Biotechnology Advances xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Biotechnology Advances

journal homepage: www.elsevier.com/locate/biotechadv

Research review paper

Polishing the craft of genetic diversity creation in directed evolution ☆

Kang Lan Tee [a], Tuck Seng Wong [b,*]

[a] Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester M1 7DN, England, United Kingdom
[b] ChELSI Institute, Department of Chemical and Biological Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, England, United Kingdom

ARTICLE INFO

Article history:
Received 28 May 2013
Received in revised form 31 August 2013
Accepted 31 August 2013
Available online xxxx

Keywords:
Directed evolution
Genetic diversity
Random mutagenesis
Site-directed mutagenesis
DNA recombination
Synthetic biology
Genome engineering
Metabolic engineering

ABSTRACT

Genetic diversity creation is a core technology in directed evolution where a high quality mutant library is crucial to its success. Owing to its importance, the technology in genetic diversity creation has seen rapid development over the years and its application has diversified into other fields of scientific research. The advances in molecular cloning and mutagenesis since 2008 were reviewed. Specifically, new cloning techniques were classified based on their principles of complementary overhangs, homologous sequences, overlapping PCR and megaprimers and the advantages, drawbacks and performances of these methods were highlighted. New mutagenesis methods developed for random mutagenesis, focused mutagenesis and DNA recombination were surveyed. The technical requirements of these methods and the mutational spectra were compared and discussed with references to commonly used techniques. The trends of mutant library preparation were summarised. Challenges in genetic diversity creation were discussed with emphases on creating "smart" libraries, controlling the mutagenesis spectrum and specific challenges in each group of mutagenesis methods. An outline of the wider applications of genetic diversity creation includes genome engineering, viral evolution, metagenomics and a study of protein functions. The review ends with an outlook for genetic diversity creation and the prospective developments that can have future impact in this field.

Contents

Abbreviations: 8-oxo-dGTP, 8-Oxo-2′-deoxyguanosine-5′-triphosphate; AAV, Adeno-associated virus; ABI-REC, Asymmetric Bridge PCR with Intramolecular Homologous Recombination; bp, Base pairs; CAST, Combinatorial Active-Site Saturation Test; CPEC, Circular Polymerase Extension Cloning; CRP, cAMP receptor protein; $dA^{TESE}TP$, 7-Deaza-7-(triethylsilylethynyl)deoxyadenosine triphosphate; DGRs, Diversity-generating retroelements; dITP, 2′-Deoxyinosine 5′-triphosphate; DNA, Deoxyribonucleic acid; dNTP, Deoxyribonucleotide triphosphate; dPTP, 2′-Deoxy-P-nucleoside-5′-triphosphate; dsDNA, Double-stranded DNA; DuARCheM, Dual Approach to Random Chemical Mutagenesis; EMP, Exponential Megapriming PCR; EMS, Ethyl methane sulfonate; enoyl ACP reductase, Enoyl-acyl carrier protein reductase; epPCR, Error-prone polymerase chain reaction; epRCA, Error-prone rolling circle amplification; GOI, Gene of interest; GST, Glutathione-S-transferase; InDel, Insertion and deletion; ITCHY, Incremental Truncation for the Creation of Hybrid Enzymes; KF, Klenow fragment; MAP, Mutagenesis Assistant Program; MEGAWHOP, Megaprimer PCR of Whole Plasmid; MGS, Mutation Generation System; MLF-SDM, Megaprimed and Ligase-Free PCR-based Method for Site-Directed Mutagenesis; NEB, New England Biolabs; NiDE, Nicking DNA Endonuclease; NRR, Non-homologous random recombination; nt, Nucleotide; OLTA, OverLap extension PCR and TA cloning; OSCARR, One-pot Simple Methodology for Cassette Randomization and Recombination; pBpa, *p*-Benzoylphenylalanine; PCR, Polymerase chain reaction; PERMUTE, PERMutation Using Transposase Engineering; PFLF-MSDM, Phosphorylation-Free and Ligase-Free PCR-based Method for Multiple SDM; PLICing, Phosphorothioate-based Ligase-Independent Gene Cloning; PPCP, PCR Production of Circular Plasmid; PS, Phosphorothioate; PTRec, Phosphorothioate-based DNA Recombination; RCA, Rolling circle amplification; REs, Restriction enzymes; RF cloning, Restriction-Free cloning; RGEN, RNA-guided Endonuclease; SDM, Site-directed mutagenesis; SEFC, Seamless Enzyme-Free Cloning; SeSaM, Sequence Saturation Mutagenesis; SHIPREC, Sequence Homology-Independent Protein RECombination; SLiCE, Seamless Ligation Cloning Extract; SPRINP, Single-Primer Reactions In Parallel; ssDNA, Single-stranded DNA; StEP, Staggered Extension Process; STRU-Cloning, Single-Tube Restriction-based Ultrafiltration Cloning; TaGTEAM, Targeting Glycosylases To Embedded Arrays for Mutagenesis; TALENs, Transcription activator-like effector nucleases; TAM, Transcription-associated mutation; TIM, Transposon Integration mediated Mutagenesis; TMGS-PCR, Truncated Metagenomic Gene-Specific PCR; TPCR, Transfer-PCR; TriNEx, TriNucleotide EXchange; TRINS, Tandem Repeat Insertion; T_s, Transitions; T_v, Transversions; UDG, Uracil-DNA glycosylase; USERec, USER Friendly DNA Recombination; ZFN, Zinc finger nuclease.

* Corresponding author. Tel.: +44 114 2227591; fax: +44 114 2227501.
E-mail address: t.wong@sheffield.ac.uk (T.S. Wong).

0734-9750/$ – see front matter © 2013 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.biotechadv.2013.08.021

1. Introduction

Directed evolution has emerged as a key enabling technology for tailoring or altering the properties of biomolecules (*e.g.*, proteins and nucleic acids) and of microorganisms to satisfy a wide range of biotechnological applications [*e.g.*, industrial biocatalysis, biotransformation, bioremediation and synthetic biology (Cobb et al., 2012)]. Rooted in the Darwinian theory of evolution, a typical directed evolution experiment encompasses iterative rounds of gene mutagenesis and phenotype selection through high-throughput screening, until the desired trait is attained (Bloom and Arnold, 2009).

Creating a good mutant library is arguably the most critical component in all directed evolution exercises and it requires a combination of the right mutagenesis method and an efficient cloning system. Methods of genetic diversity creation have previously been reviewed by various research groups (Bornscheuer and Kazlauskas, 2011; Shivange et al., 2009; Wong et al., 2006b). Nonetheless, the rapidly transforming field of molecular biology has fuelled creativity in scientists and we continue to see innovations in the way mutant libraries are prepared. For instance, new discoveries or better understanding of the underlying mechanisms of enzymes (*e.g.*, recombinase) and genetic systems (*e.g.*, transposition) have expanded the systems and methodologies used in mutagenesis. Advancement in cloning technology has led to simplification of the 2-step gene mutagenesis and cloning into a 1-step protocol.

In this review, we would like to provide a critical update of the cloning techniques and the genetic diversity creation methods developed for mutant library preparation over the past six years (since 2008). Specifically, the review summarises new cloning strategies that attempt to improve the conventional restriction–ligation cloning method to make it more amendable to mutant library creation. This is followed by an update of the methodologies in random mutagenesis, focused mutagenesis and DNA recombination, as well as the challenges these methods address by comparisons to more widely applied methods [*e.g.*, error-prone polymerase chain reaction (epPCR), QuikChange mutagenesis and DNA shuffling]. This update would provide a useful guide to both new and experienced directed evolutionists when developing strategies in mutant library creation. Importantly, the method comparison allows us to identify current key challenges. Mutant libraries have now seen applications beyond protein engineering. This review will survey its wider applications and conclude with a perspective on the future developments in the field of genetic diversity creation.

2. Cloning mutant libraries

In virtually all directed evolution campaigns, experimental work commences with molecular cloning of the gene of interest (GOI) into a vector for subsequent gene mutagenesis or into an expression vector for protein synthesis in an appropriate host organism (*e.g.*, *Escherichia coli*, *Bacillus subtilis*, *Pichia pastoris* and *Saccharomyces cerevisiae*). Traditional PCR-based gene mutagenesis methods (*e.g.*, epPCR and DNA shuffling) also require cloning of the mutagenized genes.

Conventionally, directional gene cloning relies on creating sticky ends (or cohesive ends) on both ends of an insert using a pair of type II restriction enzymes (REs), followed by joining the digested insert with a recipient vector pre-treated with the same pair of REs using a DNA ligase (Fig. 1A). Despite being a technique still widely employed in many research laboratories for cloning a GOI, this lengthy and time-consuming process has its challenges for cloning large mutant libraries. Incomplete restrictive digestion and poor ligation efficiency, for instance, reduce cloning efficiency. Further, suitable unique restriction sites might not be readily available and the addition of restriction sites might introduce undesired extra amino acids in the resultant recombinant protein. To overcome some or all of the aforementioned drawbacks, new ideas have been proposed and some have further been developed into commercial kits. These recent cloning methods are based on four strategies depicted in Fig. 1: (1) complementary overhangs, (2) homologous sequences, (3) overlapping PCR and (4) megaprimers.

2.1. Molecular cloning based on complementary overhangs

Among the 4 strategies, cloning based on complementary overhangs (Fig. 1A) most resembles the conventional restriction–ligation cloning method. It varies from the conventional cloning in the ways complementary overhangs between the gene insert and vector were generated. Four out of the 5 recently reported methods that used this strategy bypass the use of Type II REs, which typically generates 2–4 bp overhangs on both ends of an insert. One of the methods replaced this step by two parallel asymmetric PCRs of the GOI; one had excess reverse primer with tailing bases and the other had excess forward primer with tailing bases. As such, single-stranded DNA (ssDNA) was produced in each asymmetric PCR. ssDNAs from both PCRs were pooled and annealed to form a double-stranded fragment bearing overhangs at both ends that corresponded to the restriction overhangs of cloning vector (Wang et al., 2009a). This method however remained dependent on REs which were used to prepare the vector. Shinomiya et al. demonstrated unidirectional cloning by cleaving two distinct cloning sites with a single engineered zinc finger nuclease (ZFN) for both the GOI and vector (Mori et al., 2009; Shinomiya et al., 2011). The ZFN recognizes and cleaves a 26-bp DNA target site, generating a 2-nucleotide (nt) overhang (Shinomiya et al., 2011). Despite the replacement of RE either in GOI preparation or in both GOI and vector preparation, both methods above used DNA ligase to covalently link GOI with recipient vector,

Please cite this article as: Tee KL, Wong TS, Polishing the craft of genetic diversity creation in directed evolution, Biotechnol Adv (2013), http://dx.doi.org/10.1016/j.biotechadv.2013.08.021