EcoN I, EcoO109 I, Fse I, Fsp I, Hae II, Hha I, HinP1 I, Hind III, Hpa I, Kas I, Kpn I, Mfe I, Mlu I, Nae I, Nar I, Nde I, NgoM IV, Nhe I, Nru I, Nsi I, Nsp I, Pac I, PaeR7 I, PflF I, Ple I, Pml I, PpuM I, PshA I, Pst I, Pvu I, Rsr II, Sac I, Sac II, SanD I, Sap I, Sbf I, SexA I, Sfc I, Sfi I, Sfo I, Sgf I, SgrA I, Sma I, Sml I, SnaB I, Sph I, Srf I, Stu I, Tse I, Tth111 I, Xba I, Xcm I, Xho I, Xma I, Xmn I

Enzymes that cut once in pGPS2.1 but do not cut in Transprimer-2 element:

Aat II, Acc65 I, Bcg I, BciV I, BsaB I, Bsg I, BspM I, BsrB I, BssS I, BstAP I, Cla I, EcoN I, Kpn I, Nhe I, Nru I, Ple I, PshA I, Sac I, Sfc I, SgrA I, Sph I

■ ***Approximate sizes of selectable marker genes and origins of replication***

Ampicillin resistance (β -lactamase, *bla*; including promoter): 950 bp

Tetracycline resistance (Tet, Tc): 1,200 bp

Neomycin, kanamycin resistance (neomycin phosphotransferase, *npt*, Kan, Neo) from Tn5: 1,000 bp (found in cosmids); from Tn903: 850 bp (found in pACYC184 and the Transprimer)

Chloramphenicol resistance (chloramphenicol transacetylase, *cat*, Cam): 700 bp
Note: We have observed an insertion in the penultimate codon of cat that retained drug resistance.

ColEI origin of replication: 600 bp

p15A origin of replication: 810 bp

References

1. Craig, N. L. (1996) *Curr. Top. Microbiol. Immunol.* 204, 27–48.
2. Stellwagen, A. E. and Craig, N. L. (1997) *Genetics* 145, 573–585.
3. Biery, M. C., Stewart, F. J., Stellwagen, A. E., Raleigh, E. A. and Craig, N. L. (2000) *Nucl. Acids Res.* 28, 1067–1077.
4. Stellwagen, A. E. and Craig, N. L. (1997) *EMBO J.* 16, 6823–6834.
5. Kolter, R., Inuzuka, M. and Helinski, D. R. (1978) *Cell* 15, 1199–1208.
6. Metcalf, W. W., Jiang, W. and Wanner, B. L. (1994) *Gene* 138, 1–7.
7. Hanahan, D. (1983) *J. Mol. Biol.* 166, 557.
8. Hanahan, D., Jessee, J. and Bloom, F. R. (1991) *Methods Enzymol.* 204, 63–113.
9. Bodenteich, A. and Chissoe, S. (1994) *Automated DNA Sequence and Analysis*. (Adams, M. D., Fields, C. and Venter, J. C., Eds) Academic Press, 42–50.

Kit Components Sold Separately:

pGPS1.1	
#N7111S	1 µg
pGPS2.1	
#N7121S	1 µg
TnsABC* Transposase	
#P7190S	10 µl (10 reactions)
Primer N	
#S1267S	0.5 A ₂₆₀ units
Primer S	
#S1266S	0.5 A ₂₆₀ units
10X GPS Buffer and Start Solution	
#B7100S	

Companion Products:

1 kb DNA Ladder	
#N3232S	100 µg
#N3232L	500 µg
I-Ceu I	
#R0699S	250 units
#R0699L	1,250 units
I-Sce I	
#R0694S	250 units
#R0694L	1,250 units
Not I	
#R0189S	500 units
#R0189L	2,500 units
PI-Sce I (VDE)	
#R0696S	100 units
#R0696L	500 units
Pme I	
#R0560S	500 units
#R0560L	2,500 units
Spe I	
#R0133S	500 units
#R0133L	2,500 units
Swa I	
#R0604S	2,000 units
#R0604L	10,000 units



New England Biolabs, Inc.

32 Tozer Road
Beverly, MA 01915-5599
USA
Telephone (978) 927-5054
Toll Free (USA Orders) 1-800-632-5227
Toll Free (USA Tech) 1-800-632-7799
Fax (978) 921-1350
e-mail: info@neb.com
www.neb.com

New England Biolabs, Ltd.

1815 Ironstone Manor, Unit 6
Pickering, Ontario, L1W 3W9
Canada
Telephone (905) 837-2234
Toll Free 1-800-387-1095
Fax (905) 837-2994
Fax Toll Free 1-800-563-3789
e-mail: info@ca.neb.com

New England Biolabs (Beijing), Ltd.

Wangzhuang Lu No. 1
Tsinghua Tongfang High-Tech Plaza
Building B, 6th Floor B
Haidian District, Beijing 100083
China, People's Republic
Telephone 010-82378266
Fax 010-82378262
e-mail: beijing@neb-china.com

New England Biolabs GmbH

Brüningstrasse 50, Geb.G 810
65926 Frankfurt am Main
Germany
Telephone +49/(0)69/305 23140
Free Call 0800/246 5227 (Germany)
Fax +49(0)69/305 23149
Free Fax 0800/246 5229 (Germany)
e-mail: info@de.neb.com

New England Biolabs (UK) Ltd.

73 Knowl Piece, Wilbury Way
Hitchin, Hertfordshire SG4 0TY
England, UK
Telephone (01462) 420616
Call Free 0800 318486
Fax (01462) 421057
Fax Free 0800 435682
e-mail: info@uk.neb.com



NEW ENGLAND
BioLabs[®]
Inc.

the leader in enzyme technology