

Synthetic customized scFv libraries

Gautier Robin^{1,2} and Pierre Martineau¹

¹ IRCM, Institut de Recherche en Cancérologie de Montpellier, Montpellier, F-34298, France ; INSERM, U896, Montpellier, F-34298, France ; Université Montpellier1, Montpellier, F-34298, France ; CRLC Val d'Aurelle Paul Lamarque, Montpellier, F-34298, France.

² BioXtal SA, 163 Avenue de Luminy, 13288 Marseille Cedex 09, France.

pierre.martineau@inserm.fr

Web: <http://www.ircm.fr/>

Summary

Libraries of antibody fragments displayed on filamentous phages have proved their value to generate human antibodies against virtually any target. We describe here a simple protocol to make large and diverse libraries based on a single or few frameworks. Diversity is introduced in the third hypervariable loops using degenerate synthetic oligonucleotides and PCR assembly. Because all the antibody fragments isolated from the library will share their framework sequence, their stability and physical properties will be more consistent and customizable than when antibody fragments are isolated from a library prepared from human donors.

Key Words

PCR assembly, antibody fragment, single-chain Fv, phage-display, synthetic library

Running Head

Synthetic scFv libraries

1. Introduction

Antibodies have proved their value as therapeutic molecules and several of them have already reached FDA approval and generate annual sales over 1 billion dollars a year (**1**). These antibodies are mainly issued from mouse hybridoma or phage-displayed antibody fragments screening. The latter allows to directly isolate human antibody fragments, without animal immunization, from a unique and inexhaustible source (**2**).

Phage-displayed antibody libraries are usually made either from a natural source of recombinant antibodies obtained from a pool of human donors (**2**) or entirely *in vitro* using cloned antibody genes. The large sequence diversity in libraries made from natural source leads to a large variability in stability, expression levels and physical properties in the pool of antibodies (**3**). On the other hand, the *in vitro* approach allows the construction of libraries built on a selected limited set of sequences leading to more consistent and optimized characteristics. Despite their lower structural diversity, the possibility to select antibody fragments with high specificity and high affinity against any antigen from such libraries has been demonstrated (**4-6**).

In the presented protocol, we will use the *in vitro* approach to construct a library based on the single hyperstable scFv13R4 framework (**7**) (see **Note 1**). We will first introduce random sequences in the two CDR3 loops using degenerate synthetic oligonucleotides. Diversity of natural antibodies is mainly due to the sequence variability of the CDR3 loops and restraining the diversity to these loops is sufficient to generate very efficient and diverse libraries of binders (**5,6**).

The protocol presented here that presents insertion of a single loop length in an unique framework may also be used to introduce multiple CDR3 lengths in multiple frameworks, simply by applying it to each pair of framework/length, and then by pooling the libraries obtained. The pooling may be completed either in equimolar ratio or following the natural distribution of loop lengths (5).

2. Materials

All buffers must be prepared with ultra-pure water and ACS grade chemicals, and stored at room temperature unless otherwise indicated.

2.1. Introduction of random CDR3 loops

1. Thermocycler.
2. Phusion High-Fidelity DNA Polymerase and reagents (Thermo Scientific #F530) (*see Note 3*).
3. Oligonucleotides (**Table 1** and **Fig. 1**) (*see Notes 1 & 2*).
4. SYBR Safe™ DNA Gel Stain 10,000x concentrate (Invitrogen #S33102) (*see Note 4*).
5. Safe Imager™ blue light transilluminator (Invitrogen #G6600EU) (*see Note 4*).
6. TAE Buffer. TAEx50: 242 g Tris, 57.1 mL glacial acetic acid, 37.2 g Na₂EDTA.2H₂O, make up to 1 liter (pH ~ 8.5). Dilute 50 times with ultra-pure water before use.
7. 1% agarose gel: add 1 g of agarose to 100 mL of TAE buffer, boil in a microwave oven until completely melted, cool down to 45-50°C, add 10 µL of SYBR Safe™.
8. Macherey-Nagel NucleoSpin Extract II (#740609).

2.2. Cloning

1. *NcoI* (10 u/µL, NEB #R0193) and *NotI* (10 u/µL, NEB #R0189) restriction enzymes. 10xNEB3 buffer and 100xBSA from NEB.
2. 100 bp and 1 kb DNA Ladder (NEB #N3231 and #N3232) (*see Note 5*).
3. pCANTAB6 vector (1 µg/µL) (*see Note 6*).
4. T4 DNA Ligase (6 Weiss u/µL) (Fermentas #EL0014) (*see Note 7*).
5. Glucose 40% (w/v), autoclaved.
6. LBGA plates: 10 g tryptone (peptone), 5 g yeast extract, 10 g NaCl, make up to 1 liter with water, adjust pH to 7.0 with 5 N NaOH, add 15 g of agar and autoclave. Allow the solution to cool to 55-60°C, add 50 mL of 40% glucose solution, 100 µg/mL ampicillin, then pour the plates.

2.3. Transformation in electrocompetent bacteria

1. LB/Tet plates: prepare as for LBGA but instead of glucose and ampicillin add 1 mL of a 12 mg/mL tetracycline solution in 70% ethanol.
2. Cmax5αF², freshly streaked on a LB/Tet plate (*see Note 8*).
3. 2xTY: 16 g tryptone (peptone), 10 g yeast extract, 5 g NaCl, make up to 1 liter with water, adjust pH to 7.0 with 5 N NaOH, autoclave.
4. HEPES 1M: weight 2.38 g of HEPES, add 8 mL of H₂O, adjust pH to 7.0 and the volume to 10 mL. Sterilize by filtration.
5. Glycerol/HEPES: Weight 10 g of glycerol, make up to 1 liter with water, autoclave. Add 1 mL of sterile HEPES 1M.

6. H₂O/HEPES: Add 1 mL of sterile HEPES 1M to 1 liter of autoclaved ultra-pure water.
7. SOC: 20 g tryptone (peptone), 5 g yeast extract, 0.5 g NaCl, 10 mL of 250 mM KCl (18.6 g/l), make up to 1 liter with water, adjust pH to 7.0 with 5 N NaOH, autoclave. Allow to cool and add 5 mL of sterile 2M MgCl₂ (190.4 g/l, autoclaved) and 9 mL of sterile 40% glucose.
8. Magnetic stir bars, 2-3 cm long. Sterilize by autoclaving and store at 4°C.
9. Biorad GENE PULSER II and 0.2 cm gap cuvettes (*see Note 9*).
10. Large 245 mm x 245 mm square Petri dishes (Nunc #240835). Use 250 mL of LBGA per plate.
11. 14 mL sterile polypropylene round-bottom culture tubes (17 mm x 100 mm. BD Falcon #352059).
12. Glycerol 40%: Weight 40 g of glycerol, make up to 0.1 liter with water, autoclave.

3. Methods

3.1. Introduction of random CDR3 loops

1. Prepare 3 PCR by mixing on ice (*see Note 10*).

Template (pAB1-13R4)	10 ng
HF buffer (x5)	5 µL
dNTP (10 mM)	0.5 µL
Primer 1 (10 pMol/µL)	1.25 µL
Primer 2 (10 pMol/µL)	1.25 µL
H ₂ O	to 24.75 µL
Phusion	0.25 µL

Using the 3 primer pairs (**Fig. 2**): M13.rev/H3.rev; VH.FR4.for/VL.FR3.rev; L3.for/M13.uni.

2. Cycle 30 times: 98°C 30 s / (98°C 10 s / 65°C 10 s / 72°C 15 s) x 30 / 72°C 5 min / 4°C hold (*see Note 11*).
3. Separate the PCR on a 1% agarose gel. Cut the 3 bands and pool them.
4. Purify the mix on a column of NucleoSpin Extract II. Elute in 20 µL.
5. Add 5 µL of HF Buffer (x5), 0.5 µL of dNTP (10 mM), 0.25 µL of Phusion (Final volume = 25 µL).
6. Cycle 10 times: 98°C 30 s / (98°C 10 s / 65°C 10 s / 72°C 15 s) x 10 / 72°C 5 min / 4°C hold.
7. Add in the following order on ice: 310 µL H₂O, 100 µL HF (x5) buffer, 10 µL dNTP (10 mM), 25 µL M13.rev, 25 µL M13.uni, and 5 µL of Phusion (Final volume = 500 µL).
8. Distribute in 10 PCR tubes and cycle 30 times: 98°C 30 s / (98°C 10 s / 70°C 30 s / 72°C 15 s) x 10 / 72°C 5 min / 4°C hold.
9. Pool the 10 tubes and analyze 2 µL on a 1% agarose gel using 0.5 µg of 100 bp DNA Ladder as marker.

10. Purify on a NucleoSpin Extract II column using the “Protocol for PCR clean-up”. Elute in 41 μL .

3.2. Cloning in phage-display vector

Insert digestion

1. Add to the insert prepared in Sect. 3.1: 5 μL of 10xNEB3 buffer, 0.5 μL of 100xBSA, 2 μL of *NcoI*, and 2 μL of *NotI*. Incubate for 3 h at 37°C.
2. Purify on a 1% agarose gel using a NucleoSpin Extract II column. Elute in 100 μL .
3. Run 1 μL on a 1% agarose gel using 0.5 μg of 100 bp DNA Ladder, to check insert size and to estimate its amount (*see Note 12*)

Vector digestion

4. Mix the following: 20 μL of pCANTAB6 vector (20 μg), 5 μL of 10xNEB3 buffer, 0.5 μL of 100xBSA, 2 μL of *NcoI*, and 2 μL of *NotI*, 20.5 μL of H_2O . Incubate for 3 h at 37°C.
5. Purify on a 1% agarose gel using a NucleoSpin Extract II column. Elute in 100 μL .
6. Estimate the vector concentration using 1 μg of the 1 kb DNA ladder as a reference (*see Note 12*).

Ligation

Prepare 3 ligations using 1:0.5, 1:1 and 1:2 vector:insert molar ratios (*see Note 13*).

7. Prepare 3 ligations:

Digested pCANTAB6 (4500 bp)	50 ng (0.3 - 0.5 μL)
Digested insert (750 bp)	4.15 ng / 8.3 ng / 16.7 ng
Ligase buffer (x10)	0.5 μL
Ligase diluted 1/10 in 1x ligase buffer	1 μL (0.5 Weiss units)
H_2O	to 5 μL

9. Incubate overnight at 16°C in a water bath.
10. Heat for 10 min. at 65°C.
11. Transform 2 μL of the ligation by electroporation (*see Sect. 3.3 and Note 14*).
12. Plate 100 μL of $10^{-2}/10^{-3}/10^{-4}$ dilutions on LBGA plates and incubate overnight at 37°C.
13. Count the clones and measure the cloning efficiency by colony PCR (M13.rev/M13.for) or restriction (*see Note 15*).
14. Prepare the final ligation in 500 μL using the best determined ratio:

Digested pCANTAB6 (4500 bp)	5 μg
Digested insert (750 bp)	0.42, 0.83 <i>or</i> 1.67 μg
Ligase buffer (x10)	50 μL
Ligase (5 u/ μL)	10 μL (50 Weiss units)

H₂O

to 500 µL

15. Incubate overnight at 16°C in a water bath. Heat for 10 min. at 65°C.
16. Purify on a NucleoSpin Extract II column using the “Protocol for PCR clean-up”. Elute in 80 µL.

3.3. Transformation in electrocompetent bacteria

Use *freshly-prepared electrocompetent cells* following the protocol below in order to obtain the very high transformation efficiency (typically $5 \cdot 10^9$ - $2 \cdot 10^{10}$ transformants/µg of supercoiled pUC18 plasmid) required for the final library transformation (steps 11 – 30). We advise against frozen-thawed cells, even from commercial source. If not familiar with electrocompetent cell preparation, training to reliably prepare highly competent cells before performing the final transformation should be considered in order to avoid having to prepare the ligation all over again. (*see Note 16*).

All material must be pre-cooled and kept as close to 4°C as possible in an ice/water bath throughout the preparation. If possible, work in a cold room. The centrifuge and the rotor must be pre-cooled to 4°C.

Preparation of electrocompetent cells

1. Pick 1 *fresh* colony of Cmax5αF' in a 50 mL flask containing 10 mL of 2xTY and 12 µg/mL of tetracycline, and grow overnight at 37°C with vigorous shaking (220 rpm) (*see Note 8*).
2. Pour the flask content in a 5 liter flask containing 1 liter of 2xTY and 12 µg/mL of tetracycline, and grow at 37°C with vigorous shaking until OD_{600nm} reaches 0.7 (*see Note 17*).
3. Pour the flask content in two 500 mL centrifuge bottles and cool down in an ice/water bath for 30 min. Mix regularly and gently the bottles.
4. Centrifuge at 5,000 g for 5 min at 4°C and discard the supernatant.
5. Add a cold and sterile magnetic stir bar and 500 mL of cold H₂O/HEPES to each bottle. Resuspend the pellet using a magnetic stirrer. Start with a vigorous stirring until the pellet detaches from the bottle; continue with a slower rotation rate until all the bacteria are completely resuspended. You may also gently mix the bottle by turning it upside down several times.
6. Centrifuge at 5,000 g for 10 min at 4°C and discard the supernatant gently, carefully avoiding to disturb the pellet *containing the stir bar*.
7. Repeat steps 5 and 6.
8. Resuspend, as in step 5, in 50 mL of cold glycerol/HEPES. Pool the two bottles in a new centrifuge bottle. *Do not transfer the stir bars*.
9. Centrifuge at 5,000 g for 15 min at 2°C and discard the supernatant.
10. Resuspend the pellet in 1 mL of cold glycerol/HEPES *using a 10 mL pipette*. The final volume should be about 2 mL. (*see Notes 8 & 18*).

Library electroporation

11. Prepare four sterile 100 mL Erlenmeyer flasks containing 25 mL of SOC and two 14 mL sterile polypropylene culture tubes containing 0.95 mL of SOC.
12. Warm to 37°C for at least 1 h.
13. Cool on ice: 5 electroporation cuvettes, 5 Eppendorf tubes, and the slide that holds the cuvette in the electroporator (*see Note 19*).

14. In a pre-chilled Eppendorf tube, mix 300 μL of competent cells and 20 μL of the purified ligation (Sect. 3.2).
15. Transfer the mix in a pre-chilled electroporation cuvette. Be sure to put all the sample at the bottom of the cuvette by gently taping the bottom of the cuvette on a flat surface.
16. Apply an electric pulse using the following conditions: 2500 V, 25 μF , 200 Ω .
17. Immediately transfer the cells to one of the pre-warmed Erlenmeyer flask by washing the sample with 1 mL of outgrowth medium using a Pasteur pipette. (*see Note 20*).
18. Immediately transfer the flask to a 37°C incubator and shake vigorously (220 rpm) for 1 h.
19. Repeat steps 5-8 three times for the rest of the ligation (3 x 20 μL).
20. Negative control: Add 40 μL of competent cells to one of the 2 tubes of SOC.
21. Positive control: Add 1 μL of highly purified supercoiled pUC18 (10 $\text{pg}/\mu\text{L}$) plasmid to 40 μL of competent cells in one of the pre-chilled Eppendorf tube. Follow steps 5-8 but resuspend in 0.95 mL of SOC using the second pre-warmed tube. (*see Note 21*).
22. Incubate the flasks and the tubes for 1 h at 37°C with vigorous shaking.
23. Pool the 4 flasks. Plate on LBGA plates: 100 μL of the negative control; 100 μL of 10^{-1} and 10^{-2} dilutions of the positive control; 100 μL of 10^{-2} , 10^{-3} , 10^{-4} and 10^{-5} dilutions of the library.
24. Centrifuge the 100 mL of cells at 5,000 g for 10 min at 8°C and discard the supernatant.
25. Resuspend the pellet in 4 mL of SOC and plate on 4 large square Petri dishes of LBGA and incubate overnight at 37°C.
26. Scrap the cells from each square plate in 10 mL of 2xTY/glycerol (7.5 mL 2xTY / 2.5 mL 40% glycerol) and pool them.
27. Measure the $\text{OD}_{600\text{nm}}$ of a 1/200 dilution in 2xTY. Calculate the number of cells/mL assuming that a cell culture at $5 \cdot 10^8$ cells/mL has an $\text{OD}_{600\text{nm}}$ of 1 for a 1 cm optical path length. (*see Note 22*).
28. Calculate the size of the library using the series of dilutions plated in step 23.
29. Aliquot the library so that the number of cells in each tube is at least 20 times the library size, flash-freeze and store at -80 °C (*see Note 23*).
30. Measure the percentage of ampicillin resistant clones by plating serial dilutions of the library on LBGA and LBG plates. Incubate overnight at 37°C. (*see Notes 24 – 26*).

4. Notes

1. ScFv13R4 is a highly expressed human scFv. The scFv gene is cloned in a pUC119 derived plasmid. You will have to modify the oligonucleotides in Table 1 according to your scFv and vector sequences.
2. We will introduce random sequences using degenerate codon NNK (K = T or G). This degenerate codon will code for the 20 amino acids and the amber stop codon (TAG). The advantage of such a design is that the oligonucleotide can be easily synthesized at low cost. More efficient/subtle designs to optimize the expressed and functional clones are possible using either tri-nucleotides (4) or spiked oligonucleotides (5).
3. The polymerase used must generate blunt-end fragments for the assembly to succeed. It is also better to use a high-fidelity polymerase to avoid mutations. We obtain very high yields and low error rates with the Phusion polymerase but you may substitute it with your favorite high-fidelity polymerase.

4. Using SYBR safe™ and a blue light transilluminator results in a much higher cloning efficiency. Indeed, exposition to UV light, when working with EtBr, damages DNA and thus lowers cloning efficiency. If you do not have access to a blue light transilluminator, DNA damages may be reduced by using long-wave UV transilluminator (365 nm) and/or by adding a pile of transparent plastic papers between the transilluminator and the gel to decrease the UV intensity. Alternative sources for blue transilluminators and non-UV intercalating reagents are available (Clare Chemical Research <http://www.clarechemical.com/transilluminator.htm>, http://openwetware.org/wiki/User:Norman_Wang/Blue_Light_Transilluminator).
5. Any DNA marker can be used but the amount of each marker band needs to be known to quantify the purified PCR products and vector. For 0.5 µg of the two markers used in this protocol, the band amounts are between 18 ng and 125 ng.
6. pCANTAB6 is a phagemid vector derived from pUC119. It allows to display the scFv at the surface of the M13 filamentous phage (5). If your cloning vector does not have restriction sites compatible with your scFv, you will need to redesign the oligonucleotides M13.rev and M13.uni in Table 1 to introduce new restriction sites at the two extremities of your scFv gene. In such a case, check in the NEB or Fermentas catalogs that the corresponding restriction enzymes are able to cut DNA close to the end of a DNA fragment.
7. NEB does not use Weiss but cohesive end units. You can convert between the two units using the formula: 1 Weiss unit = 200 NEB cohesive end units. Always aliquot the 10x ligase buffer since it contains ATP. You can freeze/thaw an aliquot up to 10 times but it is better to use a fresh aliquot to perform the final ligation (Sect. 3.2 step 14).
8. Cmax5αF' is Φ80dlacZΔM15Δ(lacZYA – argF)U169 recA1 endA1 hsdR17 (r-k , m+k) *phoA supE44 λ- thi-1 gyrA96 relA1* / F' [*lacI*^Q Tn10 (Tet)] and can be purchased from BioRad (#170-3341). Always streak on a plate the day before use. We have tested several *E. coli* strains for transformation efficiency. Cmax5αF' has given the best results for a male, restriction (*hsdR*) and recombination minus (*recA*) strain giving efficiencies between 5.10⁹ and 2.10¹⁰ clones per µg of supercoiled pUC18 plasmid. Since Bio-Rad does not commercialize the strain anymore, a good widely-available alternative is XL1-Blue. It is much better to use a *recA* strain to avoid library instability but if instability is not a concern, then the most efficient strain is MC1061 (http://www.bio-rad.com/cmcc_upload/Literature/12864/M1652101C.pdf) (8). We have always obtained better results with freshly prepared electrocompetent cells than with commercial ones. In addition, commercial electrocompetent cells are very expensive given the high volume used (1.2 mL). If the library is constructed to make a phage-displayed library using a phagemid then the strain must be male (F', F+ or Hfr) to allow helper phage infection.
9. BioRad 0.2 cm gap cuvettes (#165-2086) allow electroporation of 400 µL of cells. Higher efficiencies can be obtained by using 0.1 cm gap cuvettes; however their maximum volume capacity is only 80 uL which implies having to perform 4-5 times more electroporation experiments.
10. Since high-fidelity polymerase enzymes exhibit 3' -> 5' exonuclease activity, it is important to keep everything on ice and to add the polymerase at the end, after the dNTP mix.
11. PCR conditions will depend on the enzyme and the oligonucleotide sequences. Adapt the conditions according to the manufacturer's instructions.
12. Visual comparison of the band intensities is usually sufficient. However softwares such as ImageJ (<http://rsbweb.nih.gov/ij/>) can be used for more accurate quantifications. Expected concentration should be approximately 30 ng/µL for the insert and 100-150 ng/µL for the vector, after digestion and purification. The band intensity should be comparable to those obtained with 0.5-1 µg of the DNA markers (see Note 5).
13. Use 50 ng of vector DNA and adjust the amount of insert according to the size of the vector in order to have the correct molar ratios. In our case, the vector and insert lengths are

respectively 4500 and 750 bp. For 50 ng of vector and a 1:1 molar ratio, the amount of insert to use is $50 \times 750 / 4500 = 8.3$ ng.

14. Any batch of competent cells is appropriate since we just want to compare the three ligation ratios. We usually use 40 μ L of frozen electrocompetent cells ($\sim 10^8$ - 10^9 transformants/ μ g).
15. Choose the vector:insert ratio that gives the highest number of clones with 100% of positive clones. Most of the time, the 1:1 ratio gives the best results with about 10^5 - 10^6 transformants using frozen electrocompetent cells. In addition, you should obtain 100% of positive clones by colony-PCR. If it is not the case, it is presumably because your vector and/or your insert were not completely digested. In such a case, make new digestions and purifications, but increase the digestion time (5-16 h). Some authors use a phosphatase treatment to reduce vector self-ligation, but we found that this treatment decreases the number of clones obtained and is thus not suitable for making libraries. It is of course required to use two different enzyme sites for cloning and trying to make libraries using blunt-end or two identical enzymes will always result in a much lower number of productive clones.
16. Obtaining very high transformation efficiencies, especially during the final library transformation, is essential to maintain the sequence diversity throughout the library construction. The protocol described here is adapted from the method of Sidhu and collaborators (8). We have tested several commercial brands of competent cells and several cell lines and we have always obtained the best results with this protocol and Cmax5 α F' strain. It is possible to obtain a two-fold higher efficiency with MC1061 strain and its male derivatives (*see Note 8*), but we prefer to use Cmax5 α F' because it is *recA* deficient and produces high quality DNA (*endA*). If not familiar with electrocompetent cell preparation, then we highly recommend to train to reliably prepare highly competent cells since you cannot use frozen cells for the final library transformation. We indeed found that, whatever the method used, freezing the competent cells results in a 5-10 fold decrease in the transformation efficiency. This is not a problem for most experiments (as in Sect. 3.2 step 11), but would result in a smaller library in Sect. 3.3. However, the amount of insert and vector produced will allow you to make about 2-3 ligations and you can thus reproduce easily Sect 3.3 if your first transformation is not good enough (*see Note 26*). Always include the positive control to measure the transformation efficiency of your cells to determine if a low number of clones is due to the electroporation experiment or to the ligation. Good electrocompetent Cmax5 α F' cells should give between $5 \cdot 10^9$ and $2 \cdot 10^{10}$ transformants per μ g of supercoiled pUC18 plasmid.
17. The cells must reach this absorbance in about 3 h. Slower growth (> 5 h) are often due to traces of detergents in flasks. You must then restart a fresh culture in a new clean sterile flask washed several times with sterile water or use a disposable plastic flask.
18. Competent cells can be frozen but the transformation efficiency will decrease 5-10 times. Aliquot in sterile pre-chilled 0.5 mL Eppendorf tubes, flash-freeze in a dry ice/ethanol bath/liquid nitrogen, and store at -80°C .
19. See figure 7 in the Bio-rad technical note MC1652101C (*see Note 8*).
20. Minimizing the duration between the pulse and the cell transfer to outgrowth medium is crucial for high recovery of *E. coli* transformants. Increasing the duration by as little as 1 minute causes a 3-fold drop in transformation efficiency.
21. Do not keep the 10 pg/ μ L dilution of pUC18 for later experiments, even frozen. Stocks must be kept frozen at at least 10 ng/ μ L and thawed to make a fresh dilution on the day of the experiment.
22. Although the formula should provide with a good estimate, the exact formula depends on the spectrophotometer and the cell type. For a more accurate measurement you can calibrate

your spectrophotometer by measuring the OD_{600nm} of a sample and determining the corresponding number of colonies obtained after plating several dilutions of the sample.

23. If the library diversity is D , and the clone number in your tube is N , the probability not to miss a given clone is $(1-(1-1/D)^N)$. If the library size is 10^9 (D), using an aliquot of $20 \cdot D$ ($N = 2 \cdot 10^{10}$) results in a probability of 0.999999998 ($1 - 2 \cdot 10^{-9}$). You can increase the aliquot size to increase the chances of having at least 1 copy of each clone in every aliquot.
24. Since the library will grow to confluence on the plate, it is important to check that ampicillin has not been exhausted too early and that a substantial fraction of ampicillin-resistant clones is present. We usually obtain close to 100% of ampicillin-resistant clones even with 10^9 clones on a single large square plate. If this is not the case, you must correct the value in step 19 (N value in **Note 23**). For instance, if you get only 70% of ampicillin-resistant cells, then you should consider that your aliquots contain only 14 times the library size instead of 20 times (70% of 20).
25. If all the steps are correctly performed, you should obtain at least 10^9 clones. Of course, this number is estimated from the dilution plates. The real diversity can be determined more accurately using high throughput sequencing (**9**).
26. You may pool the result of several independent experiments to get a higher diversity. In such a case, you must pool the different libraries in amounts corresponding to their diversity. This is easy if you have always made aliquots of 20 times the library size in Sect. 3.3 step 29 since you can simply pool together one aliquot of each sub-library: The diversity of the resulting library will be the sum of the diversities and the pool will contain a number of cells of 20 times the library size.

5. References

1. Scolnik, P. A. (2009) mAbs: a business perspective. *MAbs* 1, 179-184.
2. Schofield, D. J., Pope, A. R., Clementel, V., Buckell, J., Chapple, S. D., Clarke, K. F., Conquer, J. S., Crofts, A. M., Crowther, S. R. E., Dyson, M. R., Flack, G., Griffin, G. J., Hooks, Y., Howat, W. J., Kolb-Kokocinski, A., Kunze, S., Martin, C. D., Maslen, G. L., Mitchell, J. N., O'Sullivan, M., Perera, R. L., Roake, W., Shadbolt, S. P., Vincent, K. J., Warford, A., Wilson, W. E., Xie, J., Young, J. L. and McCafferty, J. (2007) Application of phage display to high throughput antibody generation and characterization. *Genome Biol* 8, R254.
3. Demarest, S. J. and Glaser, S. M. (2008) Antibody therapeutics, antibody engineering, and the merits of protein stability. *Curr Opin Drug Discov Devel* 11, 675-687.
4. Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wolle, J., Pluckthun, A. and Virnekas, B. (2000) Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J Mol Biol* 296, 57-86.
5. Philibert, P., Stoessel, A., Wang, W., Sibler, A., Bec, N., Larroque, C., Saven, J. G., Courtete, J., Weiss, E. and Martineau, P. (2007) A focused antibody library for selecting scFvs expressed at high levels in the cytoplasm. *BMC Biotechnol* 7, 81.
6. Silacci, M., Brack, S., Schirru, G., Marland, J., Ettore, A., Merlo, A., Viti, F. and Neri, D. (2005) Design, construction, and characterization of a large synthetic human antibody phage display library. *Proteomics* 5, 2340-2350.
7. Martineau, P., Jones, P. and Winter, G. (1998) Expression of an antibody fragment at high levels in the bacterial cytoplasm. *J Mol Biol* 280, 117-127.
8. Sidhu, S. S., Lowman, H. B., Cunningham, B. C. and Wells, J. A. (2000) Phage display for selection of novel binding peptides. *Methods Enzymol* 328, 333-363.
9. Ravn, U., Gueneau, F., Baerlocher, L., Osteras, M., Desmurs, M., Malinge, P., Magistrelli, G., Farinelli, L., Kosco-Vilbois, M. and Fischer, N. (2010) By-passing in vitro screening—next

generation sequencing technologies applied to antibody display and in silico candidate selection. *Nucleic Acids Res* 38, e193-e193.

Figure captions

Figure 1. Design of the mutagenic oligonucleotides. The H3.rev degenerate oligonucleotide (Table 1) to introduce diversity in the CDR3 VH loop is designed in three parts: 1) A 5' sequence that matches the reverse of the 5' sequence of the FR4 VH gene with a T_m of 60-70°C (usually 18-25 bp); 2) A degenerate core that will encode the CDR3 loop. We use here a fully degenerated sequence encoded by NNK codons (NNM on the reverse strand) (*see Note 2*); 3) A 3' sequence that matches the reverse of the 3' sequence of the FR4 VH gene with a T_m of 60-70°C. Oligonucleotide VH.FR4.for is the complementary DNA of the 5' end of H3.rev and allows the assembly of the fragments by PCR. The L3.for (Table 1) oligonucleotide was designed following the same principles, except that it was the complementary DNA of the reverse strand. In the case of the VL it is however usually better not to degenerate the 5' part of the loop since it is germline encoded and not highly variable (**4,6**). T_m values were calculated using Finnzymes' recommended method for the Phusion polymerase enzyme used in the protocol (https://www.finnzymes.fi/tm_determination.html).

Figure 2. Library construction. The figure outlines the main steps to introduce diversity in the CDR3 loops of a single scFv. Several variations of the protocol are possible. For instance, diversity may also be introduced in other CDR loops by designing additional oligonucleotide pairs and we have successfully assembled scFv genes from ten fragments using this method. The first PCR could also be omitted and replaced by synthetic oligonucleotides, resulting in a fully synthetic scFv library.

Table 1. Oligonucleotides ^a

Name	Sequence (5' -> 3')
M13.rev	GAGCGGATAACAATTTTCACACAGG
M13.uni	AGGGTTTTCCCAGTCACGACGTT
VH.FR4.for	TGGGGCAGAGGCACCCT
VL.FR3.rev	GCAGTAATAATCAGCCTCGTCC
H3.rev ^b	AGGGTGCCTCTGCCCCA (MNN) ₅₋₂₀ TCTCACACAGTAATAAACAGCCG
L3.for ^b	GGACGAGGCTGATTATTACTGC (NNK) ₉₋₁₁ TTCGGCGGAGGGACCAAG

^a The T_m of the oligonucleotides are all between 60 and 70 °C. ^a Design of the degenerated oligonucleotides is explained in **Fig. 1** (see **Note 2**).