

Functional enrichment by direct plasmid recovery after fluorescence activated cell sorting

Balakrishnan Ramesh, Christopher S Frei, Patrick C Cirino, and Navin Varadarajan

Department of Chemical and Biomolecular Engineering, University of Houston, Houston, TX

BioTechniques 59:157-161 (September 2015) doi 10.2144/000114329

Keywords: plasmid recovery; library screening; toxicity

Supplementary material for this article is available at www.BioTechniques.com/article/114329.

Iterative screening of expressed protein libraries using fluorescence-activated cell sorting (FACS) typically involves culturing the pooled clones after each sort. In these experiments, if cell viability is compromised by the sort conditions and/or expression of the target protein(s), rescue PCR provides an alternative to culturing but requires re-cloning and can introduce amplification bias. We have optimized a simple protocol using commercially available reagents to directly recover plasmid DNA from sorted cells for subsequent transformation. We tested our protocol with 2 different screening systems in which <10% of sorted cells survive culturing and demonstrate that >60% of the sorted cell population was recovered.

Fluorescence-activated cell sorting (FACS) is a powerful high-throughput screening tool for protein engineering, provided a protein's function of interest can be coupled to and correlated with cell fluorescence. Screening of libraries for proteins with improved traits is typically accomplished by expression in a suitable microbial host and iterative sorting, pooling and culturing of the sorted clones (1,2). Unfortunately, enhancements in the protein's function of interest (e.g., enzymatic activity) and/or cell fluorescence (e.g., fluorescent protein expression) may compromise cell viability or growth (3,4). Furthermore, some whole-cell protein assays require the cells to be permeabilized and thus

rendered inviable (5). In such cases, rescue PCR is used to amplify the corresponding genes from the sorted cells. Apart from the subsequent re-cloning being time-consuming and laborious, PCR has the potential to alter the frequency distribution of DNA sequences by amplification bias and homologous recombination (6). Alternate methods that allow for the direct recovery of plasmid DNA from sorted cells with high efficiency enable considerable savings in time and cost while potentially preserving the genetic diversity and fitness improvements being sought.

Two key requisites of any method used for recovery during iterative library screening are: maximal retention of

genetic diversity for subsequent sorting and high yield of clones displaying the desired phenotype. Recently, a method to directly recover high-copy plasmid (~1000 copies/cell) from small numbers of cells (10^3 – 10^5) has been reported, but its effectiveness after FACS, especially for low- and medium-copy plasmids, is unknown (7). In the present study, we have quantitatively compared a direct plasmid recovery method with the more traditional method of culturing sorted cells following FACS, focusing on maximizing the yield of the subpopulations expressing the desired phenotype. We demonstrate enrichment and recovery using two disparate protein systems: *E. coli* MC1061 cells expressing the cytotoxic protease chymotrypsin and *E. coli* HF19 cells expressing variants of the AraC regulatory protein that activate expression of GFP (8,9). In both systems, direct plasmid recovery was found to be superior in comparison to recovery by culturing of sorted cells.

As a first step, we evaluated three different lysis conditions for the cells expressing chymotrypsin B: (i) 1× Bugbuster protein extraction reagent (Cat. No. 70921, EMD Millipore, Billerica, MA) with 1 mg/mL lysozyme (Sigma-Aldrich, St. Louis, MO) in 50 mM Tris-HCl, pH 7.5 (buffer 1); (ii) 100 mM EDTA and 1% Triton-X in 50 mM Tris-HCl, pH 7.5 (buffer 2); and (iii) 1× lysis buffer (buffer 3, contains 0.1 M NaOH, 0.7% SDS, 0.7 mM EDTA, and 1% isopropanol, pH 13.5) for plasmid miniprep (Cat. No. D4036-1-6; Zymo Research, Irvine, CA) (10). For evaluating buffers 1 and 2, *E. coli* MC1061 cells ($N_0 = 50,000$ in ~200 μ L) harboring plasmid pBAD_AChy_700 (p15A origin, 15–20 copies/cell) were sorted and incubated with an equal volume of 2× lysis buffer at 25°C for 15 min (8,11). Plasmids were purified using the Zymo DNA Clean & Concentrator kit. For buffer 3, an equal number of cells were sorted, and the protocol prescribed for Zippy Plasmid Miniprep Kit (Zymo Research) was followed using the DNA binding columns (Cat. No. D4004, Zymo Research). Plasmid DNA

METHOD SUMMARY

E. coli cells were sorted using flow cytometry, and plasmid DNA directly isolated from the sorted cells (10^5 – 10^6) was subsequently retransformed into competent cells. Our robust method yields high phenotype enrichment and provides an alternative to culturing sorted cells when screening or target protein expression compromises host cell viability.

was eluted with 10 μ L of Zippy elution buffer (10 mM Tris, 1 mM EDTA, adjusted to pH 8.0), and 2 μ L of eluent was used to transform electrocompetent *E. coli* MC1061 cells [transformation efficiency > 10^9 colony forming units (cfu)/ μ g] (12). The total number of transformants (N_t) was estimated by plating transformation dilutions onto LB-agar plates supplemented with chloramphenicol (LB-Cm plates). The recovery efficiency, defined as the ratio of total number of transformants obtained to total number of sorted cells used for plasmid recovery (N_t/N_o), was highest when using alkaline lysis buffer 3 ($44 \pm 24\%$) in comparison to buffer 1 ($16 \pm 13\%$) or buffer 2 ($4 \pm 2\%$). A step-by-step experimental protocol to perform DNA isolation using buffer 3 is available in the Supplementary Material.

A library of chymotrypsin B (chylib1) variants containing a C-terminal FLAG tag was constructed using overlap-extension PCR and cloned into vector pBAD_700 as described previously (8). Residues S189, G216, S217, and A226 play a critical role in determining the substrate specificity of the protease and were targeted for randomization (13). For protein expression, cells grown to an OD_{595} of 0.5 were induced with 0.2% arabinose at 37°C for 2 h. Cells expressing chylib1 were mixed with 1% of cells expressing wild-type (WT) chymotrypsin (ratio determined by OD_{595}) and incubated with 20 nM Chy-BQ7 (a positively charged peptide substrate that contains a chymotrypsin-sensitive linker sandwiched between a FRET pair) for 15 min at 25°C (8). The labeled population was analyzed and sorted using a BD FACSJazz cell sorter (BD Biosciences, San Jose, CA) at an event rate of ~ 7000 /second to isolate cells showing the highest (top 1%) fluorescence in the 530/40 nm channel (Figure 1A). In the absence of spiking, the frequency of cells expressing chylib1 in the sort gate was 0.3%. One thousand cells in the sort gate were collected in 100 μ L 2xYT medium and plated onto LB-Cm plates. The resulting cfu total was <50. Poor viability of sorted cells ($4 \pm 1\%$) is possibly due to both the proteolytic activity of chymotrypsin and labeling with peptide substrate in 1% sucrose.

To assess the effectiveness of our plasmid recovery method for sorted cells with poor viability, 50,000 cells in the sort gate were collected in a microcentrifuge

tube, and plasmid DNA was isolated using the buffer 3 protocol. Subsequent to transformation into *E. coli* MC1061 cells, the recovery efficiency (N_t/N_o) was estimated (by plating) at $63 \pm 20\%$. Next, transformants recovered with 1 mL SOC medium (20 g/L tryptone, 5 g/L yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$, 10 mM $MgCl_2$, and 20 mM glucose) (plasmid recovery method) or 50,000 sorted cells collected in 250 μ L 2xYT (culturing method) were directly grown in 100 mL 2xYT medium supplemented with 0.5% glucose and 25 μ g/mL chloramphenicol at 37°C for 10 h and then used to seed a subculture. After 2 h of induction with 0.2% arabinose, cells were labeled with Chy-BQ7 and analyzed by flow cytometry. The presence of full-length chymotrypsin variants on the bacterial surface was characterized by incubating induced cells with 40 nM anti-FLAG R-Phycoerythrin (ProZyme Inc., Hayward, CA) for 30 min at 25°C. The frequency of subpopulation in the sort gate increased by 35-fold in comparison to the parent population when plasmid recovery was employed (Figure 1A). On the other hand, the culturing method lost a majority of the cells expressing full-length chymotrypsin variants after just one round of sorting (Figure 1B).

We next performed similar experiments with our AraC-based biosensor system in which cells expressing variant AraC-TAL activate GFPuv expression upon sensing triacetic acid lactone (TAL), while cells expressing wild-type

AraC do not. *E. coli* HF19 cells harboring pPCC442 (P_{BAD} -gfpuv, modified RSF1030 origin, ~ 200 copies/cell) and pPCC423 (P_{tac} -araC, pBR322 Δ ROP origin, 30–60 copies/cell) or pPCC1202 (P_{tac} -araC-TAL, pBR322 Δ ROP origin) were induced with 5 mM TAL for 6 h at 37°C (9,14–16). Cells expressing AraC were spiked with cells expressing AraC-TAL (final composition of AraC-TAL = 10% as determined by OD_{595}) and sorted for cells showing the highest (top 2.5%) GFPuv fluorescence (Figure 2A). Sorted cells ($N_o = 500,000$) were collected in 500 μ L SOC, and the resulting 2 mL of collected sample was centrifuged for 10 min at $17,900 \times g$. Decanting the supernatant left behind a residual volume ($\sim 200 \mu$ L) of concentrated cells for DNA isolation using the plasmid recovery method (buffer 3 protocol). Eluted plasmid was transformed into HF19 cells harboring pPCC442 with a recovery efficiency of $106 \pm 52\%$. An aliquot of sorted cells was also directly plated to estimate viability. It is to be noted that the AraC system ($6 \pm 4\%$) had similar viability (based on cfu) to the chymotrypsin system ($4 \pm 1\%$). Transformants (plasmid recovery method) or sorted cells (culturing method) were directly grown in LB-Cm medium supplemented with 50 μ g/mL apramycin for 10 h at 37°C and used for inoculating fresh 10 mL LB-Cm culture supplemented with 50 μ g/mL apramycin, 100 μ M IPTG, 50 mM TES, and 1% glycerol to OD_{595} 0.2. Cells were induced with 5 mM TAL for 6 h at 37°C (OD_{595} 10) and washed once

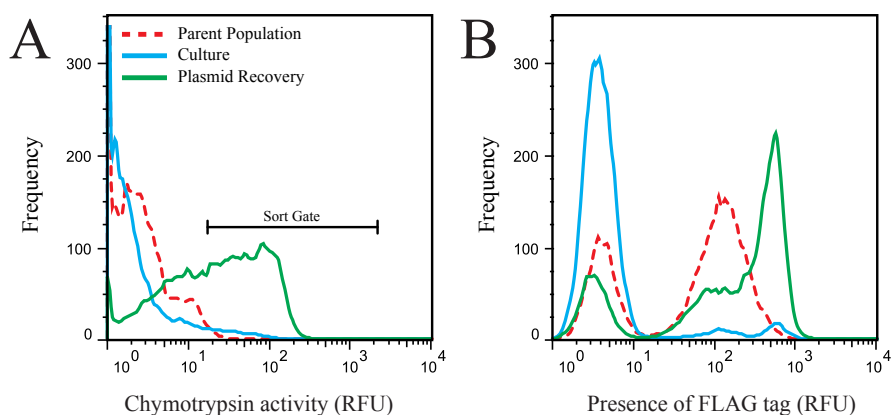
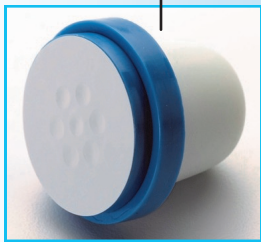


Figure 1. Comparison of phenotype enrichment by plasmid recovery and culturing methods with the chymotrypsin system. *E. coli* MC1061 cells expressing a chymotrypsin B library (chylib1) were spiked with 1% cells expressing wild-type chymotrypsin and sorted to collect the top 1% (determined by fluorescence) of the parent population (sort gate marked). Shown are chymotrypsin activity (A) and presence of FLAG tag at the C terminus of full-length chymotrypsin (B) of the parent population prior to sorting (red, dashed) and sorted populations recovered using plasmid recovery (green) and culturing (blue) methods measured using flow cytometry.

TAKE ACTION AGAINST CONTAMINATION WITH QMI® SAFE-SEPTUM

Sample, Inoculate Or Add Nutrients
To Your Bioreactor Aseptically !



The QMI Safe-Septum is:

- Aseptic
- Pressure & Temperature Safe
- Pre-Sterilized
- Easy To Retrofit
- Validated



651-501-2337

Email: info@qmisystems.com

www.qmisystems.com

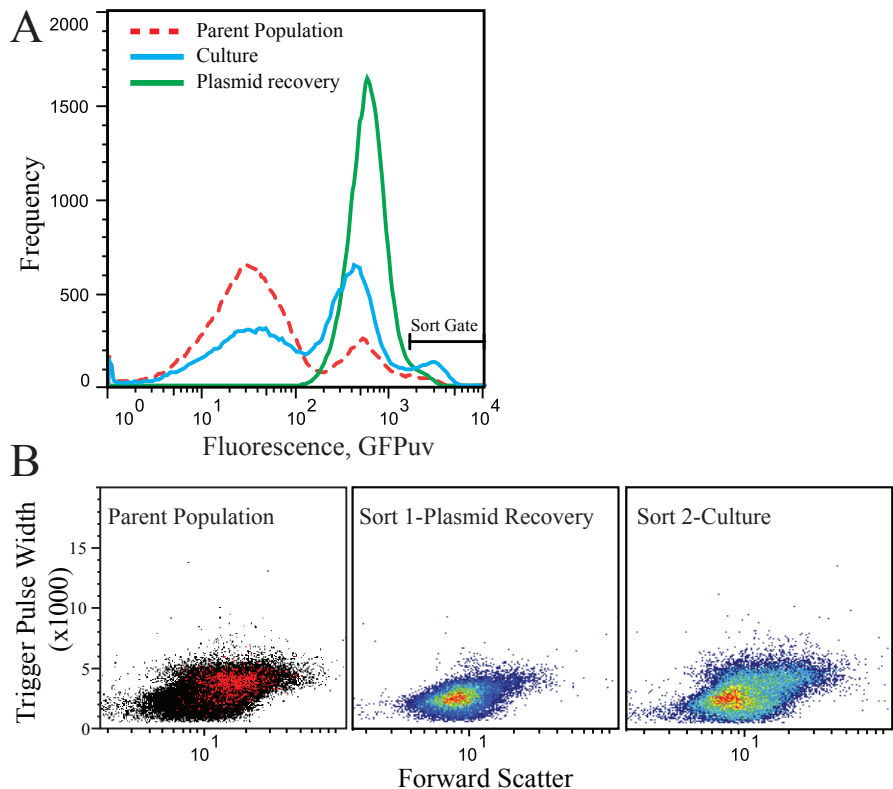


Figure 2. Comparison of phenotype retained by plasmid recovery and culturing methods with the AraC biosensor system. (A) *E. coli* HF19 cells expressing wild-type AraC from plasmid pPCC423 were spiked with 10% cells expressing AraC-TAL. Cells were then sorted to collect the top 2.5% (determined by fluorescence) of the parent population (sort gate marked). Histograms represent the frequency of cells showing fluorescence in response to 5 mM TAL prior to sorting (red, dashed) and sorted populations recovered using plasmid recovery (green) and culturing (blue) methods. (B) Scatter plots representing the size of measured events, as determined by forward scatter and trigger pulse width. Plots show all events of the parent population (black) and the sorted subset in the sort gate (red). Representative density plots of populations recovered by plasmid recovery and culturing methods show frequencies of size distribution of populations.

with an equal volume of PBS for flow cytometry.

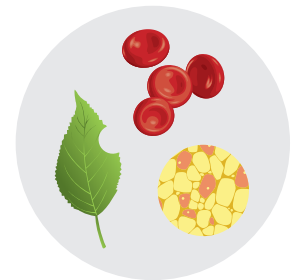
Using the geometric mean (GM) of clonal AraC-TAL fluorescence response to 5 mM TAL as the reference, a significant reduction ($P < 0.05$) in GFPuv fluorescence signal ($46 \pm 18\%$ of clonal AraC-TAL GM) was observed with the culturing method in comparison to the plasmid recovery method ($89 \pm 18\%$) (Figure 2A). Surprisingly, the culture method led to 43% of the population being unresponsive to both 5 mM TAL and 100 μ M L-arabinose. Upon analysis of forward scatter-trigger pulse width plots, we observed that a majority of the sorted cells appeared larger (Figure 2B). While the size of the cell can be linked to its growth rate in single-celled organisms such as *E. coli*, cell size also plays a fundamental role in fitness, both directly and indirectly (17,18). Since FACS-based screening aims to select for the cells with the highest fluorescence, it is important

to recognize that, given the complex relationships between cell size, fitness and resistance, gene expression, and plasticity (19, 20), it is possible to select for cells displaying high fluorescence that is not directly related to genetic variations in the protein of interest but rather due to variation in the host cell itself. While a more detailed mechanistic investigation needs to be undertaken to determine the relationships between high fluorescence, cell size, and biphasic response to ligands as demonstrated in our model system, we note that the plasmid recovery method, because of retransformation into fresh cells after every sort, minimizes any advantages conferred by the host cell during culturing that are unrelated to genetic variations in the protein of interest.

In order to quantify the genetic diversity of gene variants obtained by plasmid recovery, we have undertaken high-throughput sequencing analysis

Introducing KAPA PROBE FORCE

Direct qPCR analysis
without sample purification



Direct analysis from crude extracts



Streamlined workflows



Superior inhibitor tolerance



Multiplex compatibility

of AraC libraries being screened for responsiveness to TAL. In comparison to recovery of sorted cells by culturing, the number of unique AraC variants identified after sorting a library using the plasmid recovery method was 5–6 fold higher (manuscript in preparation).

In summary, we have developed a technique for the direct isolation of medium to high copy-number plasmid DNA from FACS-sorted cells with recovery efficiency sufficiently high for comprehensive screening of large, diverse libraries. This simple method for phenotype enrichment is well-suited as an alternative to culturing sorted cells or genotype recovery by PCR, especially when protein expression compromises host cell growth or viability.

Author contributions

B.R and C.S.F performed the experiments for the chymotrypsin system and the AraC system, respectively. All authors designed the research and wrote the manuscript.

Acknowledgments

This publication was supported by the Welch Foundation (E1774), NIH R01 (CA174385), CPRIT (RP130570), NSF (CBET 1159398); and MRA Stewart-Rahr Young Investigator Award to NV. This paper is subject to NIH public access policy.

Competing interests

The authors declare no competing interests.

References

1. **Yoo, T.H., M. Pogson, B.L. Iverson, and G. Georgiou.** 2012. Directed evolution of highly selective proteases by using a novel FACS-based screen that capitalizes on the p53 regulator MDM2. *ChemBioChem* 13:649–653.
2. **Sjostrom, S.L., Y. Bai, M. Huang, Z. Liu, J. Nielsen, H.N. Joensson, and H. Andersson Svahn.** 2014. High-throughput screening for industrial enzyme production hosts by droplet microfluidics. *Lab Chip* 14:806–813.
3. **Dumon-Seignovert, L., G. Cariot, and L. Vuillard.** 2004. The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr. Purif.* 37:203–206.
4. **Tu, R., R. Martinez, R. Prodanovic, M. Klein, and U. Schwaneberg.** 2011. A flow cytometry-based screening system for directed evolution of proteases. *J. Biomol. Screen.* 16:285–294.
5. **Harvey, B.R., G. Georgiou, A. Hayhurst, K.J. Jeong, B.L. Iverson, and G.K. Rogers.** 2004. Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from *Escherichia coli*-expressed libraries. *Proc. Natl. Acad. Sci. USA* 101:9193–9198.
6. **Kalle, E., M. Kubista, and C. Rensing.** 2014. Multi-template polymerase chain reaction. *Biomol. Detect. Quantif.* 2:11–29.
7. **Kintses, B., C. Hein, M.F. Mohamed, M. Fischlechner, F. Courtois, C. Lainé, and F. Hoffelder.** 2012. Picoliter cell lysate assays in microfluidic droplet compartments for directed enzyme evolution. *Chem. Biol.* 19:1001–1009.
8. **Ramesh, B., V.G. Sendra, P.C. Cirino, and N. Varadarajan.** 2012. Single-cell characterization of autotransporter-mediated *Escherichia coli* surface display of disulfide bond-containing proteins. *J. Biol. Chem.* 287:38580–38589.
9. **Tang, S.-Y., S. Qian, O. Akinterinwa, C.S. Frei, J.A. Gredell, and P.C. Cirino.** 2013. Screening for enhanced triacetic acid lactone production by recombinant *Escherichia coli* expressing a designed triacetic acid lactone reporter. *J. Am. Chem. Soc.* 135:10099–10103.
10. **Jia, X.Y.** 2011. Plasmid DNA Isolation. US Patent 20110117641A1.
11. **Chang, A.C. and S.N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.* 134:1141–1156.
12. **Varadarajan, N., J.R. Cantor, G. Georgiou, and B.L. Iverson.** 2009. Construction and flow cytometric screening of targeted enzyme libraries. *Nat. Protoc.* 4:893–901.
13. **Jelinek, B., G. Katona, K. Fodor, I. Venekei, and L. Gráf.** 2008. The crystal structure of a trypsin-like mutant chymotrypsin: the role of position 226 in the activity and specificity of S189D chymotrypsin. *Protein J.* 27:79–87.
14. **Tang, S.-Y., H. Fazelinia, and P.C. Cirino.** 2008. AraC regulatory protein mutants with altered effector specificity. *J. Am. Chem. Soc.* 130:5267–5271.
15. **Phillips, G.J., S.K. Park, and D. Huber.** 2000. High copy number plasmids compatible with commonly used cloning vectors. *Biotechniques* 28:400–402, 404, 406 passim.
16. **Twigg, A.J. and D. Sherratt.** 1980. Trans-complementable copy-number mutants of plasmid ColE1. *Nature* 283:216–218.
17. **Hill, N.S., P.J. Buske, Y. Shi, and P.A. Levin.** 2013. A moonlighting enzyme links *Escherichia coli* cell size with central metabolism. *PLoS Genet.* 9:e1003663.
18. **Yoshida, M., S. Tsuru, N. Hirata, S. Seno, H. Matsuda, B.-W. Ying, and T. Yomo.** 2014. Directed evolution of cell size in *Escherichia coli*. *BMC Evol. Biol.* 14:257.
19. **St-Pierre, F. and D. Endy.** 2008. Determination of cell fate selection during phage lambda infection. *Proc. Natl. Acad. Sci. USA* 105:20705–20710.
20. **Lehner, B.** 2010. Conflict between noise and plasticity in yeast. *PLoS Genet.* 6:e1001185.

Received 09 April 2015; accepted 02 July 2015.

Address correspondence to Navin Varadarajan, Department of Chemical and Biomolecular Engineering, University of Houston, Houston, Texas, TX. E-mail: nvaradar@central.uh.edu

To purchase reprints of this article, contact: biotechniques@fosterprinting.com