Chapter 16

ITCHY: Incremental Truncation for the Creation of Hybrid Enzymes

Wayne M. Patrick and Monica L. Gerth

Abstract

Incremental Truncation for the Creation of Hybrid enzYmes (ITCHY) is a directed evolution technique for randomly recombining two genes. The chief advantage of ITCHY is that there is no requirement for the two genes to share any sequence similarity. This distinguishes ITCHY from directed evolution methods that are based on homologous recombination, such as DNA shuffling. In ITCHY, *Escherichia coli* exonuclease III is used to incrementally truncate one of the parental genes from its 3′ end and the other from its 5′ end. Ligation of the randomly truncated gene fragments yields a combinatorial library of chimeras. In this chapter, we provide detailed protocols for constructing libraries using both the user-friendly thio-ITCHY method and also time-dependent incremental truncation. We illustrate the protocols with the data that we obtained when we recombined two alcohol dehydrogenase genes that only share 47 % sequence identity.

Key words Incremental truncation, Exonuclease III, Nonhomologous recombination, Protein engineering, Thio-ITCHY, Time-dependent truncation

1 Introduction

In vitro recombination techniques have yielded many success stories in the field of directed evolution [1]. Most methods for in vitro recombination either rely on a high level of sequence identity between the genes being recombined, or they are underpinned by computational predictions of favored crossover locations; these methods are described elsewhere in this volume.

In contrast, Incremental Truncation for the Creation of Hybrid enzYmes (ITCHY) is a method for randomly recombining any two genes, with no requirement that they share sequence similarity [2]. ITCHY has provided a means to utilize the power of recombination for sampling sequence space and finding unexpected solutions, in cases where the two parental proteins are structural homologues but their genes are <70 % identical. For example, ITCHY was used to recombine two homologous deoxynucleoside

Elizabeth M.J. Gillam et al. (eds.), *Directed Evolution Library Creation: Methods and Protocols*, Methods in Molecular Biology, vol. 1179, DOI 10.1007/978-1-4939-1053-3_16,

kinases with 48 % sequence identity, yielding a chimeric enzyme that could phosphorylate the nucleoside analogue prodrug, d4T—an activity that neither parent possessed [3]. ITCHY is also a tractable method for exploring the role of nonhomologous end joining in the (natural) evolution of protein folds and functions [4].

The basic ITCHY protocol involves incrementally truncating one parental gene from its 3′ end, creating a set of sequences in which every single-base truncation is represented. Simultaneously, the second parental gene is incrementally truncated from its 5′ end. In practice, the two parental genes are cloned into a single plasmid, in tandem. The plasmid is linearized with a restriction enzyme that cleaves between the two genes, and then the sets of incrementally truncated sequences are generated by digestion with *Escherichia coli* exonuclease III. In the original ITCHY method [2], time-dependent exonuclease III digestion yielded incrementally truncated genes. An important advance was the development of thio-ITCHY [5]. In this user-friendly method, the linearized plasmid is amplified in a PCR that uses a mixture of dNTPs and α-phosphorothioate dNTPs (αS-dNTPs). The phosphorothioate-containing linkages that are incorporated randomly into the PCR product are resistant to cleavage by exonuclease III, thus enabling sets of incrementally truncated genes to be generated. In each protocol, exonuclease III digestion is followed by the preparation of blunt ends and intramolecular ligation to recircularize the plasmid, thereby fusing together randomly sized fragments of the two parental genes. An overview of the steps in the thio-ITCHY protocol, using an example from our laboratory, is shown in Fig. 1.

A number of variations and alterations to the basic ITCHY and thio-ITCHY methodologies have also been implemented. Hart and coworkers successfully incorporated the time-dependent truncation protocol into their method for finding folded fragments of recalcitrant proteins, ESPRIT (Expression of Soluble Proteins by Random Incremental Truncation) [6]. Rather than cloning the two parental genes into the same plasmid, one of us (MLG) explored the possibility of beginning with each gene residing on a different plasmid (albeit with identical backbones). The full-length thio-ITCHY template (with both parental genes on the same, linear molecule) could then be assembled in an overlap extension PCR [3]. However, this protocol is difficult to optimize, and subsequently we have returned to constructing two-gene templates such as the one shown in Fig. 1. Finally, the pINSALECT vector was designed to remove out-of-frame chimeras from ITCHY libraries [7]. Two-thirds of the variants in an ITCHY library will contain a frameshift at the crossover point. Removing frameshifted variants is particularly important if the ITCHY library is to be used as a starting point for further diversification by DNA shuffling (a method named SCRATCHY [8]) or in ORF-selector ESPRIT, where a single gene is truncated from both termini [9]. The pINSALECT vector contains the gene for the *Saccharomyces cerevisiae* VMA split intein, sandwiched between the Tat signal sequence and

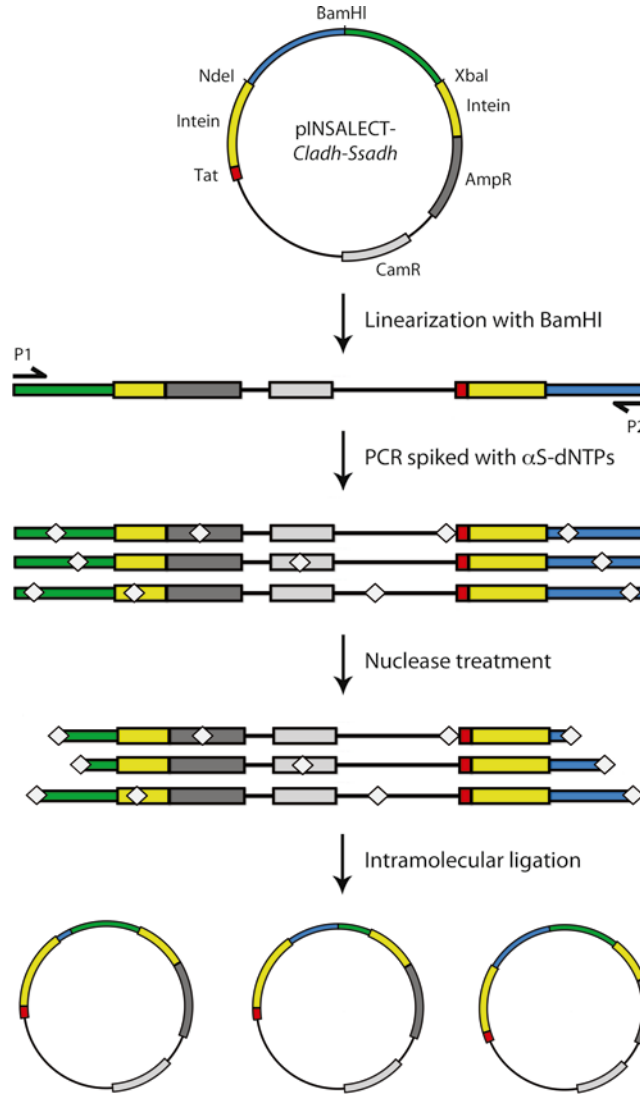

Fig. 1 Overview of the thio-ITCHY protocol. In this example, the genes for the *C. ljungdahlii* alcohol dehydrogenase (*Cladh*, *blue*) and the *S. solfataricus* alcohol dehydrogenase (*Ssadh*, *green*) were cloned in tandem into the pINSALECT backbone, using the NdeI, BamHI, and XbaI sites indicated. The *adh* genes form part of a continuous open reading frame that begins with the Tat signal sequence (*red*), includes the VMA split intein (*yellow*), and ends with the gene for β-lactamase (AmpR). The AmpR marker is used for reading frame selection. The plasmid also contains a chloramphenicol resistance marker (CamR) for maintenance during library construction and sequence analysis. After the plasmid is linearized, it is amplified using the primers indicated (P1 and P2), with a mixture of dNTPs and αS-dNTPs. The (random) sites of αS-dNTP incorporation are shown as *white diamonds*

the TEM-1 β-lactamase gene (Fig. 1). The two genes to be recombined are cloned into a loop of the intein homing endonuclease domain. Only in-frame ITCHY variants will give rise to full-length polyproteins, in which the intein ligates Tat to β-lactamase and the host cell displays an ampicillin resistance phenotype.

Here, we present our protocols for thio-ITCHY and the time-dependent truncation method. We have illustrated the protocols using data from a thio-ITCHY library that we have recently constructed. In this example, we recombined genes for the alcohol dehydrogenases from *Clostridium ljungdahlii* (GenBank accession CP001666; locus tag CLJU_c24860) and *Sulfolobus solfataricus* (GenBank accession AJ010590). These genes share 47 % sequence identity; therefore, it would have been impossible to recombine them using DNA shuffling. As recommended above, our starting point was the two genes cloned in tandem, in the pINSALECT backbone (Fig. 1).

2 Materials

2.1 Thio-ITCHY

1. Restriction enzyme that cleaves between the two parental genes in the template plasmid, e.g., BamHI-HF from New England Biolabs (Ipswich, MA, USA) (Fig. 1).
2. E.Z.N.A. Cycle Pure Kit (Omega Bio-Tek, Norcross, GA, USA) for purifying the products of PCRs and digestion reactions. Equivalent kits from other suppliers are also suitable. We recommend preheating the elution buffer to 60 °C before use, to increase the DNA yield from each spin column.
3. Spectrophotometer and cuvettes for measuring DNA concentration, e.g., a Biophotometer and UVettes from Eppendorf (Hamburg, Germany).
4. Oligonucleotide primers that are suitable for the thio-ITCHY PCR (*see* **Note 1**).
5. Deoxyribonucleotide triphosphates (dNTPs): stock solution containing each dNTP at 5 mM. Store aliquots of this stock solution at −20 °C.
6. α-Phosphorothioate dNTPs (αS-dNTPs): stock solution containing each αS-dNTP at 1 mM. Store aliquots of this stock solution at −20 °C. High-quality αS-dNTPs can be purchased from Glen Research Corporation (Sterling, VA, USA).
7. Taq DNA polymerase and its reaction buffer, e.g., Kapa Taq with 10× Kapa Taq Buffer A (Kapa Biosystems, Woburn, MA, USA) (*see* **Note 2**).
8. Thermocycler with a heated lid (e.g., an MJ Mini from Bio-Rad (Hercules, CA, USA)).
9. An agarose gel (0.8 % or 1.0 % w/v agarose), stained with ethidium bromide at 0.5 μg/mL.
10. A suitable DNA ladder and apparatus for agarose gel electrophoresis.

11. *E. coli* exonuclease III (100 U/μL) and its accompanying reaction buffer, NEBuffer 1 (New England Biolabs, Ipswich, MA, USA).
12. Mung bean nuclease (10 U/μL) and 10× mung bean nuclease reaction buffer (New England Biolabs, Ipswich, MA, USA).
13. DNA polymerase I, Klenow fragment (5 U/μL) and its accompanying reaction buffer, NEBuffer 2 (New England Biolabs, Ipswich, MA, USA).
14. Ethylenediaminetetraacetic acid (EDTA) stock solution: 0.5 M, pH 8.0.

2.2 Size Selection and Library Construction

1. An agarose gel (0.8 % w/v agarose), stained with 1× SYBR Safe DNA gel stain (Invitrogen, Carlsbad, CA, USA).
2. A size standard, constructed by linearizing a plasmid that has the same backbone as the ITCHY template plasmid, but only one of the two parental gene inserts (*see* **Note 3**).
3. Gel loading dye, a suitable DNA ladder, and apparatus for agarose gel electrophoresis.
4. Safe Imager 2.0 Blue-Light Transilluminator (Invitrogen, Carlsbad, CA, USA).
5. Clean razor blade for excising bands from a gel.
6. GenElute Gel Extraction Kit (Sigma-Aldrich, St. Louis, MO, USA). Equivalent kits from other manufacturers are also suitable.
7. T4 DNA ligase (400 cohesive end units per microliter) and 10× T4 DNA ligase reaction buffer (New England Biolabs, Ipswich, MA, USA). Other T4 DNA ligases are also suitable (*see* **Note 4**).
8. E.Z.N.A. Cycle Pure Kit (Omega Bio-Tek, Norcross, GA, USA), or equivalent, with elution buffer pre-warmed to 60 °C.
9. Aliquots (50 μL) of electrocompetent *E. coli* cells (*see* **Note 5**).
10. Gene Pulser electroporation cuvettes (0.2 cm electrode gap, Bio-Rad, Hercules, CA, USA).
11. MicroPulser electroporator (Bio-Rad, Hercules, CA, USA).
12. Sterile SOC medium: 20 g/L tryptone; 5 g/L yeast extract; 10 mM NaCl; 2.5 mM KCl; 20 mM glucose.
13. LB-agar plates containing the correct antibiotic for selecting plasmid-containing cells.

2.3 Sequence Analysis and Scale-Up

1. Thermocycler with a heated lid.
2. Plasmid backbone-specific primers for amplifying the recombined ITCHY inserts.

3. Reagents for a standard PCR screen. While there are many alternate suppliers, we routinely use the 2× Kapa Taq ReadyMix with dye (Kapa Biosystems, Woburn, MA, USA). This contains reaction buffer, dNTPs, and Taq polymerase.
4. Ethidium bromide-stained agarose gels and electrophoresis equipment.
5. E.Z.N.A. Cycle Pure Kit (Omega Bio-Tek, Norcross, GA, USA), or equivalent, with elution buffer pre-warmed to 60 °C.
6. Standard LB-agar plates (circular, 85–90 mm diameter) containing the appropriate antibiotic for maintaining the ITCHY plasmid.
7. Two square bioassay dishes (245 mm × 245 mm) from Corning (Tewksbury, MA, USA) or Nunc (Thermo Fisher Scientific, Waltham, MA, USA). Each bioassay dish holds 200 mL of LB-agar, supplemented with the appropriate antibiotic for selecting plasmid-containing cells.
8. A fresh batch of electrocompetent *E. coli* cells (*see* **Note 5**).
9. Electroporation cuvettes, an electroporator, and SOC medium, as described in Subheading 2.2.
10. LB medium, containing the appropriate antibiotic (20 mL).
11. Sterile glycerol (50 % v/v).
12. Cryogenic vials, suitable for storage at −80 °C.

2.4 ITCHY by Time-Dependent Truncation

1. Materials for ITCHY, as listed in Subheading 2.1: **items 1–3** and **8–14**.
2. NaCl stock solution: 120 mM, sterilized by autoclaving.
3. A laboratory timer (or a smart phone; *see* **Note 6**).

3 Methods

3.1 Thio-ITCHY

1. Construct a plasmid containing the two genes that are to be recombined, separated by a sequence that contains a unique restriction site (as shown in Fig. 1).
2. Purify the plasmid DNA and linearize ~3 μg of it by digestion with the restriction enzyme that cleaves between the two parental genes. We typically use BamHI-HF for this purpose (Fig. 1), though other enzymes are also suitable (*see* **Note 7**). After digestion to completion (e.g., 4 h incubation at 37 °C, with 20 U enzyme), purify the DNA with the Cycle Pure Kit. Elute the DNA from the spin column in 50 μL of elution buffer and determine its concentration spectrophotometrically. Use water to dilute an aliquot to 10 ng/μL.
3. In thin-walled 0.2 mL tubes, prepare three PCRs with different ratios of dNTPs to αS-dNTPs, as shown in Table 1 (*see* **Note 8**).

Table 1
Components of thio-ITCHY PCRs

Reagent	Control	1:7 ratio	1:9 ratio
10× polymerase reaction buffer	5.0 µL	10.0 µL	10.0 µL
dNTP mix (from 5 mM stock)	2.0 µL	3.5 µL	3.6 µL
αS-dNTP mix (from 1 mM stock)	0.0 µL	2.5 µL	2.0 µL
Forward primer (from 10 µM stock)	2.0 µL	4.0 µL	4.0 µL
Reverse primer (from 10 µM stock)	2.0 µL	4.0 µL	4.0 µL
Linearized plasmid DNA (from 10 ng/µL stock)	1.0 µL	2.0 µL	2.0 µL
Taq DNA polymerase (5 U/µL)	0.2 µL	0.4 µL	0.4 µL
Water	37.8 µL	73.6 µL	74.0 µL

Table 2
Recommended PCR cycling conditions for amplifying long templates

Step 1	2 min	94 °C
Step 2	10 s	94 °C
Step 3	20 s	Annealing temperature for primers (*see* **Note 1**)
Step 4	1 min/kb	68 °C
Step 5	Repeat steps 2–4 for an additional 24 cycles	
Step 6	5 min	68 °C
Step 7	Hold	4 °C (for product storage, if necessary)

4. Mix the reactions well, and then split each of the two 100 µL reactions into 2×50 µL aliquots. Place all five tubes in the thermocycler.
5. Run a PCR program that is suitable for amplifying long templates (*see* **Note 9**). In our experience, the program listed in Table 2 works well.
6. Pool the 2×50 µL aliquots of each αS-dNTP-containing reaction back together.
7. Run 3 µL of each PCR product on an ethidium bromide-stained agarose gel, alongside a DNA ladder. All three reactions ("control," "1:7 ratio," and "1:9 ratio") should yield amplified products, although the presence of αS-dNTPs may decrease yields marginally, particularly when present at their highest concentration (i.e., the 1:7 ratio) (*see* **Note 10**).

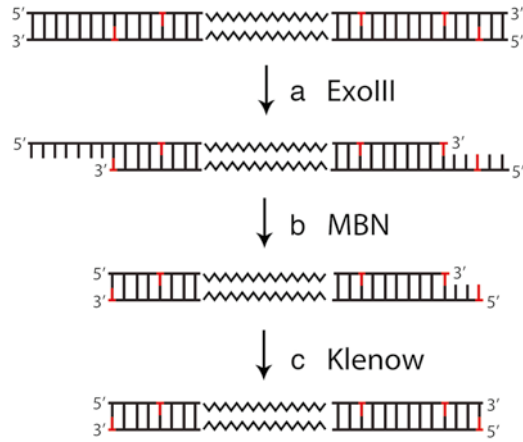

Fig. 2 Critical steps in the thio-ITCHY protocol. The thio-ITCHY PCR product has randomly incorporated phosphorothioate-containing linkages (shown in *red*), due to the presence of αS-dNTPs in the PCR. (**a**) Treatment of the PCR product with exonuclease III removes nucleotides from the 3′ termini. DNA strand hydrolysis is halted when exonuclease III encounters a phosphorothioate-containing linkage. (**b**) Mung bean nuclease (MBN) removes nucleotides from the single-stranded 5′ overhangs. Like exonuclease III, it halts at phosphorothioate-containing linkages (e.g., right-hand end). (**c**) Treatment with Klenow DNA polymerase fills in the remaining gaps, making blunt ends that are suitable for intramolecular ligation

8. Pool the remainder of the "1:7 ratio" and "1:9 ratio" reaction products and purify them using the Cycle Pure Kit. Elute the DNA from the spin column in 50 μL elution buffer and determine its concentration spectrophotometrically. A typical yield is 4–5 μg of DNA (i.e., 50 μL at 80–100 ng/μL).
9. Set up an exonuclease III digestion reaction containing all of the DNA from **step 8**, 1× NEBuffer 1, and 100 U of exonuclease III per microgram of DNA. A typical total reaction volume is 100 μL. Add the exonuclease III last, mix the reaction components well, and incubate at 37 °C for 30 min. This step removes nucleotides from the 3′ ends of the αS-dNTP-containing PCR products (Fig. 2a).
10. Stop the exonuclease III digestion reaction by adding 5 volumes of buffer CP from the Cycle Pure Kit. Purify the DNA using this kit, and elute from the spin column in 50 μL elution buffer. Determine the DNA concentration using a spectrophotometer. A typical yield at this stage is ~2 μg of purified DNA.
11. Set up a mung bean nuclease digestion reaction containing all of the DNA from step 10, 1× mung bean nuclease reaction buffer, and 1.5 U of mung bean nuclease per microgram of DNA (*see* **Note 11**). A typical total reaction volume is 60 μL. Add the mung bean nuclease last, mix the reaction components well,

and incubate at 30 °C for 30 min. This step removes the single-stranded overhangs at the 5′ ends of the αS-dNTP-containing PCR products, although it is important to note that mung bean nuclease is unable to cleave phosphorothioate-containing linkages on the 5′ overhangs (Fig. 2b).

12. Stop the mung bean nuclease digestion reaction by adding 5 volumes of buffer CP from the Cycle Pure Kit. Use the kit to purify the digested DNA and elute from the spin column in 50 μL elution buffer. Determine the DNA concentration using a spectrophotometer. A typical yield is ~1.5 μg of purified DNA.
13. Set up an end-polishing reaction containing all of the DNA from **step 12**, 1×NEBuffer 2, dNTPs at a final concentration of 100 μM (i.e., a 50-fold dilution of the stock used in **step 3**), and 1 U of Klenow DNA polymerase per microgram of DNA. A typical total reaction volume is 60 μL. Add the Klenow DNA polymerase last, mix the reaction components well, and incubate at 25 °C for 15 min. This step removes any 3′ overhangs, and fills in any 5′ overhangs, that may remain after treatment with mung bean nuclease (Fig. 2c).
14. Stop the end-polishing reaction by adding EDTA to a final concentration of 10 mM and heating at 75 °C for 20 min.
15. Use the Cycle Pure Kit to purify the DNA and elute from the spin column in 50 μL elution buffer. Determine the DNA concentration spectrophotometrically. A typical yield is ~1.1 μg of purified DNA.

3.2 Size Selection and Library Construction

1. Mix the purified DNA from the thio-ITCHY protocol (above) with gel loading dye (1× final concentration). Separately, mix ~100 ng of the size standard (*see* **Note 3**) with an aliquot of the same gel loading dye. Run the two samples, together with a DNA ladder, on a 0.8 % w/v agarose gel. A reduced voltage is highly recommended, to improve the resolution of the randomly sized thio-ITCHY DNA fragments. We typically electrophorese at ~13 V/cm (i.e., 80 V for a 6 cm wide mini-gel) for at least 60 min.
2. By comparison with the DNA ladder and the size standard, excise the part of the smear that corresponds to the size standard, ±500 bp (or whatever size range is desired in your library). We strongly recommend using a blue-light transilluminator and SYBR Safe DNA stain, rather than a UV transilluminator and ethidium bromide, for this step (*see* **Note 12**).
3. Purify the DNA from the excised gel band. We use the GenElute Gel Extraction Kit and elute the DNA from the spin column in 50 μL Elution Solution that has been pre-warmed to 60 °C. Use an aliquot (1–2 μL) to determine the DNA concentration spectrophotometrically. A typical yield at this stage is ~600 ng of purified DNA.

4. Set up an intramolecular ligation reaction, using all of the purified DNA from **step 3**. It is critical to use a DNA concentration < 3 ng/μL in the reaction, for high-efficiency intramolecular ligation [10]. We would typically set up a 300 μL ligation reaction that contains ~600 ng DNA, 1×T4 DNA ligase reaction buffer, and T4 DNA ligase (400 cohesive end units; *see* **Note 4**). Incubate the ligation reaction at 16 °C for 12–16 h, and then inactivate the ligase by heating to 65 °C for 10 min.
5. Use the Cycle Pure Kit to purify the DNA from the ligation reaction, and elute from the spin column using 30 μL pre-warmed elution buffer.
6. Use a 2 μL aliquot of the ligated DNA to transform a 50 μL aliquot of *E. coli*, by electroporation (*see* **Note 13**).
7. Immediately after electroporation, add 500 μL of SOC medium to the cuvette, mix by pipetting, and transfer the cells to a sterilized, capped test tube or a 15 mL tube (Falcon or equivalent). Allow the cells to recover by incubating them at 37 °C, with shaking, for 1 h.
8. Store the remainder of the purified ligation reaction (from **step 6**, above) at −20 °C.
9. Spread aliquots (5 μL and 50 μL) of the recovery culture on LB-agar plates. Incubate the plates at 37 °C for 12–16 h.
10. Count the number of colonies on each plate. The colony counts can be used to estimate the size of the final, scaled-up library (*see* Subheading 3.3, **step 10**). The final library is likely to be ~1,500 times larger than the number of colonies on the plate on which 5 μL of the recovery culture was spread. Our scaled-up libraries typically contain 10^5–10^6 clones (*see* **Notes 14** and **15**).

3.3 Sequence Analysis and Scale-Up (See Note 16)

1. Use sterile toothpicks (or 2 μL pipette tips) to pick ~30 colonies at random from the plates described above (Subheading 3.2, **step 10**). Transfer each colony into a thin-walled, 0.2 mL tube containing 5 μL of sterile water.
2. Lyse the cells by incubating the tubes in a thermocycler, at 95 °C for 5 min.
3. Amplify the ITCHY inserts by PCR. While many variations are possible, we have listed our standard protocol below. We typically set up 25 μL PCRs in thin-walled 0.2 mL tubes, as per Table 3.
4. Mix each sample and place the tubes in the thermocycler. Run an appropriate PCR program, with an extension time that is sufficient to amplify inserts with a maximum size that is equal to the combined lengths of the two parental genes.
5. Run 3 μL aliquots of each PCR product on agarose gels, to confirm successful amplification. We typically observe amplified inserts from 60 to 90 % of the colonies screened.

Table 3
Components for a standard colony PCR

Reagent	Volume
Water	9.5 μL
2× Kapa Taq ReadyMix with dye	12.5 μL
Forward primer (from 10 μM stock)	1.0 μL
Reverse primer (from 10 μM stock)	1.0 μL
Cell lysate (from **step 2**, above)	1.0 μL

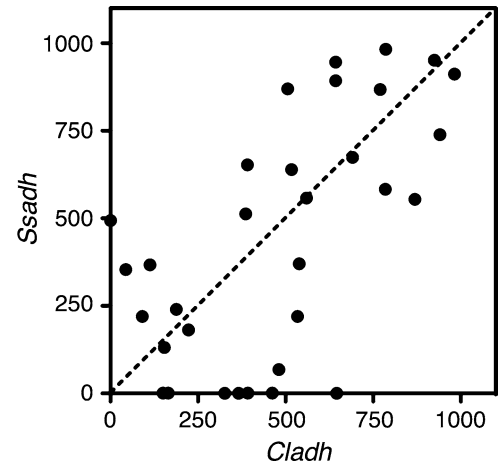

Fig. 3 Crossover distribution of a library made by recombining *Cladh* and *Ssadh*. Sequences from the library are plotted such that the coordinates (*x*, *y*) represent ("last base from *Cladh*," "first base from *Ssadh*"). The *Cladh* gene is 1,056 bp and the *Ssadh* is 1,050 bp. Therefore, points on the diagonal (*dashed line*) correspond to hybrid genes that are the same length as one parent. Points above the *dashed line* indicate hybrid genes that are shorter than either parent, with points below the *dashed line* indicating longer hybrid genes

6. Purify the remainder of each PCR product using the Cycle Pure Kit. Elute the DNA from each spin column in 30 μL of elution buffer.
7. Sequence each PCR product, using the forward and/or reverse primers from the PCR as the sequencing primer(s).
8. Align the sequence of each PCR product with the sequences of the two parental genes, and determine the location of the crossover in each ITCHY variant. Software applications such as MacVector are useful for this analysis.
9. Assess the ITCHY library for potential bias, by plotting the crossover data as shown in Fig. 3. Sequences from an unbiased

library, with effective size selection (Subheading 3.2, **step 2**), should give points that are scattered along the diagonal, as shown in Fig. 3.

10. If the results from Subheading 3.2 (**step 10**) and Subheading 3.3 (**step 9**) are satisfactory, then proceed with constructing the full-sized library. Begin by preparing the LB-agar plates on which the transformed cells of the library will be spread. For each ITCHY library, we typically use two 245 mm × 245 mm square bioassay dishes (*see* **Note 17**) and two regular Petri dishes.
11. Prepare a fresh batch of electrocompetent *E. coli* cells (*see* **Note 5**). We observe transformation efficiencies for freshly prepared cells that are two- to fourfold higher than cells that have undergone a freeze/thaw cycle.
12. Thaw the stored library ligation mixture (Subheading 3.2, **step 8**). Add 2 μL aliquots of it to 14 × 50 μL aliquots of electrocompetent *E. coli* cells. Transform each aliquot and recover the transformed cells, as described above (Subheading 3.2, **steps 6** and **7**).
13. Pool all of the transformed cells in a sterile 15 mL tube, and mix by inverting the tube 2–3 times. The total volume should be 7.7 mL (14 recovery cultures; 550 μL per culture).
14. Spread 2 μL and 20 μL aliquots on the regular LB-agar plates (diluting as necessary to obtain a spreadable volume). Incubate the plates at 37 °C for 12–16 h.
15. Spread the remainder of the library on the two large plates, ~3.85 mL per plate. We recommend incubating these plates at 30 °C, to avoid the possibility of a confluent lawn.
16. Count the number of colonies on the plates from **step 14** above, and extrapolate to estimate the number of colonies on the large plates.
17. The large library plates should be covered in thousands (or millions) of small colonies. To recover them, pipette 4 mL of LB medium (with the appropriate antibiotic) onto the center of the plate. Use a sterilized glass spreader to scrape the cells off the surface and pool them in one corner.
18. Remove the resuspended cells with a P1000 pipettor and transfer them to a sterile 50 mL tube (Falcon tube or similar).
19. Pipette a second 4 mL aliquot of LB onto the same plate and repeat the scraping step.
20. Repeat **steps 17–19** with the second library plate. Pool all of the recovered cells from both plates in a single 50 mL tube.
21. Mix the pooled cells well, by pipetting up and down repeatedly with a P1000 pipettor.
22. Centrifuge the cells at 3,000 × *g* and 4 °C for 15 min. Use a P1000 pipettor to remove the supernatant.

23. Resuspend the cell pellet in 1 mL of LB medium, with antibiotic. Split the library into 100 µL aliquots and transfer each aliquot to a cryogenic vial. Add 50 µL of sterile glycerol (50 % v/v) to each aliquot and mix well by pipetting.
24. Store the aliquots at −80 °C until you are ready to proceed with downstream screening or selection.

3.4 ITCHY by Time-Dependent Truncation (See Note 18)

1. If time-dependent truncation is preferred to thio-ITCHY, then we strongly recommend that users determine the exonuclease III truncation rate under the conditions in their own laboratory. To begin, purify the plasmid containing the two genes that are to be recombined.
2. Linearize the plasmid (~5 µg) by digestion with a restriction enzyme that cleaves between the two parental genes, such as BamHI-HF in Fig. 1 (*see* **Note 19**). After digestion to completion, purify the DNA with the Cycle Pure Kit. Elute from the spin column in 50 µL of pre-warmed elution buffer and determine the DNA concentration using a spectrophotometer. A typical yield of DNA is 4 µg (i.e., 50 µL at 80 ng/µL).
3. A standard truncation reaction should contain DNA at 33 ng/µL and 100 U of exonuclease III per microgram of DNA. The rate of truncation can be controlled by varying the temperature and the concentration of NaCl (*see* **Note 20**). We recommend determining the truncation rate empirically, using a 22 °C incubation and 30 mM NaCl. Mix reagents in a thin-walled 0.2 mL tube, as shown in Table 4.
4. Set a thermocycler (or heating block) to incubate at a constant temperature of 22 °C. Add the tube to the thermocycler and allow the reaction to equilibrate for 2–5 min.
5. Dispense 75 µL aliquots of buffer CP from the Cycle Pure Kit into 8 thin-walled 0.2 mL tubes. Label the tubes "A" to "H" and place them on ice.
6. Remove a 15 µL sample ($t = 0$) from the tube that is incubating at 22 °C. Add the sample to tube A. Mix thoroughly, but quickly, by pipetting up and down.

Table 4
Reaction components for exonuclease III truncation of DNA

Reagent	Volume
Linearized plasmid (80 ng/µL)	50 µL
10× NEBuffer 1	12 µL
NaCl (from 120 mM stock)	30 µL
Water	28 µL

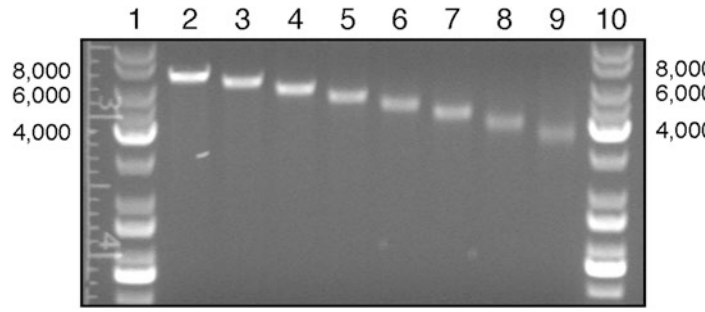

Fig. 4 Empirical determination of exonuclease III truncation rate. A linearized plasmid (pINSALECT-*Cladh*-*Ssadh*, 6,683 bp) was truncated as described in the text. Samples were taken at 5 min intervals, quenched, purified, and run on an ethidium bromide-stained gel (0.8 % w/v agarose). Lanes 1 and 10 contain a DNA ladder, with the sizes of key bands indicated in base pairs. Lanes 2–9 contain the samples, from $t = 0$ (lane 2) to $t = 35$ min (lane 9). The average fragment size decreased from 6,683 bp at $t = 0$ to ~3,900 bp at $t = 35$ min. Therefore, the truncation rate was estimated to be ~80 bases/min

7. There is now 3.5 μg of DNA remaining in the tube at 22 °C. Add 3.5 μL of exonuclease III (i.e., 350 U) to this tube. Mix quickly and thoroughly by pipetting up and down. Start a timer (*see* **Note 6**). Leave the cap of the tube open, to minimize temperature fluctuations due to handling.
8. After incubation for 5 min at 22 °C, remove a 15 μL sample of the truncation reaction and add it to tube B (on ice). Mix with the CP buffer in tube B, by pipetting.
9. Repeat **step 8** at 5 min intervals, until the final sample is taken at $t = 35$ min (and added to tube H, on ice).
10. Purify the DNA in tubes A–H using the Cycle Pure Kit. Elute from each of the eight spin columns in 30 μL of pre-warmed elution buffer.
11. Run 10 μL aliquots of each sample on a 0.8 % w/v agarose gel, alongside a suitable ladder. An example of the expected result is shown in Fig. 4. Use the gel to estimate the truncation rate in bases per minute. For example, from Fig. 4, we estimate the truncation rate under our conditions to be ~80 bases/min. In contrast, the previously reported formula for truncation rate (*see* **Note 20**) would have predicted a truncation rate of ~30 bases/min.
12. With an accurate estimate of the truncation rate in hand, it is now possible to design an optimized strategy for library construction. As with thio-ITCHY, we typically aim to construct hybrids that are the length of one parental gene, ±500 bp (Subheading 3.2, **step 2**). Therefore, in our example (recombining *Cladh* and *Ssadh*, which are 1,050 bp and 1,056 bp,

respectively), we would aim to truncate a total of ~550–1,550 bases from the two genes (which have a combined length of 2,106 bp). A worked example, designed to accomplish this goal, follows (*see* **Note 21**).

13. Prepare more linearized plasmid DNA, as described in **steps 1** and **2** above. Set up a truncation reaction, as described in **steps 3** and **4** above.
14. Dispense 75 μL aliquots of buffer CP from the Cycle Pure Kit into two thin-walled 0.2 mL tubes, labeled "A" and "C." Dispense a 450 μL aliquot of buffer CP into a 1.5 mL tube, labeled "B." Place the three tubes on ice.
15. Remove a 15 μL sample ($t = 0$) from the tube that is incubating at 22 °C. Add the sample to tube A, and mix thoroughly by pipetting up and down.
16. Add 3.5 μL of exonuclease III (i.e., 350 U) to the tube that is incubating at 22 °C. Mix quickly and thoroughly by pipetting up and down. Start a timer (*see* **Note 6**). Leave the cap of the tube open, to minimize temperature fluctuations due to handling.
17. After incubation for 30 s at 22 °C, remove a 2.25 μL sample of the truncation reaction and add it to tube B (on ice). Mix with the CP buffer in tube B, by pipetting.
18. Repeat **step 17** at 30 s intervals, adding each successive sample to tube B. Take the final sample for tube B at $t = 20$ min.
19. Immediately after adding the final 2.25 μL aliquot to tube B, take one final sample of 15 μL and add it to tube C.
20. Purify the DNA in tubes A–C using the Cycle Pure Kit. Elute from spin columns A and C in 30 μL of elution buffer. Elute from spin column B in 50 μL of elution buffer.
21. Check that the truncation has worked as desired, by running 10 μL aliquots of samples A and C (i.e., $t = 0$ and $t =$ final) on an ethidium bromide-stained agarose gel.
22. Use a spectrophotometer to determine the DNA concentration in sample B. Proceed with mung bean nuclease and Klenow DNA polymerase treatment, as described in Subheading 3.1, **steps 11–15** (*see* **Note 22**).
23. The gel in **step 21** (above) will confirm whether or not a size selection step is necessary, prior to setting up the intramolecular ligation reaction. In most cases, careful control of the time-dependent truncation will obviate the need for size selection. Therefore, library construction can proceed as described in Subheading 3.2, **steps 4–10** (*see* **Note 23**).
24. Library sequence analysis, scale-up, and storage should proceed exactly as described in Subheading 3.3, above.

4 Notes

1. Primers should be designed to effect amplification of the entire plasmid, after it has been linearized. The annealing locations of the forward and reverse primers are indicated in Fig. 1 (P1 and P2, respectively). The two primers should be noncomplementary and should have melting temperatures that are within 5 °C of each other. We estimate melting temperatures with the OligoAnalyzer tool from Integrated DNA Technologies (http://www.idtdna.com/analyzer/applications/oligoanalyzer). The optimal annealing temperature to use in the PCR is typically 3–5 °C cooler than the lower of the two primer melting temperatures.
2. As noted previously [10], it is important to use a polymerase that lacks 3′–5′ exonuclease activity. When strand synthesis is completed, proofreading polymerases (Phusion, Vent, etc.) "idle": they continually remove and resynthesize 3′ ends. However, these polymerases are unable to hydrolyze phosphorothioate-containing linkages. Therefore, in the thio-ITCHY PCR, idling would lead to the accumulation of phosphorothioates at the 3′ ends of the amplified products and result in a library that is biased toward genes with small (or nonexistent) truncations.
3. In most scenarios, the user is likely to be recombining two genes and looking for recombinants that are approximately the same length as one parent or the other. In order to enrich the library for these variants, it is important to carry out a size selection step (Subheading 3.2). For accurate size selection, we construct a standard of the desired size (i.e., a linearized plasmid that contains one of the two parental genes).
4. We have had success with the T4 DNA ligases supplied by New England Biolabs (Ipswich, MA, USA), Fermentas (Vilnius, Lithuania), and Enzymatics Inc. (Beverly, MA, USA). We strongly recommend using a ligase reaction buffer that does not contain crowding agents such as polyethylene glycol (PEG), because these agents favor intermolecular ligation rather than the desired intramolecular ligation events. Note that New England Biolabs uses cohesive end units to quantify ligase activity, while other suppliers use Weiss units. We have recommended using 400 cohesive end units of ligase in the ITCHY ligation. This is equivalent to ~6 Weiss units.
5. Within the constraints of downstream screening or selection steps, strains with high transformation efficiencies (such as *E. coli* MC1061) are preferred. We prepare electrocompetent cells according to standard protocols [11].

6. A regular lab timer or stopwatch can be used to determine when samples should be withdrawn from the exonuclease III digestion reaction. We have also found the VersaTimer iPhone app (Vivendo Software; http://www.vivendosoftware.com), to be useful, because it can be programmed to sound an audible alarm at any given interval.
7. It is important that the plasmid DNA is digested to completion; otherwise, the uncut vector can be carried forward through all successive steps and will contaminate the final library with non-truncated variants. We find the high-fidelity restriction enzymes from New England Biolabs (Ipswich, MA, USA) to be preferable for ensuring complete digestion. If digestion is incomplete, it will be necessary to run the digested DNA on an agarose gel and excise the band that corresponds to the linearized plasmid.
8. Varying the ratio of αS-dNTPs to dNTPs affords some control over the lengths of the truncations in the library. Increasing the concentration of αS-dNTPs over dNTPs leads to shorter truncations. However, we have found that PCR product yields begin to decrease when the ratio of αS-dNTPs to dNTPs is increased above 1:7. Empirically, the mixtures that we recommend (1:7 and 1:9 ratios) give an even distribution of truncation lengths, over a range that is suitable for many applications (truncations from 0 to ~1,500 base pairs).
9. Thio-ITCHY amplicons are typically >4 kb. Therefore, optimization (in the absence of αS-dNTPs) may be required. In particular, long incubations at elevated temperatures can result in depurination of the template strands [12]. We recommend reducing the extension temperature from 72 °C to 68 °C and minimizing the denaturation time in each cycle.
10. If the dNTP-only control reaction does not yield an amplified product, then the primers and/or cycling conditions will need to be optimized further. Previous protocols have recommended treating this control in parallel with the αS-dNTP-containing products, for the remainder of the thio-ITCHY protocol [10]. This is good practice for first-time users and controls for the activity of exonuclease III. However, experienced users may wish to forego these steps in the interests of expediency and discard the remainder of the "control" reaction after it has been used to verify the success of the PCR amplification.
11. Note that mung bean nuclease is supplied at a concentration of 10 U/μL. Therefore, it may be necessary to dilute it, prior to adding it to the digestion reaction. We have found 1× mung bean nuclease reaction buffer to be a suitable diluent.

12. UV transillumination of ethidium bromide-stained DNA can induce damage, even when exposure is kept to less than 60 s. This damage can decrease the efficiencies of downstream ligation and transformation steps [13]. Constructing a large ITCHY library requires the highest possible quality of DNA. In our hands, the use of a blue-light transilluminator results in libraries that are five- to tenfold larger than those prepared with DNA that is excised from UV-illuminated gels.
13. Thaw the aliquot of electrocompetent cells on ice. Add the aliquot of DNA to the thawed cells, and transfer the mixture to a pre-chilled, sterile Gene Pulser cuvette. We electroporate using a MicroPulser unit, set to 2.5 kV.
14. In an ITCHY library, the total number of possible variants is found by multiplying the lengths of the two parental genes. For example, we recombined the *C. ljungdahlii* alcohol dehydrogenase gene (1,056 bp) and the *S. solfataricus* alcohol dehydrogenase gene (1,050 bp). There are $1{,}056 \times 1{,}050 = 1.1 \times 10^6$ possible variants. Assuming that each variant is equally likely, our library analysis program GLUE [14] can be used to estimate the number of clones that are required to achieve a desired level of sampling. For example, a library containing ~3.3×10^6 clones is required to sample 95 % of all the possible variants in our alcohol dehydrogenase library.
15. If the test transformation yields an unsuitably low number of colonies, then we recommend scaling up the thio-ITCHY PCRs (and subsequent steps). For example, it is straightforward to set up 200 μL reactions with each ratio of αS-dNTPs to dNTPs (1:7 and 1:9) and split each into 4×50 μL aliquots for thermocycling (Subheading 3.1, **steps 3** and **4**). We have also had success using drop dialysis, rather than silica-based spin columns, for desalting the ligation reaction prior to electroporation [15].
16. We recommend sequencing randomly chosen ITCHY variants from the initial transformation (Subheading 3.2) to assess the distribution of crossovers in the library. This avoids wasting time and resources on a scaled-up library that contains unwanted biases (e.g., too many clones that are under- or over-truncated). We have summarized our protocol for library analysis in Subheading 3.3, though we note that it has also been published in an earlier volume [16].
17. The large volume (200 mL) and high surface area of each bioassay dish make it liable to "sweat" when incubated at 30–37 °C. This can be minimized by cooling the LB-agar to ~55 °C, before pouring the plates. However, the plates may also need to be pre-warmed (37 °C, 4–6 h) and/or dried in a laminar flow hood or class II biosafety cabinet (10–15 min), before they are dry enough to use.

18. In most cases, we recommend the thio-ITCHY protocol, both for its ease of use and for the unbiased libraries that typically result. However, the original incarnation of ITCHY [2], which uses time-dependent exonuclease III digestions, remains useful in cases where the user desires more control over the truncation process. Our protocol modifies and updates the one that has been described previously [10].
19. Exonuclease III will act on blunt or 3′-recessed ends, but not on 3′ overhangs. Ensure that the selected restriction enzyme does not cleave to leave 3′ overhangs. Common enzymes that are unsuitable include KpnI, PstI, and SphI.
20. It is well known that the activity of exonuclease III is temperature and NaCl dependent [17]. Previously, Ostermeier and Lutz reported a formula for estimating the truncation rate at 22 °C, with 33 ng/μL DNA and 100 U of exonuclease III per microgram of DNA [10]:

$$\text{Rate}\left(\text{bases} / \text{min}\right) = 47.9 \times 10^{(-0.00644N)}$$

where N is the concentration of NaCl in millimoles/L (for concentrations in the range 0–150 mM). However, we often observe truncation rates to be higher than this formula suggests, so we recommend estimating truncation rates empirically, to avoid biasing the library toward over-truncated variants.

21. In designing a sampling strategy, consider the maximum truncation length (~1,550 bases in our worked example), the observed truncation rate (~80 bases/min), and the total volume of the exonuclease III digestion reaction (120 μL). We recommend taking 15 μL $t = 0$ and $t =$ final samples for gel analysis. Therefore, in our example, we aimed to sample a total volume of 90 μL, over the course of 20 min (20 min × 80 bases/min = maximum truncation of ~1,600 bases). Sampling every 30 s suggests that each sample should be 2.25 μL. Note that exonuclease III digestion produces a Poisson distribution of truncation lengths [17]. Both our empirical data and a theoretical treatment of the method [18] confirm that uniform distributions of truncation lengths can result, even when the sampling interval (e.g., 30 s) is such that there is an average truncation of more than one base between samples.
22. In theory, the Klenow DNA polymerase step is not necessary in the time-dependent truncation protocol because there are no phosphorothioate-containing linkages to stop mung bean nuclease from digesting the 5′ overhangs (viz., Fig. 2b, c). In practice, we and others [6, 10] have found an end-polishing step to be critical for constructing large libraries.

23. Previous protocols for ITCHY by time-dependent truncation [10] have recommended digestion of the truncated products with a restriction enzyme (e.g., XbaI in Fig. 1), prior to an intermolecular ligation step. The idea was to uncouple the fragments that result from synchronized truncations. However, in our hands, this extra restriction digest leads to ten- to 100-fold decreases in library size. Moreover, we have observed variants with little evidence of synchronization in the truncation lengths for each parental gene. On balance, we recommend omitting the restriction digest.

Acknowledgments

The authors gratefully acknowledge financial support for this work from the University of Otago Priming Partnerships fund. W.M.P. is the recipient of a Rutherford Discovery Fellowship.

References

1. Lutz S, Patrick WM (2004) Novel methods for directed evolution of enzymes: quality, not quantity. Curr Opin Biotechnol 15:291–297
2. Ostermeier M, Shim JH, Benkovic SJ (1999) A combinatorial approach to hybrid enzymes independent of DNA homology. Nat Biotechnol 17:1205–1209
3. Gerth ML, Lutz S (2007) Non-homologous recombination of deoxyribonucleoside kinases from human and *Drosophila melanogaster* yields human-like enzymes with novel activities. J Mol Biol 370:742–751
4. Eisenbeis S, Höcker B (2010) Evolutionary mechanism as a template for protein engineering. J Pept Sci 16:538–544
5. Lutz S, Ostermeier M, Benkovic SJ (2001) Rapid generation of incremental truncation libraries for protein engineering using α-phosphothioate nucleotides. Nucleic Acids Res 29:e16
6. Tarendeau F, Boudet J, Guilligay D, Mas PJ, Bougault CM, Boulo S, Baudin F, Ruigrok RW, Daigle N, Ellenberg J, Cusack S, Simorre JP, Hart DJ (2007) Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. Nat Struct Mol Biol 14:229–233
7. Gerth ML, Patrick WM, Lutz S (2004) A second-generation system for unbiased reading frame selection. Protein Eng Des Sel 17:595–602
8. Lutz S, Ostermeier M (2003) Preparation of SCRATCHY hybrid protein libraries: size- and in-frame selection of nucleic acid sequences. Methods Mol Biol 231:143–151
9. An Y, Yumerefendi H, Mas PJ, Chesneau A, Hart DJ (2011) ORF-selector ESPRIT: a second generation library screen for soluble protein expression employing precise open reading frame selection. J Struct Biol 175:189–197
10. Ostermeier M, Lutz S (2003) The creation of ITCHY hybrid protein libraries. Methods Mol Biol 231:129–141
11. Hanahan D, Jessee J, Bloom FR (1991) Plasmid transformation of *Escherichia coli* and other bacteria. Methods Enzymol 204:63–113
12. Cheng S, Fockler C, Barnes WM, Higuchi R (1994) Effective amplification of long targets from cloned inserts and human genomic DNA. Proc Natl Acad Sci U S A 91:5695–5699
13. Hartman PS (1991) Transillumination can profoundly reduce transformation frequencies. Biotechniques 11:747–748
14. Firth AE, Patrick WM (2005) Statistics of protein library construction. Bioinformatics 21: 3314–3315
15. Saraswat M, Grand RS, Patrick WM (2013) Desalting DNA by drop dialysis increases library size upon transformation. Biosci Biotechnol Biochem 77:402–404
16. Hanson-Manful P, Patrick WM (2013) Construction and analysis of randomized protein-encoding libraries using error-prone PCR. Methods Mol Biol 996:251–267
17. Hoheisel JD (1993) On the activities of *Escherichia coli* exonuclease III. Anal Biochem 209:238–246
18. Ostermeier M (2003) Theoretical distribution of truncation lengths in incremental truncation libraries. Biotechnol Bioeng 82:564–577