

Short Technical Reports

- tension and mantle closure in *Xenopus laevis* gastrulation: combined roles for integrin $\alpha(5)\beta(1)$, fibronectin, and tissue geometry. Dev. Biol. 242:109-129.
6. **Drake, C.J., L.A. Davis, and C.D. Little.** 1992. Antibodies to β 1-integrins cause alterations of aortic vasculogenesis, *in vivo*. Dev. Dyn. 193:83-91.
7. **Hamburger, V. and H.L. Hamilton.** 1992. A series of normal stages in the development of the chick embryo. Dev. Dyn. 195:231-272.
8. **Isogai, S., M. Horiguchi, and B.M. Weinstein.** 2001. The vascular anatomy of the developing zebrafish: an atlas of embryonic and early larval development. Dev. Biol. 230:278-301.
9. **Kulesa, P.M. and S.E. Fraser.** 1999. Confocal imaging of living cells in intact embryos. Methods Mol. Biol. 122:205-222.
10. **Kulesa, P.M. and S.E. Fraser.** 2000. *In ovo* time-lapse analysis of chick hindbrain neural crest cell migration shows cell interactions during migration to the branchial arches. Development 127:1161-1172.
11. **Little, C.D. and C.J. Drake.** 2000. Whole-mount immunolabeling of embryos by microinjection. Increased detection levels of extracellular and cell surface epitopes. Methods Mol. Biol. 135:183-189.
12. **New, D.A.T.** 1955. A new technique for the cultivation of the chick embryo *in vitro*. J. Embryol. Exp. Morph. 3:326-331.
13. **Schnabel, R., H. Hutter, D. Moerman, and H. Schnabel.** 1997. Assessing normal embryogenesis in *Caenorhabditis elegans* using a 4D microscope: variability of development and regional specification. Dev. Biol. 184:234-265.
14. **Stockdale, F.E., W. Nikovits, Jr., and B. Christ.** 2000. Molecular and cellular biology of avian somite development. Dev. Dyn. 219:304-321.
15. **Thomas, C., P. DeVries, J. Hardin, and J. White.** 1996. Four-dimensional imaging: computer visualization of 3D movements in living specimens. Science 273:603-607.

We would like to thank Adrienn Gabor for her exceptional assistance with software development and the design and construction of the incubator. This work was supported by National Institutes of Health grant nos. 5 F32 HL68457-02 to P.A.R. and 5P01 HL52813-08 to C.D.L., The G. Harold & Leila Y. Mathers Charitable Foundation (to C.D.L.), the Heartland Affiliate of the American Heart Association (Beginning Grant-in Aid 0160369Z to B.J.R.), and NATO-NSF DGE-0075179 to A.C. Address correspondence to Dr. Charles D. Little, Department of Anatomy and Cell Biology, University of Kansas Medical Center, 3901 Rainbow Blvd., Kansas City, KS 66160, USA. e-mail: clittle@kumc.edu

Received 30 August 2002; accepted 5 November 2002.

Paul A. Rupp¹, Brenda J. Rongish¹, Andras Czirok^{1,2}, and Charles D. Little¹

¹University of Kansas Medical Center

Kansas City, KS, USA

²Eotvos University
Budapest, Hungary

ferent crossover sites were found; most occurred when there was a stretch of eight or more identical base pairs in both sequences, and many were concentrated in the regions important for studying ligand binding and transactivation. This method should prove to be useful for generating chimeric gene products from parent templates that share relatively low sequence identity.

INTRODUCTION

Chimeric gene products can be produced from two homologous parent genes using restriction enzyme digestion and fragment ligation (15,19). However, the number and position of corresponding restriction sites are often very limited, making this strategy less general. Restriction sites can be engineered into the genes to facilitate gene fragment swapping, but this quickly becomes tedious when numerous chimeras are needed. Also, it is not always possible to introduce a restriction site at a desired position without changing the coded amino acids. Blunt-end ligation of PCR-generated fragments can be used to generate chimeras at any position, but this becomes cumbersome for more than one crossover and requires the rational design of crossover positions.

DNA shuffling involves *in vitro* or *in vivo* recombination methods for generating chimeric genes that rely on short stretches of identical DNA sequence in the homologous parent genes. DNA shuffling for chimeragenesis in gene families (13,17,18) has been widely used in directed evolution, examples of which include changing enzyme substrate specificity (24), improving enzyme thermostability (3), distinguishing functional and nonfunctional mutations (25), and probing protein structure-function relationships (4). The original Stemmer method for DNA shuffling uses DNase I fragmentation and gene reassembly (17); however, a variety of modified methods has been developed to improve the efficiency of *in vitro* recombination: staggered extension (5,26), shuffling with restriction enzyme-cleaved DNA fragments (6), shuffling using ssDNA (7,23), random chimeragenesis on transient templates (RACHITT) (1), the homology independent ITCHY method (12), and the

Short Technical Reports

SHIPREC method of Sieber et al. (16). These modified methods are often technically more challenging and require additional steps; some of them can only be applied to highly homologous genes. Some DNA shuffling methods have higher tolerance for parent genes with relatively low homology, but generally the efficiency of chimera generation decreases significantly as nucleotide sequence identities decline, and special approaches are needed.

We were motivated to generate a series of estrogen receptor (ER) chimeric constructs using DNA shuffling techniques to facilitate the study of the molecular basis of the subtype selectivity of some novel ER ligands. Since our targets, which are the ligand binding domains from ER α and ER β share only 63% identity at the DNA level, they would be considered non-ideal substrates for application of the DNA shuffling technique. Here we describe the development of a simple but efficient and generally applicable DNA shuffling method to generate ER chimera constructs in a yeast system, with the sequences of their ligand binding domains derived from the two parent sequences in ER α and ER β .

MATERIALS AND METHODS

Yeast Strain and Plasmid Constructs

The yeast reporter strain BJ-ECZ (*ura3*, *trp1*, *leu2*, *pep4*, *prc1*, *prb1*) containing nine tandem copies of an estrogen responsive *LacZ* reporter construct integrated into the *ura3* locus was generated in our laboratory previously (22). To construct the human ER β yeast expression vector, a 1.5-kb *Bam*HI fragment of human ER β cDNA from pNGV1-ER β (10) was inserted into YE ρ ER α (22) by replacing the 1.8-kb *Bam*HI fragment containing human ER α cDNA.

DNA Shuffling

The strategy used is illustrated in Figure 1. It employs DNA shuffling (21) with several modifications to optimize the generation of crossovers and production of chimeras. Primers ER α A289f (5'-GCCAACCTTGGC-

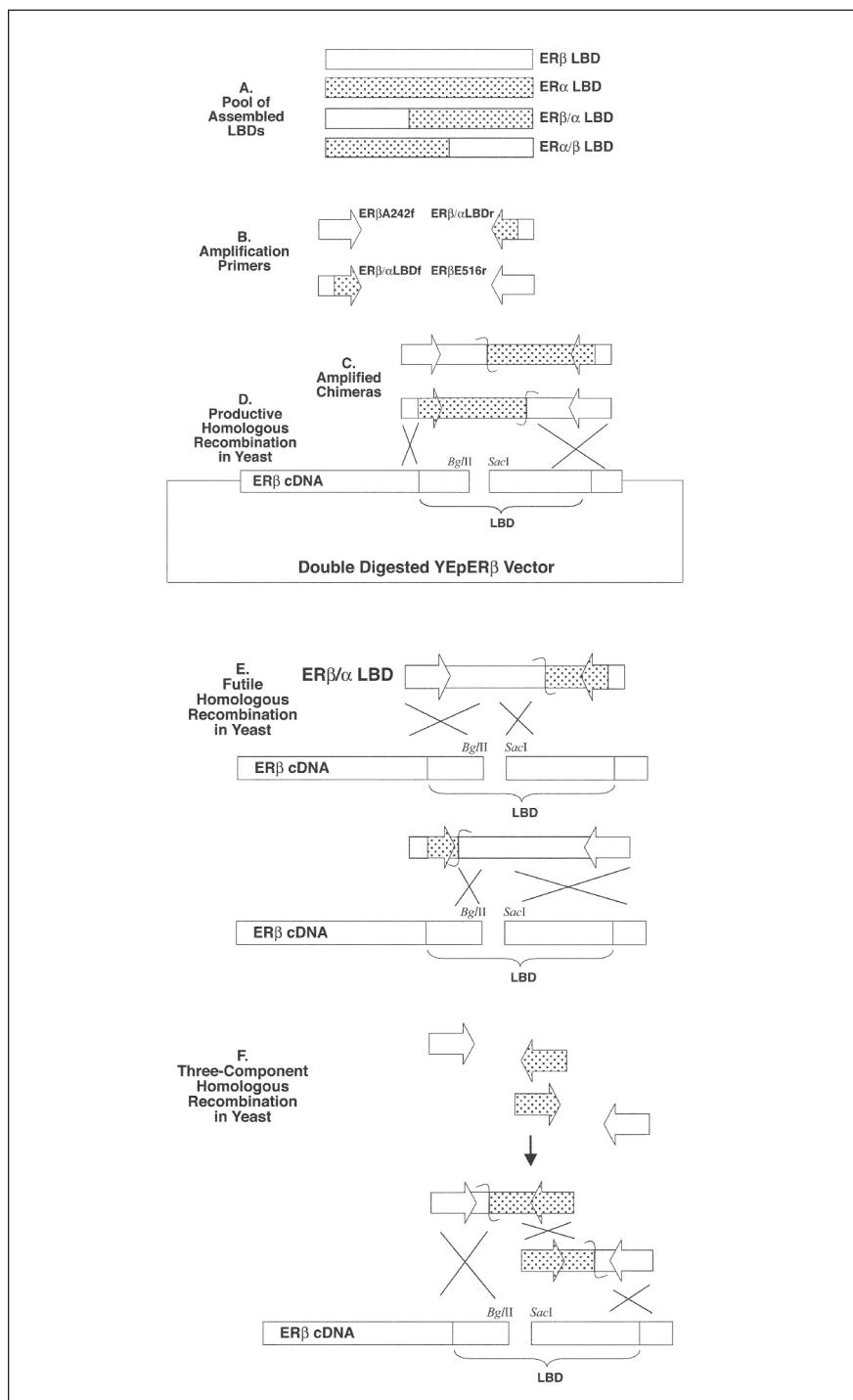


Figure 1. DNA shuffling /homologous recombination system. (A) The assembled ER ligand binding domain (LBD) pool after DNA shuffling. The full-length assembled ligand binding domains are mostly non-chimeric, either ER α or ER β , with a small portion of chimeric ER α/β and ER β/α . (B) Use of primer pairs ER α A242f and ER β/α LBDr or ER β/α LBDf and ER β E516r in the PCR amplification from assembled ligand binding domain products. (C) Only chimeric sequences can be successfully amplified using the above primer pairs. (D) In vivo homologous recombination between the chimeric ligand binding domain sequences and digested YEpER β vector, based on the identical sequences at both ends of the ligand binding domains and sequences in the vector, as indicated by the crossing lines. (E) In vivo homologous recombination within the ligand binding domain sequence results in wild-type products. (F) Two pairs of primers used for PCR generate two partial but overlapping chimeric ligand binding domain fragments; those fragments undergo homologous recombination with a linearized vector to form an ER expression vector carrying two crossovers in the ligand binding domain.

Short Technical Reports

CAAGCCC-3') and ER α S566r (5'-GCTTGGTCCGTCTCCTCCAC-3') were used to amplify the cDNA of the ER α ligand binding domain from a plasmid template pCMV5-ER α . To amplify ER β ligand binding domain cDNA, primers ER β A242f (5'-GCCG-GCAAGGCCAAGAGAAGT-3') and ER β E516r (5'-CTCTGCCGGGCTG-CACTC-3') were used in a PCR using pCMV5-ER β (9) as a template. After gel purification using the QIAquick™ Gel Extraction Kit (Qiagen, Valencia, CA, USA), 2 μ g ER α and ER β ligand binding domain were mixed, and 2.5 μ L 1 M Tris-HCl (pH 7.5) and 2.5 μ L 200 mM MnCl₂ were added. The volume was brought to 49 μ L with distilled water. The mixture was equilibrated at 15°C for 5 min. DNase I (1 U/ μ L; Invitrogen, Carlsbad, CA, USA) was freshly diluted 1:10 in distilled water, and 1 μ L of it was added to the reaction. The digestion was performed at 15°C for 90 s and was stopped by adding 5 μ L ice-cold stop buffer containing 50 mM EDTA and 30% (v/v) glycerol. The digested products were separated by electrophoresis in a 2% (w/v) agarose gel. The DNA fragments of about 100 bp were recovered from the gel. About 0.5 μ g of them was combined with 0.5 μ L (2.5 U) *Pfu* DNA polymerase (Stratagene, La Jolla, CA, USA), 3 μ L 10 \times *Pfu* DNA polymerase reaction buffer, 1.5 μ L 2.5 mM dNTP (Applied Biosystems, Foster City, CA, USA) in a total volume of 30 μ L. Proofreading DNA polymerase was used so as to minimize the frequency of point mutations that could be introduced into the progeny (shuffled genes). These point mutations would complicate functional studies of the chimeric genes because some of the point mutations may affect protein functions. The reassembling reaction was performed in a PCR machine (PTC-100™; MJ Research, Waltham, MA, USA) with the following conditions: 60 s at 94°C, followed by 40 cycles of 30 s at 94°C, 1 min at 40°C and 1 min at 72°C. The recombinant products were amplified in a standard PCR (40 cycles of 45 s at 94°C, 30 s at 50°C, and 1 min at 72°C), using 1 μ L 1:10 diluted reassembled products as template, ER β /αLBDf (5'-GCCAAGAGAAGT-GCGGCCACGCCCGAAAGA-

AGAACAGCCTGGCC-3'; underlined sequences match ER α ligand binding domain cDNA, sequences in italics match ER β cDNA outside the ligand binding domain) and ER β E516r, or ER β A242f and ER β /αLBDr (5'-*GC-TGCACTCGGACCCCGTGATGGAGGACTTGGCGCATGTAGGCGGTG*-3') as primers. After gel purification of the properly sized products, they were transformed with the YE β vector fragment (*Bgl*II and *Sac*I digested) (100 ng PCR products with 25 ng linearized vector) into the yeast strain BJ-E CZ using a LiOAc-based method (2). Transformed cells were plated onto synthetic minimal medium lacking tryptophan (SD-Trp). After 2–3 days at 30°C, the individual transformants were transferred to approximately 3 mL liquid SD-Trp medium and incubated at 30°C for two days. The plasmids were rescued from yeast cells using Y-DER yeast DNA extraction reagent (Pierce Biotechnology, Rockford, IL, USA) and transformed into *E. coli* DH5 α cells. The chimeric sequences were checked by DNA sequencing performed from plasmids purified from *E. coli* cells using BigDye™ and an ABI PRISM® 3700 sequencer (Applied Biosystems).

RESULTS AND DISCUSSION

We applied the DNA shuffling technique to generate ERs containing chimeric ligand binding domain sequences from the two ER subtypes, ER α and ER β , to facilitate our studies of the molecular basis for the ER subtype-selectivity of novel, non-steroidal ligands. The nucleotide sequence identity between the two ER ligand binding domains is 63% when the alignment is based on the amino acid sequence alignment. Hence, our initial results indicated that most of the assembled products were wild-type sequences instead of the chimeric products in which we were interested, presumably because of the low percent sequence identity between the two parent sequences. Another obstacle we encountered was that there are no convenient restriction enzyme sites available at both ends of the ligand binding domain for introducing the chimeric sequences into the ER

expression vector. Therefore, we designed a strategy (Figure 1) to overcome both of these problems, increasing the efficiency of chimera generation by amplifying only the chimeric products from a pool of reassembled gene products and integrating the chimeric sequences into the yeast expression vector YE β through homologous recombination in yeast cells.

We first amplified the two ER ligand binding domain cDNAs with an approximately 40-bp extension outside the ligand binding domain at both ends (approximately 0.8 kb) by standard PCR methods (Figure 2A). The purified PCR products underwent DNase I partial digestion, followed by separation on a 2% agarose gel (Figure 2B). After we isolated approximately 100-bp fragments from the agarose gel, we assembled them in a PCR without adding additional primers. The size of the assembled products ranged from approximately 300 bp to well over 1 kb. Since the latter length is longer than the input ligand binding domain sequence, this is an indication of the complexity of the reaction mixture (Figure 2C). We also tried to use approximately 50-bp fragments for the assembling reaction, but we failed to get the spectrum of the assembled products as shown in Figure 2C, most likely because of the low sequence identity between the two parent genes.

After PCR amplification of this assembled product using primer pairs ER β /αLBDf and ER β E516r or ER β A242f and ER β /αLBDr (Figure 1B), a major product having the correct size of the full-length ligand binding domain was obtained (Figure 2D). Because we had designed the two amplification primers so that their 3' ends specifically match only one and the other of the two parent genes, respectively, the PCR products are guaranteed to be chimeric, and wild-type sequences are completely eliminated (Figure 1C). Taking advantage of the plasmid construction by homologous recombination in yeast (8), we integrated the chimeric ligand binding domain sequences into the yeast expression vector by co-transforming the purified products with *Bgl*II and *Sac*I double-digested YE β vector into the yeast

Short Technical Reports

test strain BJ-ECZ (Figure 1D). The homologous sequences between the ends of the products and the linearized vector are at least 40 bp long, which is enough to elicit efficient recombination in yeast (11,14).

By functional assays, we found that the majority of the yeast transformants responded to the hormone 17 β -estradiol in a β -galactosidase plate assay. By DNA sequencing, we found that about half of them had chimeric ligand binding domains with one crossover occurring at those sites indicated in Figure 3. A similar approach was used to generate ER α expression vectors containing chimeric ligand binding domains. Using this recombination-based method, we have also constructed ER expression vectors with their whole ligand binding domain swapped.

Throughout the approximately 0.75-kb ligand binding domain regions, 22 crossover positions have been found in the sequenced chimeric clones (Figure 3). Most of these sites are concentrated at both ends of the domain, corresponding to the regions having the highest sequence identity between the two ligand binding domains. The encoding residues from most of these sites are in the ligand binding pocket and are close to the ligand binding sites. These chimeric constructs have proved to be useful to us in studying the molecular basis of action of ER subtype-selective ligands through structure-function analyses of the activity of ER α - or ER β -specific ligands on ER β/α chimeric proteins (J. Sun Ph.D. thesis, University of Illinois, Urbana, 2002; Reference 20; and manuscripts in preparation). We have found that a stretch of eight or more base pairs that are identical in both the ER α and ER β sequences is sufficient to trigger an efficient recombination process in the shuffling reaction. In regions that share, overall, a very high DNA sequence similarity between the two ligand binding domains, as few as two identical base pairs sufficed to generate a crossover (Figure 3).

Parent sequence contamination is a significant problem associated with DNA shuffling. Here, in the PCR amplification step, we use primers that are specific to the two different ER subtype sequences so as to amplify only

chimeric sequences out of the pool of reassembled products. Nevertheless, during our screen for ER expression vectors that contain shuffled ligand binding domain sequences, we found that half of the clones were wild type. Although we used double enzyme digestion to excise a small piece fragment in the wild-type ligand binding domain cDNAs, so as to make the DNA pool free from the complete wild-type sequence, some recombination did occur between the remaining ligand binding domain cDNA sequences in the plasmid and the amplified shuffled ligand binding domain cDNAs, and this gave a significant background of wild-type sequence (Figure 1E). One approach to reducing this background might be to remove most of the ligand binding domain sequence from the parent expression plasmid.

Theoretically, all of the amplified products should carry an odd number of crossovers. In reality, we found only a single crossover for each chimeric sequence; sequences containing three crossovers must be very rare, if they exist at all. This may be largely due to the annealing of DNA fragments derived from the same parental genes (homoduplex formation), whose probability is much higher than that of heteroduplex formation. By using the strategy illustrated in Figure 1F, which is a three-component homologous recombination system in yeast, we were able to generate chimeric ligand binding domains carrying two crossovers.

The strategies described here for generating chimeric gene products from the two homologous ER subtype genes provide a quick and efficacious way to perform DNA shuffling on tar-

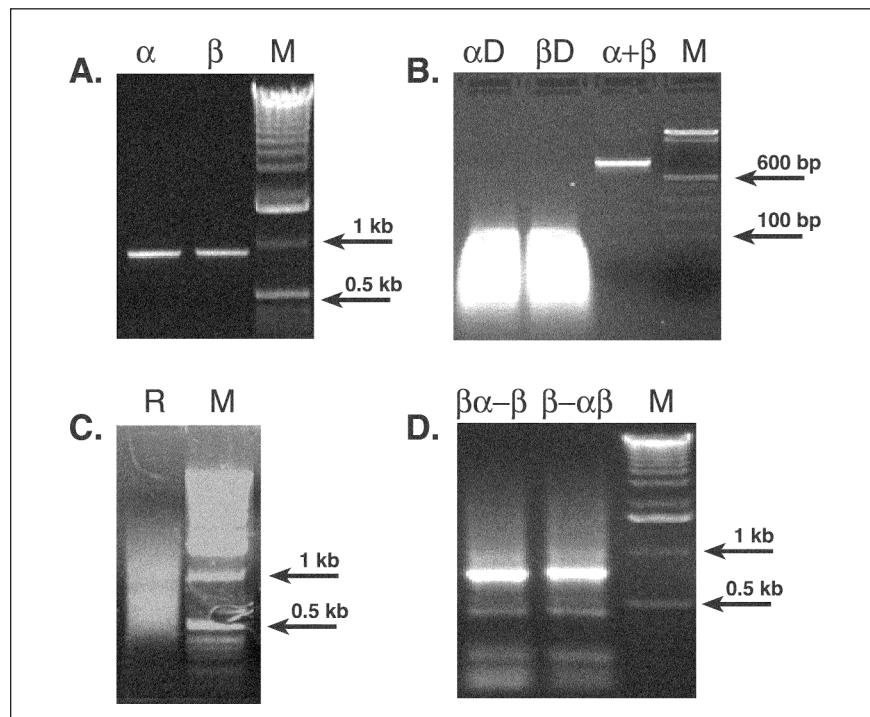


Figure 2. DNA shuffling between ER α and ER β ligand binding domains. (A) The ligand binding domain cDNAs from ER α and ER β were amplified with primer pairs ER α A289f and ER α S566r (lane α) or ER β A242f and ER β E516r (lane β), respectively. Lane M, 1-kb DNA ladder (Invitrogen). (B) The partial DNase I digestion of cDNAs of ER α ligand binding domain (lane α D) and ER β LBD (lane β D). Their original size is shown in Lane α + β for comparison. Lane M, 100-bp DNA ladder (Invitrogen). (C) Lane R, the products after assembly; lane M, 1-kb DNA ladder (Invitrogen). (D) The amplifications of chimeric products from the assembled products. Primers ER β /αLBDf and ER β E516r were used in the PCR to amplify chimeric ligand binding domains that have N-terminal sequences from ER β and C-terminal sequences from ER β (lane β - α - β). Primers ER β A242f and ER β /αLBDr were used in the PCR to amplify chimeric ligand binding domains that have N-terminal sequences from ER β and C-terminal sequences from ER α (lane β - α - β). Lane M, 1-kb DNA ladder (Invitrogen).

Short Technical Reports

gets with low sequence identity that would not be good candidates for DNA shuffling by standard methodologies. This approach should be applicable to other homologous genes, including those for other nuclear hormone receptor subtypes.

REFERENCES

- REFERENCES**

 1. Coco, W.M., W.E. Levinson, M.J. Crist, H.J. Hektor, A. Darzins, P.T. Pienkos, C.H. Squires, and D.J. Monticello. 2001. DNA shuffling method for generating highly recombinant genes and evolved enzymes. *Nat. Biotechnol.* 19:354-359.
 2. Gietz, R.D., R.H. Schiestl, A.R. Willems, and R.A. Woods. 1995. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* 11:355-360.
 3. Giver, L., A. Gershenson, P.O. Freskgard, and F.H. Arnold. 1998. Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. USA* 95:12809-12813.

Figure 3. The results of DNA shuffling from cDNAs of the ER α and ER β ligand binding domains. DNA sequence alignment between the ER α and ER β ligand binding domains is shown based on their amino acid sequence alignment. The sites of crossover are highlighted.

Short Technical Reports

4. **Hansson, L.O. and B. Mannervik.** 2000. Use of chimeras generated by DNA shuffling: Probing structure-function relationships among glutathione transferases, p. 463-477. In Applications of Chimeric Genes and Hybrid Proteins, Pt C.
5. **Judo, M.S.B., A.B. Wedel, and C. Wilson.** 1998. Stimulation and suppression of PCR-mediated recombination. *Nucleic Acids Res.* 26:1819-1825.
6. **Kikuchi, M., K. Ohnishi, and S. Harayama.** 1999. Novel family shuffling methods for the in vitro evolution of enzymes. *Gene* 236:159-167.
7. **Kikuchi, M., K. Ohnishi, and S. Harayama.** 2000. An effective family shuffling method using single-stranded DNA. *Gene* 243:133-137.
8. **Ma, H., S. Kunes, P.J. Schatz, and D. Botstein.** 1987. Plasmid construction by homologous recombination in yeast. *Gene* 58:201-216.
9. **McInerney, E.M., K.E. Weis, J. Sun, S. Mosselman, and B.S. Katzenellenbogen.** 1998. Transcription activation by the human estrogen receptor subtype β (ER β) studied with ER β and ER α receptor chimeras. *Endocrinology* 139:4513-4522.
10. **Mosselman, S., J. Polman, and R. Dijkema.** 1996. ER β : identification and characterization of a novel human estrogen receptor. *FEBS Lett.* 392:49-53.
11. **Oldenburg, K.R., K.T. Vo, S. Michaelis, and C. Paddon.** 1997. Recombination-mediated PCR-directed plasmid construction in vivo in yeast. *Nucleic Acids Res.* 25:451-452.
12. **Ostermeier, M., J.H. Shim, and S.J. Benkovic.** 1999. A combinatorial approach to hybrid enzymes independent of DNA homology. *Nat. Biotechnol.* 17:1205-1209.
13. **Pompon, D. and A. Nicolas.** 1989. Protein engineering by cDNA recombination in yeasts: shuffling of mammalian cytochrome P-450 functions. *Gene* 83:15-24.
14. **Raymond, C.K., T.A. Pownder, and S.L. Sexson.** 1999. General method for plasmid construction using homologous recombination. *BioTechniques* 26:134-141.
15. **Rehberg, E., B. Kelder, E.G. Hoal, and S. Pestka.** 1982. Specific molecular activities of recombinant and hybrid leukocyte interferons. *J. Biol. Chem.* 257:11497-11502.
16. **Sieber, V., C.A. Martinez, and F.H. Arnold.** 2001. Libraries of hybrid proteins from distantly related sequences. *Nat. Biotechnol.* 19:456-460.
17. **Stemmer, W.P.C.** 1994. DNA shuffling by random fragmentation and reassembly—in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* 91:10747-10751.
18. **Stemmer, W.P.C.** 1994. Rapid evolution of a protein in-vitro by DNA shuffling. *Nature* 370:389-391.
19. **Streuli, M., A. Hall, W. Boll, W.E. Stewart, II, S. Nagata, and C. Weissmann.** 1981. Target cell specificity of two species of human interferon- α produced in *Escherichia coli* and of hybrid molecules derived from them. *Proc. Natl. Acad. Sci. USA* 78:2848-2852.
20. **Sun, J., J. Baudry, J.A. Katzenellenbogen, and B.S. Katzenellenbogen.** Molecular basis for the subtype discrimination of the estrogen receptor β selective ligand, diarylpropionitrile. *Mol. Endocrinol.* doi:10.1210/me.2002-0341.
21. **Volkov, A.A. and F.H. Arnold.** 2000. Methods for in vitro DNA recombination and random chimeragenesis. *Methods Enzymol.* 328:447-456.
22. **Wrenn, C.K. and B.S. Katzenellenbogen.** 1993. Structure-function analysis of the hormone-binding domain of the human estrogen-receptor by region-specific mutagenesis and phenotypic screening in yeast. *J. Biol. Chem.* 268:24089-24098.
23. **Zha, W., T. Zhu, and H. Zhao.** Family shuffling with single-stranded DNA. In F.H. Arnold and G. Georgiou (Eds.), *Methods in Molecular Biology*, Directed Evolution Library Creation: Methods and Protocols. Humana Press, Totowa, NJ. (In Press.)
24. **Zhang, J.H., G. Dawes, and W.P.C. Stemmer.** 1997. Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc. Natl. Acad. Sci. USA* 94:4504-4509.
25. **Zhao, H. and F.H. Arnold.** 1997. Functional and nonfunctional mutations distinguished by random recombination of homologous genes. *Proc. Natl. Acad. Sci. USA* 94:7997-8000.
26. **Zhao, H., L. Giver, Z.X. Shao, J.A. Affholter, and F.H. Arnold.** 1998. Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* 16:258-261.

This work was supported by National Institutes of Health grant nos. PHS 5R01 CA18119 (B.S.K.) and PHS 5R37 DK15556 (J.A.K.), and American Cancer Society grant no. IRG99224001 (H.Z.). Address correspondence to Dr. Benita S. Katzenellenbogen, University of Illinois, Department of Molecular and Integrative Physiology, 524 Burrill Hall, 407 South Goodwin Avenue, Urbana, IL 61801-3704, USA. e-mail: katzenel@life.uiuc.edu

Received 26 September 2002; accepted 19 November 2002.

Jun Sun, John A. Katzenellenbogen, Huimin Zhao, and Benita S. Katzenellenbogen
*University of Illinois
Urbana, IL, USA*