

TRIAD: a transposition-based approach for gene mutagenesis by random short in-frame insertions and deletions for directed protein evolution

Stephane Emond

Department of Biochemistry, University of Cambridge <https://orcid.org/0000-0002-5160-0917>

Florian Hollfelder (✉ fh111@cam.ac.uk)

Department of Biochemistry, University of Cambridge <https://orcid.org/0000-0002-1367-6312>

Method Article

Keywords: directed evolution, protein engineering, mutagenesis, insertions and deletions, in vitro transposition

Posted Date: June 14th, 2021

DOI: <https://doi.org/10.21203/rs.3.pex-1448/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Insertions and deletions (InDels) are among the most frequent changes observed in natural protein evolution, yet their potential has hardly been harnessed in directed evolution experiments. Here we describe the standard protocol for TRIAD (Transposition-based Random Insertion And Deletion mutagenesis), a simple and efficient Mu transposon mutagenesis approach for generating libraries of single InDel variants with one, two or three triplet nucleotide insertions or deletions. This method has recently been employed in three published examples of InDel-based directed evolution of proteins, including a phosphotriesterase, a scFv antibody and an ancestral luciferase.

Introduction

Access to reliable methods for gene randomization is essential for success in directed evolution experiments. Popular and practically useful methods must meet several key requirements: a good library generation protocol should create a large number of variants, avoid bias in gene composition or type of variant introduced, and be technically straightforward. When it comes to amino acid substitutions, several approaches (e.g. error-prone PCR, site-saturation mutagenesis starting with synthetic oligonucleotides) have been developed that partially or fully meet these criteria. In contrast, the use of insertions and deletions (InDels) in directed evolution experiments has been curtailed by practical limitations in existing methodologies to randomly incorporate such changes within a sequence of interest. Therefore, while the effects of substitutions are extensively documented in the literature, the application of InDels in protein engineering has been sparse, with very few directed evolution campaigns on record that originate from such libraries. For example, the random insertion/deletion (RID) mutagenesis protocol^[1], the first attempt towards creating InDel libraries, relies on a complex protocol involving random cleavage of single stranded DNA, so that random substitutions are introduced unintentionally alongside the target mutations. Two other early methods, segmental mutagenesis^[2] and Random insertional-deletional strand exchange mutagenesis (RAISE^[3]), do not control for the length of the InDel and consequently produce libraries that primarily contain frameshifted variants. In contrast, a codon-based protocol dubbed codon-based random deletion (COBARDE^[4]) gives a pool of multiple codon-based deletions with <5% frameshifts but requires custom reprogramming of an oligonucleotide synthesizer to create mutagenic oligonucleotides. Alternatively, the viability of transposon-based protocols has been established for generating deletions of various sizes, up to gene truncation variants^[5]. However, the only reported such protocol to create insertions, namely pentapeptide scanning mutagenesis^[6], merely gains access to insertions of defined size and sequence.

Here we present a set of procedures (dubbed TRIAD: Transposition-based Random Insertion And Deletion mutagenesis)^[7] for random introduction of single short in-frame InDels of one, two or three nucleotide triplets ($\pm 3, 6$ or 9 bp) into a given DNA sequence. TRIAD consists of a single transposition reaction followed by successive cloning steps for the generation of deletions or insertions (Figure 1). TRIAD's first step is an in vitro Mu transposition reaction^[8] that ultimately determines the location of the forthcoming

single InDel event in each variant. The reaction is performed using engineered mini-Mu transposons, dubbed TransDel (Figure 1A) and TransIns (Figure 1B), that are inserted randomly within the target DNA sequence during the first step of TRIAD; resulting in the generation of transposon insertion libraries. The ends of TransDel and TransIns were designed to bring about deletion and insertion libraries, respectively. TransDel is functionally equivalent to the previously described MuDel transposon^[5] with recognition sites for the type IIS restriction enzyme MlyI at both ends. Taking advantage of the transposon insertion mechanism, the positioning of MlyI sites within TransDel enables the deletion of 3 bp at random positions within the target sequence upon MlyI digestion and self-ligation (Figure 2A), as previously described^[5]. This strategy was extended to the generation of longer contiguous deletions (i.e., -6 and -9 bp) with a second stage, involving the insertion and subsequent MlyI-mediated removal of custom-made cassettes (dubbed Del2 and Del3; Figure 1A and Figure 2). For the generation of insertions, a new transposon, TransIns, was designed as – in contrast to TransDel – an asymmetric transposon (Figure 1B and Figure 2B), bearing different end sequences (NotI on one end and MlyI on the other). The latter site marks subsequent insertion sites for the ligation of custom-made shuttle cassettes: Ins1, Ins2 and Ins3 carrying one, two and three randomized nucleotide triplets, respectively. Further digestion using a type IIS restriction enzyme (AclI) removes most of the shuttle sequence but leaves triplet insertions behind (Figure 1B and Figure 2B).

The use of TRIAD has, so far, been reported in three published examples of InDel-based directed evolution of proteins, including a phosphotriesterase^[7], a therapeutic scFv antibody^[9] and, most recently, an ancestral luciferase^[10].

Reagents

Enzymes

FastDigest restriction endonucleases (Thermo Fisher Scientific). Enzymes used to prepare or remove TRIAD transposons and cloning cassettes include: BglIII (FD0083), Eco57I (FD0344; isoschizomer of AclI), SclI (FD1374; isoschizomer of MlyI), NotI (FD0593), SmaI (FD0663). Enzymes used to insert the gene of interest (GOI) withing the TRIAD-bespoke target vector may include (non-exhaustive list): NdeI (FD0583), NcoI (FD0573), BamHI (FD0054), XhoI (FD0694), HindIII (FD0504).

MuA transposase (Thermo Fisher Scientific, F750).

T4 DNA ligase (Thermo Fisher Scientific, EL0011).

DNA Polymerase I, Large (Klenow) Fragment (New England Biolabs, M210S).

DNA extraction/purification kits

GeneJET Plasmid Miniprep kit (Thermo Fischer Scientific, K0503).

GeneJET Plasmid Midiprep kit (Thermo Fischer Scientific, K0481).

Zymoclean Gel DNA Recovery kit (Zymo Research, D4001/D4002).

DNA Clean & Concentrator-5 kit (Zymo Research, D4004) for DNA purification and concentration from reaction mixtures.

Bacterial culture reagents

E. coli 10G ELITE electrocompetent cells (Lucigen, 60051-1/60051-2/60052-1/60052-2)

Culture media: LB Broth Miller (Formedium, LMM0104); LB-Agar Miller (Formedium, LMM0204); Recovery Medium (Lucigen, 80026-1).

Antibiotics: ampicillin (Melford, A40040-25); kanamycin (Melford, K22000-5); chloramphenicol (Duchefa Biochemie, C0113).

Plasmids (NB: detailed construction are described in a previous reference^[7]):

Target vectors housing the gene of interest: pID-Tet (See Supplementary file 1 for full sequence); pID-T7 (See Supplementary file 2). See also troubleshooting note 2.

Vectors housing engineered transposons: pUC57-TransDel (See Supplementary file 3); pUC57-TransIns (See Supplementary file 4);

Vectors housing selection cassettes for the deletion of 6 bp (pUC57-Del2; see Supplementary file 5) and 9 bp (pUC57-Del3; see Supplementary file 6);

Vectors housing selection cassettes for the insertion of 3 bp (pUC57-Ins1; see Supplementary file 7), 6 bp (pUC57-Ins2; see Supplementary file 8) and 9 bp (pUC57-Ins3; see Supplementary file 9).

Equipment

Eppendorf ThermoMixer F1.5 (Eppendorf)

NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific)

Electroporator Eporator (Eppendorf)

Electroporation cuvettes 1 mm gap (Geneflow, E6-0050)

Petri dishes, 14 mm diameter (Thermo Fisher, 10051622)

Incubator (*e.g.*, Heratherm 104 L, Thermo Fischer Scientific) and shaker (*e.g.*, MaxQ6000, Thermo Fischer Scientific) for bacterial cell culture.

Procedure

Prior to applying this protocol, the gene of interest (GOI) needs to be inserted in a TRIAD-bespoke target vector such as pID-T7 or pID-Tet (See troubleshooting note 1).

1. Construction of transposon insertion libraries

This part describes how to generate transposon insertion libraries, *i.e.*, with transposon (either TransDel or TransIns) incorporated only in the GOI and not elsewhere in the target vector. Use TransDel to construct deletion libraries and TransIns for insertion libraries.

1.1 Prepare TransDel or TransIns transposon.

Isolate TransDel or TransIns from pUC57 by restriction digestion with BglII, electrophoresis and agarose gel purification. Perform separate reactions with either TransDel or TransIns.

Reaction (final volume: 20 µl):

pUC57-Transposon: 1 to 2 µg

FastDigest BglII: 1 µl

10x FastDigest Green buffer: 2 µl

Nuclease-free water to 20 µl

Incubate 30 minutes at 37 °C.

Separate the resulting DNA products by agarose gel electrophoresis and extract the fragment corresponding to the transposon (around 1 kb).

1.2 *In vitro* transposition reaction.

Transposons are incorporated at each possible insertion site within the target vector during this step.

Reaction (final volume: 20 µl):

Target vector (containing GOI): 300 ng

TransDel/TransIns (BglII-digested): 50 ng

5X Reaction Buffer: 4 µl

MuA Transposase: 1 µl

Nuclease-free water to 20 µl

Incubate 1 to 2 hours at 30 °C then heat-inactivate for 10 min at 75 °C.

Purify and concentrate DNA to 10 µl in nuclease-free water.

1.3 Transform purified transposition product (corresponding to 3 to 15 ng of plasmid DNA) into electrocompetent *E. coli*.

Mix up to 2.5 µl or up to 5 µl of purified transposition product with 25 µl or 50 µl of electrocompetent cells, respectively (see troubleshooting note 2). Set voltage to 1.8 kV when using 1 mm gap electroporation cuvettes (alternatively, use up to 2.5 kV when using 2 mm gap electroporation cuvettes). Add Recovery Medium (or, alternatively, use SOC medium) to a final volume of 300 µl and incubate 1 hour under agitation at 37 °C.

Plate up to 250 µl of this transformed culture mix on LB-Agar supplemented with 100 µg/ml ampicillin and 34 µg/ml chloramphenicol (LB-Agar Amp + Cam) and incubate overnight at 37 °C. Alternatively, inoculate 10 mL liquid LB Amp + Cam with the transformed culture mix and incubate under agitation at 37 °C overnight. To assess transposition efficiency, plate 0.1 to 1% of transformed cells on LB-Agar Amp + Cam and, after overnight incubation at 37 °C, calculate the number of transformants present in the total transformation mixture from the number of counted colonies.

1.4 Extract plasmid library pool from *E. coli* transformants.

After growth, collect transformed *E. coli* cells and extract plasmid libraries. Depending on the conditions used for growth (solid or liquid growth):

- If transformed cells were plated in a 14 cm Petri Dish, scrape colonies in 6 mL LB, transfer into three 2-ml Eppendorf tubes, pellet cells by centrifugation and extract plasmid pool from one tube (store the two others as back-ups) using a Midiprep kit.
- If transformed cells were grown in a 10 mL liquid culture, collect cell pellet by centrifugation and extract plasmid directly from the collected pellet using a Midiprep kit.

This step ultimately yields an intermediate plasmid library pool in which the transposons (either TransDel or TransIns) are incorporated at any viable insertion site in the target vector (*i.e.*, the ampicillin resistance gene and the origin of replication are likely to be free of transposon at this stage).

1.5 Extract GOI containing TransDel or TransIns from transposon plasmid library.

Digest transposon plasmid library obtained during previous step by restriction endonucleases (RE) cutting at the extremities of the GOI (FastDigest RE1 and RE2).

Reaction (final volume: 20 µl):

Transposon plasmid library: 1 to 2 µg

FastDigest RE1 and RE2: 1 µl each

10x FastDigest Green buffer: 2 µl

Nuclease-free water to 20 µl

Incubate restriction digestion for 30 minutes at 37 °C.

Separate the resulting fragments by agarose gel electrophoresis. The reaction will generate four products, corresponding to (in decreasing size): (i) target vector backbone with transposon, (ii) target vector without transposon, (iii) GOI with transposon and (iv) GOI without transposon. Isolate band corresponding to fragment (iii) and purify the DNA from the agarose gel. The target vector (band ii) may also be isolated and purified at this stage (alternatively, the digested vector can be prepared extemporaneously).

1.6 Ligate the GOI containing the transposon in the target vector.

Perform cohesive end ligation based on ratio 1 vs. 3 to 5 (vector/insert) in 20 µl reactions with T4 DNA ligase. Digested vector amount should be between 10 and 50 ng (calculate insert quantity accordingly).

Reaction (final volume: 20 µl):

Target vector: 10 to 50 ng

GOI with transposon: calculate quantity depending on fragment size

T4 DNA ligase: 5 Weiss units (1 µl)

10X T4 DNA Ligase buffer: 2 µl

Nuclease-free water to 20 µl

Incubate 30 minutes at room temperature or overnight at 16 °C. Purify and concentrate ligation product in 10 µl nuclease-free water.

1.7 Transform purified ligation mix (corresponding to 2.5 to 25 ng of plasmid DNA) into electrocompetent *E. coli* cells.

See step 1.3 for transformation and incubation conditions.

To estimate the transposon insertion library size, plate 0.1 to 0.5% of transformed cells on LB-Agar Amp + Cam and extrapolate the number of transformants present in the total transformation mixture from the number of resulting colonies.

1.8 After growth, collect cells and extract plasmid pool corresponding to transposon insertion library.

Follow procedure described in step 1.4.

2. Construction of deletion libraries from TransDel insertion libraries

2.1 Construction of single nucleotide triplet deletion libraries (as outlined by Jones^[5])

2.1.1 Remove the TransDel transposon by digesting the TransDel insertion library plasmid pool with Schl.

This step yields a linearised single nucleotide triplet deletion library (*i.e.*, linear target vector containing the GOI minus 3 bp). The resulting linear plasmid library can be used to generate either single nucleotide triplet deletion library (go to step 2.1.2) or double and triple nucleotide triplet deletion libraries (go to step 2.2.2).

Reaction (final volume: 20 µl):

TransDel insertion library plasmid pool: 1 to 2 µg

FastDigest Schl: 1 µl

10x FastDigest Green buffer: 2 µl

Nuclease-free water to 20 µl

Incubate 30 minutes at 37 °C then heat-inactivate for 5 minutes at 80 °C.

Separate the resulting fragments by agarose gel electrophoresis and extract the digestion product corresponding to the target vector containing the GOI.

2.1.2 Re-circularise the linear single nucleotide triplet deletion plasmid by intramolecular ligation.

For optimal intramolecular ligation, DNA concentration should be below 1 ng/µl.

Reaction (final volume: 50 µl):

Linear DNA (vector+target gene): 10-50 ng

T4 DNA Ligase: 5 Weiss units (1 µl)

10X T4 DNA Ligase buffer: 5 µl

Nuclease-free water to 50 µl

Incubate 30 minutes at room temperature. Purify and concentrate ligation product in 10 μ l nuclease-free water.

2.1.3 Transform purified ligation mix (corresponding to 2.5 to 25 ng of plasmid DNA) into electrocompetent *E. coli* cells.

See step 1.3 for transformation conditions.

Plate up to 250 μ l of this transformed culture mix on LB-Agar Amp (100 μ g/ml) and incubate overnight at 37 °C. Alternatively, inoculate 10 mL liquid LB Amp with the transformed culture mix and incubate overnight under agitation at 37 °C. To estimate library size, plate 0.1 to 0.5% of transformed cells on LB-Agar Amp and extrapolate the number of transformants present in the total transformation mixture from the number of resulting colonies.

2.1.4 After growth, collect cells and extract plasmid library pool corresponding to single nucleotide triplet deletion library.

Follow procedure described in step 1.4. After this step, the resulting single nucleotide triplet library can be stored as a DNA solution.

2.2 Construction of double and triple nucleotide triplet deletion libraries

2.2.1 Prepare Del 2 or Del3 selection cassette by digestion of pUC57-Del2 (or Del3) with SmaI.

Reaction (final volume: 20 μ l):

pUC57-Del2 (or Del3): 1 to 2 μ g

Fast-digest SmaI: 1 μ l

10x FastDigest Green buffer: 2 μ l

Nuclease-free water to 20 μ l

Incubate 30 minutes at 37 °C

Separate the resulting fragments by agarose gel electrophoresis and extract the digestion product corresponding to the selection cassette (around 1 kb).

2.2.2 Ligate the selection cassette (Del2 or Del3) in the linear single nucleotide triplet deletion plasmid library obtained at step 2.1.1

Perform blunt-end ligation based on ratio 1 vs. 5 to 10 (vector/insert) in 20 µl reactions with T4 DNA ligase.

Reaction (final volume: 20 µl):

Single triplet deletion plasmid library: 10 to 50 ng

Selection cassette (Del2/Del3): calculate quantity based on the size of the vector + GOI fragment

T4 DNA ligase: 5 Weiss units (1 µl)

10X T4 DNA Ligase buffer: 2 µl

Nuclease-free water to 20 µl

Incubate 30 minutes at room temperature or overnight at 16 °C. Purify and concentrate ligation product in 10 µl nuclease-free water.

2.2.3 Transform purified ligation mix (corresponding to 2.5 to 25 ng of plasmid DNA) into electrocompetent *E. coli* cells.

See step 1.3 for transformation and incubation conditions. Use culture media (LB-Agar or liquid LB) supplemented with 50 µg/ml kanamycin (Kan) at this step.

2.2.4 After growth, collect cells and extract plasmid library pool corresponding to Del2 or Del3 insertion library.

Follow procedure described in step 1.4.

2.2.5 Remove Del2 (or Del3) selection cassette by digesting the cassette insertion library plasmid pool with Schl.

This step yields a linearised double or triple nucleotide triplet deletion plasmid library from Del2 or Del3 insertion plasmid library, respectively. Follow procedure outlined at step 2.1.1.

2.2.6 Re-circularise the double or triple nucleotide triplet plasmid library by intramolecular ligation. Follow procedure outlined at step 2.1.2.

2.2.7 Transform purified ligation mix (corresponding to 2.5 to 25 ng of plasmid DNA) into electrocompetent *E. coli* cells.

Follow procedure outlined at step 2.1.3.

2.2.8 After growth, collect cells and extract plasmid library pool corresponding to double or triple nucleotide triplet deletion library.

Follow procedure described in step 1.4. After this step, the resulting libraries can be stored in the form of DNA solutions.

3. Construction of insertion libraries from TransIns insertion libraries

3.1 Prepare InsX (X = 1, 2 or 3) selection cassette by double-digestion of pUC57-InsX with NotI and Schl.

Reaction (final volume: 20 µl):

pUC57-InsX: 1 to 2 µg

FastDigest NotI: 1 µl

FastDigest Schl: 1 µl

10x FastDigest Green buffer: 2 μ l

Nuclease-free water to 20 μ l

Incubate 30 minutes at 37 °C then heat-inactivate for 5 minutes at 80 °C.

This double digestion will yield 5 fragments (in descending size order: 1408 bp, 1152 bp (Ins1) / 1155 bp (Ins2) / 1158 bp (Ins3), 502 bp, 486 bp and 264 bp). Separate the fragments by agarose gel electrophoresis and extract the digestion product corresponding to the selection cassette (around 1.15 kb, second fragment in descending size order).

3.2 Remove the TransIns transposon by sequential double-digestion of pUC57-InsX with NotI and SclI.

Follow procedure outlined at step 3.1. This step yields a linearised plasmid library with cuts randomly distributed within the GOI and allowing the subsequent introduction of nucleotide triplet insertions.

Separate the resulting fragments by agarose gel electrophoresis and extract the digestion product corresponding to the target vector containing the GOI.

3.3 Ligate the InsX selection cassette in the linearised plasmid library obtained at step 3.2.

Follow procedure outlined at step 2.2.2.

3.4 Transform purified ligation mix (corresponding to 2.5 to 25 ng of plasmid DNA) into electrocompetent *E. coli* cells.

See step 2.2.3 for transformation and incubation conditions.

3.5 After growth, collect cells and extract plasmid library pool corresponding to InsX insertion library.

Follow procedure described in step 1.4.

3.6 Remove InsX selection cassette by digesting the cassette insertion library plasmid pool with Eco57I.

This step yields a linearised single, double or triple nucleotide triplet insertion plasmid library from Ins1, Ins2 or Ins3 insertion plasmid library, respectively. Digestion with Eco57I removes InsX leaving 2-base 3'-

overhangs to be removed in the next step.

Reaction (final volume: 20 µl):

InsX insertion library plasmid pool: >2 µg

Fastdigest Eco57I: 1 µl

SAM (20X): 1 µl

10x FastDigest Green buffer: 2 µl

Nuclease-free water to 20 µl

Incubate 30 minutes at 37 °C then heat-inactivate for 5 minutes at 80 °C.

Separate the resulting fragments by agarose gel electrophoresis and extract the digestion product corresponding to the target vector containing the GOI.

3.7 Removal of 3'-overhangs left by Eco57I using DNA Polymerase I, Large (Klenow) Fragment.

This step yields a linear plasmid library with random 1, 2 or 3 nucleotide triplet insertions within the GOI.

Perform reaction in NEB buffer 2 supplemented with 33 µM of each dNTP (132 µM total dNTPs). Adjust ratio enzyme/DNA accordingly by using 1 unit of enzyme per µg of DNA to be treated. Incubate reaction for 15 min at 25 °C. Stop digestion by adding EDTA (around 10 mM final) and heating for 20 min at 80 °C. Purify and concentrate resulting DNA product in 10 µl Nuclease-free water.

2.2.6 Re-circularise the insertion plasmid library by intramolecular ligation.

Follow procedure outlined at step 2.1.2.

2.2.7 Transform purified ligation mix (corresponding to 2.5 to 25 ng of plasmid DNA) into electrocompetent *E. coli* cells.

Follow procedure outlined at step 2.1.3.

2.2.8 After growth, collect cells and extract plasmid library pool corresponding to single, double or triple nucleotide triplet insertion library.

Follow procedure described in step 1.4. After this step, the resulting libraries can be stored in the form of DNA solutions.

Troubleshooting

Note 1: To enable TRIAD, any recognition sequences for MlyI (SclI), NotI and AclI (Eco57I) in the GOI must be removed either by site-directed mutagenesis (*e.g.*, by introducing a silent mutation that removes an eventual recognition sequences) or by designing and ordering a specific synthetic gene. Bespoke cloning vectors (*i.e.*, pID-Tet and pID-T7; supplementary files 1 and 2) have previously been assembled to host the GOI during the TRIAD procedure^[7] and are available from the authors.

Note 2: In our hands, 1 µl of purified and concentrated transposition reaction mix usually yields 10,000 to 50,000 CFUs using 25 µl of commercial electrocompetent cells. The use of high transformation efficiency cells ($>10^9$ CFUs/µg) is recommended to generate highly diverse libraries at each relevant step of the TRIAD procedure.

Time Taken

TRIAD should span just over 5 days when performing ligations and transformations on the same day (see the timeline in Figure 1).

Anticipated Results

References

1. H. Murakami, T. Hohsaka, M. Sisido, *Nat Biotechnol* **2002**, *20*, 76-81.
2. M. G. Pikkemaat, D. B. Janssen, *Nucleic Acids Res* **2002**, *30*, E35-35.
3. R. Fujii, M. Kitaoka, K. Hayashi, *Nucleic Acids Res* **2006**, *34*, e30.
4. J. Osuna, J. Yanez, X. Soberon, P. Gaytan, *Nucleic Acids Res* **2004**, *32*, e136.
5. D. D. Jones, *Nucleic Acids Res* **2005**, *33*, e80.
6. B. Hallet, D. J. Sherratt, F. Hayes, *Nucleic Acids Res* **1997**, *25*, 1866-1867.

7. S. Emond, M. Petek, E. J. Kay, B. Heames, S. R. A. Devenish, N. Tokuriki, F. Hollfelder, *Nat Commun* **2020**, *11*, 3469.
8. S. Haapa, S. Taira, E. Heikkinen, H. Savilahti, *Nucleic Acids Res* **1999**, *27*, 2777-2784.
9. K. Skamaki, S. Emond, M. Chodorge, J. Andrews, D. G. Rees; D. Cannon; B. Popovic, A. Buchanan, R. R. Minter, F. Hollfelder, *Proc Natl Acad Sci U S A* **2020**, *117* (44), 27307-27318.
10. A. Schenkmyerova, G. P. Pinto, M. Toul, M. Marek, L. Hernychova, J. Planas-Iglesias, V. Liskova, D. Pluskal, M. Vasina, S. Emond, M. Dörr, R. Chaloupkova, D. Bednar, Z. Prokop, F. Hollfelder, U. T. Bornscheuer & J. Damborsky, *Nature Communication* **2021** (*in press*, DOI: 10.1038/s41467-021-23450-z).

Figures

Figure 1

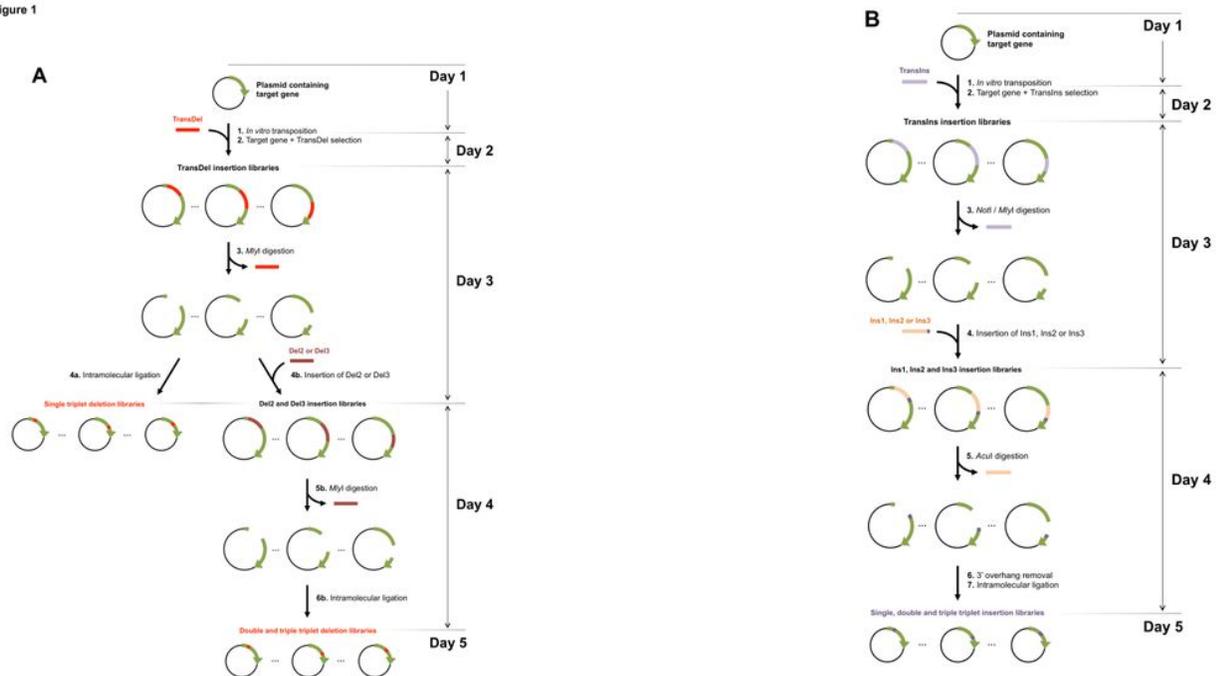


Figure 1

Schematic outline and timeline of the procedure for the generation of random InDel libraries A. Generation of deletion libraries. Step 1: The TransDel insertion library is generated by in vitro transposition of the engineered transposon TransDel into the plasmid containing the target gene followed

by the subcloning of the fragment comprising the target gene and the transposon into a fresh plasmid. Step 2: MlyI digestion removes TransDel together with 3 bp of the original target gene and generates a single break per target gene variant. Step 3a: Intramolecular ligation results in the reformation of the target gene minus 3 bp, yielding a library of single variants with a deletion of 1 triplet. Step 3b: DNA cassettes dubbed Del2 and Del3 are then inserted between the break in the target gene to generate Del2 and Del3 insertion libraries. Step 4b: MlyI digestion removes Del2 and Del3 together with 3 and 6 additional bp of the original GOI, respectively. Step 5b: Intramolecular ligation results in the reformation of the target gene minus 6 and 9 bp, yielding libraries of single variants with a deletion of 2 and 3 triplets, respectively. Red vertical lines indicate deletions.

B. Generation of insertion libraries. Step 1: The TransIns insertion library is generated by in vitro transposition of the engineered transposon into the target gene. Step 2: digestion by NotI and MlyI removes TransIns. Step 3: DNA cassettes dubbed Ins1, Ins2 and Ins3 (with respectively 1, 2 and 3 randomized bp triplets at one of their extremities; indicated in blue) are then inserted between the break in the target gene to generate the corresponding Ins1, Ins2 and Ins3 insertion libraries. Step 4: AclI digestion and 5' end digestion by the Klenow fragment remove the cassettes, leaving the randomized triplet(s) in the original target gene. Step 5: Intramolecular ligation results in the reformation of the target gene plus 3, 6 and 9 random bp, yielding libraries of single variants with an insertion of 1, 2 and 3 triplets, respectively. Purple vertical lines indicate insertions.

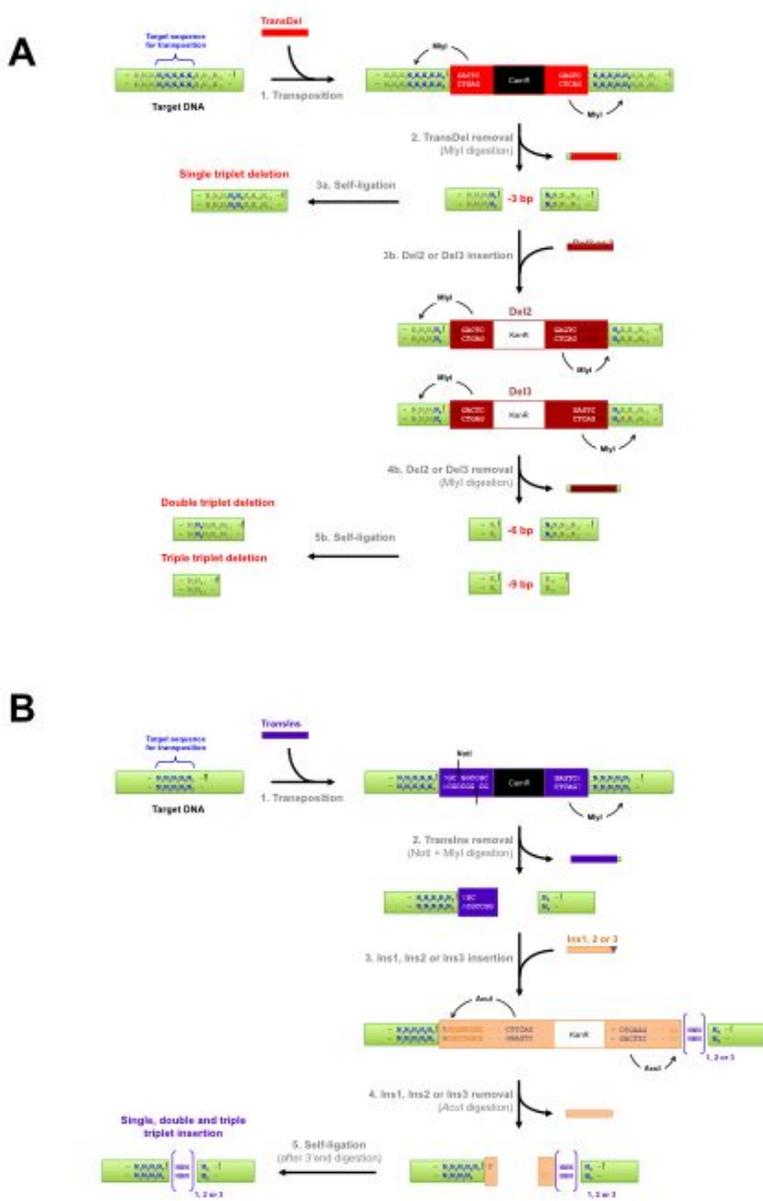


Figure 2

Mechanism for the generation of InDels using TRIAD. The target sequence is located on a plasmid with ampicillin resistance (*bla*) and neither the target sequence nor the plasmid contain any *MylI*, *NotI* or *AclI* restriction sites. (A) Generation of single, double and triple triplet nucleotide deletions. Step 1. Two *MylI* recognition sites (5'GAGTC(N)5') are positioned at each end of *TransDel*, 1 bp away from the site of transposon insertion. Transposition with *TransDel* results in the duplication of 5 bp (N4N5N6N7N8) of the

target DNA at the insertion point. TransDel carries a selection marker (resistance gene against chloramphenicol; CamR) enabling the recovery of in vitro transposition products after transformation into *E. coli*. Step 2. MlyI digestion removes TransDel together with 8 bp of the target DNA (4 bp at each end), leaving blunt ends and resulting in the removal of a contiguous 3 bp sequence from the target DNA (N5N6N7). Step 3a. Self-ligation reforms the target DNA minus 3 bp, as previously described 14. Step 3b. Alternatively, blunt-ended cassettes Del2 or Del3 are ligated into the gap left upon TransDel removal for the generation of 6 and 9 bp deletions, respectively. Both Del2 and Del3 also contain two MlyI recognition sites advantageously positioned towards the ends of the cassettes. These cassettes also contain a different marker than TransDel (resistance gene against kanamycin; KanR) to avoid cross-contamination. Step 4b. MlyI digestion removes Del2 and Del3 together with respectively 3 and 6 additional bp of the original target DNA. In the case of Del2, MlyI digestion results in the removal of a 3 bp sequence (N2N3N4) on one side of the cassette. In the case of Del3, MlyI digestion results in the removal of two 3 bp sequence (N2N3N4) on both side of the cassette (N2N3N4 and N8N9N10). Step 5b. Self-ligation reforms the target DNA minus 6 or 9 bp. (B) Generation of single, double and triple randomized triplet nucleotide insertions. Step 1. TransIns is an asymmetric transposon with MlyI at one end and NotI at the other end. Both recognition sites are positioned 1bp away from TransIns insertion site. Upon transposition, 5 bp (N1N2N3N4N5) of the target DNA are duplicated at the insertion point of TransIns. Step 2. Double digestion with NotI and MlyI results in the removal of TransIns. Digestion with MlyI removes TransIns with 4 bp (N1N2N3N4) of the duplicated sequence at the transposon insertion site. Digestion with NotI leaves a 5', 4-base cohesive overhang. Step 3. DNA cassettes Ins1, Ins2 and Ins3 (Ins1/2/3) carrying complementary ends are ligated in the NotI/MlyI digested TransIns insertion site. Ins1, Ins2 and Ins3 carry respectively 1, 2 and 3 randomized bp triplets at their blunt-ended extremities ([NNN]_{1,2 or 3}; indicated in purple). Ins1/2/3 contain two AclI recognition sites (5'CTGAAG(16/14)) strategically positioned towards their ends. One site is located so that AclI will cleave at the point where the target DNA joins Ins1/2/3. The other site is positioned so that AclI will cut inside Ins1/2/3 to leave the randomized triplet(s) with the target DNA. Step 4. Digestion with AclI removes Ins1/2/3 leaving 3', 2-base overhangs with the target DNA (i.e., 5'N5T on one end and 5'TC on the end carrying the randomized triplet(s)). Digestion with the Large Klenow fragment generates blunt ends by removing the overhangs. This step also enables to discard the extra nucleotide (N5) from the sequence duplicated during the transposition. Step 5. Self-ligation reforms the target DNA with one, two or three randomized nucleotide triplets.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [pIDTet.gb](#)
- [pIDT7.gb](#)
- [pUC57TransDel.gb](#)
- [pUC57TransIns.gb](#)

- pUC57Del2.gb
- pUC57Del3.gb
- pUC57Ins1.gb
- pUC57Ins2.gb
- pUC57Ins3.gb