NOVEL PHAGE RECOMBINATION PATHWAYS AND THEIR APPLICATIONS

Thesis submitted for the degree of Doctor of Philosophy at the University of Leicester

by

Thimma Rajkumar Reddy M.Sc

Department of Molecular and Cell Biology

University of Leicester

2016

Acknowledgements

I want to thank Dr. Shaun Cowley, Dr. Léna Fevat and Miss Sarah Munson for their support and encouragement to pursue recombineering research at Leicester. I would like to also thank Prof. Francis Stewart for critical discussions and the supply of vital recombineering reagents. I owe a great debt to Prof. John Schwabe for his support of my Ph.D candidature. I would like to also thank Dr. Sue Shackleton for her support and advice in preparing this thesis. Most important of all, I would like to acknowledge the patience and love of my wife, Léna and my three children, Lara, Miles and Kolya.

Critical Appraisal

Bacteriophage recombination systems and recombineering

Homologous recombination in *Escherichia coli* plays a central role in repairing DNA damage that arise from replication errors, UV damage and chemical mutagens (1). The RecBCD pathway is the main mechanism in *E. coli* to repair double-strand breaks (DSB) (2, 3). The RecBCD enzyme upon encountering a double-stranded DNA (dsDNA) break unwinds the DNA using the helicase activities of RecB and RecD. When the RecBCD complex reaches a specific DNA motif called the crossover hotspot instigator (Chi) site (5' GCTGGTGG 3') it nicks the DNA to release a 3' single-stranded DNA (ssDNA) tail and loads RecA protein onto the ssDNA. The RecA-ssDNA complex then invades an intact dsDNA and pairs with a homologous sequence, thus displacing the complementary strand to form a displacement loop (D-loop) structure. The D-loop is resolved either by replication primed by the 3' end invading strand or the D-loop is cut and a Holliday junction is formed. These two different resolution mechanisms are termed 'break-copy' and 'break-join' (4). The 'break-copy' mechanism requires replication and produces a recombinant DNA molecule whilst the 'break-join' mechanism does not require DNA replication and generates two recombinant molecules containing reciprocally exchanged gene markers. A secondary *E. coli* repair pathway is the RecF pathway, which is primarily used to repair single-stranded gaps (3). The 5' end of an ssDNA gap is unwound and degraded by the RecJ exonuclease in concert with the RecQ helicase. The RecF and RecOR accessory proteins then load RecA to the 3' ssDNA overhang by two different mechanisms. The following steps of strand invasion, D-loop formation and Holliday junction resolution is similar to the RecBCD pathway.

Bacteriophages encode their own recombination systems that can function independently of host recombination functions (5). The Red recombination system of the *E. coli* lambda bacteriophage encodes three proteins that together mediate efficient homologous recombination in the absence of RecA (6). λ exonuclease, encoded by *exo*, is a processive 5'-3' exonuclease (7), which degrades DNA and generates 3' ssDNA. β protein, encoded by the *bet* gene, is a DNA annealing protein that binds to ssDNA and pairs complementary DNA sequences and protects it from exonucleolytic attack (8-10). The third protein γ, encoded by *gam*, inhibits RecBCD and prevents degradation of linear dsDNA (11, 12). The RecET proteins of the *E. coli* Rac prophage are functional equivalents of λ exonuclease and β, respectively (13-15). Phage recombination systems require only short homology regions and generate high recombination efficiencies using a novel lagging strand recombination mechanism termed beta recombination, which is discussed later. Consequently, phage recombination principles have been employed in the widely used in vivo cloning technique of recombinogenic mediated engineering (recombineering).

The practical use of phage recombination systems in the form of recombineering was first demonstrated by two different reports. Zhang et al., (13) utilized the

RecET system and dsDNA cassettes containing short homology regions to efficiently perform a variety of genetic engineering tasks. They also devised a plasmid based arabinose inducible recombineering system that could be used in standard RecBCD *E. coli* strains that are commonly used for molecular cloning. The other study by Kenan Murphy (16) used a chromosomally integrated P_{lac} - Red system that replaced the *recBCD* genes to generate a controlled hyper recombination state. Recombineering methodology has been greatly improved by the development of defective prophage based recombineering systems (17, 18) and the placement of the Red operon under the control of control of arabinose or rhamnose promoters on a low copy $pSC101^{ts}$ plasmid backbone (19). These systems overcome the limitations of either toxic effects of constitutive *gam* expression, leaky nature of the *lac* promoter or the use of high copy plasmid backbones of the previous systems.

Current Recombineering Technologies and Applications

Recombineering methodology is not limited by the size and sequence restrictions of conventional DNA manipulation techniques (20). The nucleotide precision of recombineering is particularly useful in challenging DNA engineering exercises like genome and BAC mutagenesis. Oligos can be used to introduce small changes (up to 30 bp) and at high efficiencies (>1 %) that do not necessitate selection and allow screening to be performed by PCR or colony hybridization (21, 22). Gene knockouts or mutants can be rapidly created to interrogate gene function or to generate useful strains for bioprocess applications. Disabling the methyl mismatch repair system greatly enhances oligo recombination (23) but the hypermutagenic background is not suitable for many recombineering applications though some solutions have been presented (24, 25). Several genetic loci can also be simultaneously modified with Multiplex Automated Genome Engineering (MAGE) (26). The MAGE technique is discussed in greater detail later. Counterselection mutagenesis is alternate method of introducing seamless changes (27-29). A dual positive-negative selection marker is first inserted at the desired genomic location and selected using positive selection. An oligo or dsDNA fragment is then used to precisely replace the marker and applying negative selection. Counterselection mutagenesis is valuable to perform gene tagging for live cell studies and protein purification (30-32). Robust counterselection methodologies have been developed based on the finding that omission of λ exonuclease decreases the non-specific recombination rates (33).

Recombineering with dsDNA is less efficient (> 0.1%) than oligos and generally requires selection but is necessary and useful in many applications. Gene replacements, incorporation of large tags or knockouts can be made using dsDNA cassettes (34). Also, complete heterologous pathways can be transferred into a host for metabolite production (35, 36). A linked selection marker is used to select the recombinants and is often flanked by FRT or LoxP recombination sites. Expression of the Flp or Cre recombinase is used to remove the marker, leaving a small 'scar' sequence. A different dsDNA recombineering application is gap repair cloning (37). A subcloning plasmid is used to capture a DNA sequence from the genome or a plasmid, which can then be further modified with insertion

of different cassettes to generate a gene targeting/expression vector. Recombineering vector construction pipelines have been developed and have greatly facilitated the development of new genetically engineered mouse models and cell lines (38). Linear-Linear recombination, discussed later in further detail, is a variant gap repair technique that allows the cloning of a sequence directly from genomic DNA templates into heterologous expression vectors (39). This method bypasses the intermediary steps of library generation and can be used for rapid bioprospecting of various microorganisms for biotechnologically useful pathways.

Mechanistic Models of Recombineering

In a comprehensive review of the *E. coli* replication and recombination systems, Kuzminov (40) postulated that the major mechanism of Red recombination involves annealing of ssDNA to the lagging strand of the replication fork. Oligo mutagenesis experiments (41, 42) provided further support for this model with the observation of greater recombination frequencies using oligos targeting the lagging strand than the leading strand. The replication dependent bias was readily explained by the presence of single stranded gaps between Okazaki fragments on the lagging strand as opposed to the more continuous nature of DNA replication on the leading strand. However, recombination of dsDNA was suggested to involve a different mechanism involving a dual resected dsDNA intermediate (43). Two different recombination models were proposed by Court and colleagues (44) based on the size of the heterologous sequence present between the two flanking homology regions. Small mutations are incorporated by replication fork backtracking involving a chicken foot intermediate while larger gene sized insertions are achieved by the rescue of a stalled fork with an incoming fork in the opposite direction. Poteete (45) tested the Court model and the requirement of two different replications forks using unidirectionally replicating plasmids and found that recombination occurred with a lagging strand bias. An alternate replisome invasion/template switch mechanism was suggested for the replication dependency of Red recombination.

The first report to suggest the involvement of a single-stranded DNA intermediate in Red mediated dsDNA recombination came from the study of Lim et al., (46) who used asymmetric dsDNA cassettes containing two different antibiotic resistance genes separated by three regions of homology. Recombination at different genomic loci was favored by cassette configurations that required initial binding of the terminal homology region to the lagging strand. The majority of recombinants contained only one of the antibiotic markers suggesting extensive processing similar to the mechanism of Okazaki fragment generation. The ssDNA recombination model was confirmed by two separate studies using two different approaches. Mosberg and colleagues (47) used dsDNA cassettes containing terminal mismatches to track the genetic inheritance of strand specific mutations and Maresca et al., (48) used asymmetric phosphorothioate modified cassettes to promote the in vivo generation of ssDNA intermediates. Both studies showed that dsDNA cassettes are processed by λ exonuclease or host exonucleases/helicases to generate an ssDNA intermediate that is annealed by β to the lagging strand of the replication fork. These findings

unified the recombination mechanism of both ssDNA and dsDNA into a single model termed the beta model (48). The beta model also suggested methods to improve recombineering efficiencies through targeting the lagging strand (48, 49) and mutating the DnaG primase (50) to increase the availability of ssDNA regions in the replication fork. The study of Maresca et al., also identified a different ends-in Red recombination pathway that did not show a lagging strand bias, suggesting that other recombination pathways can operate at lower efficiencies.

Red recombination between a linear DNA and a circular DNA (13, 51) and between two linear DNA molecules (37, 39) have been termed linear plus circular homologous recombination (LCHR) and linear plus linear homologous recombination (LLHR), respectively. LLHR is a form of gap repair that involves the insertion of a linear DNA fragment into a linear subcloning plasmid through exonuclease digestion and annealing and does not require active DNA replication (39). In contrast, gap repair using replicating chromosomal or plasmid repair templates represents a LCHR type reaction (52). The two processes also differ in their homology requirement. LLHR shows a linear relationship with increasing homology (39) as opposed to LCHR, which shows maximal recombination at 120 bp (Red system) and 150 bp (RecET system) (53). A study by Fu et al., (39) demonstrated that full-length RecE is more specialized at performing LLHR and conversely the Red system is more efficient at mediating LCHR. The authors hypothesized that LLHR and LCHR represent two different annealing mechanisms, the former involving simple annealing and the latter involving a replication fork.

1st Published Work: Gap Repair Recombination

Background

The beta model describes how linear DNA is inserted into circular replicating DNA. Conversely, gap repair involves the replication mediated insertion of DNA into a linear gapped subcloning plasmid from a circular source like the *E. coli* chromosome or a BAC plasmid. Therefore, gap repair is an extraction process as opposed to insertion. Conceivably, these two different recombineering processes may operate through different recombination pathways. The double-strand break repair (DBSR) model has been previously suggested to explain Red mediated gap repair (54-57). In this model RecA promoted strand invasion is followed by Holliday formation leading to gene conversion with crossing-over. Indeed, these preliminary experiments were performed with expression of RecA and plasmid templates that did not require active DNA replication to mediate gap repair. In the absence of RecA, the classical DSBR pathway generated gap repair recombinants at very low frequencies (10^{-5}). But the reported efficiencies of recombineering mediated gap repair without RecA are over 1,000 fold higher (10^{-2}) (58). The seminal work of Stahl and co-workers (59) showed that Red recombination during active DNA replication utilized a predominantly single-strand annealing mechanism even in the presence of RecA. Also, data from in vitro experiments showing β mediated strand invasion (60) has not been

corroborated by in vivo studies (45, 59), which support a strand annealing mechanism. A strict replication dependency of recombineering mediated gap repair has been previously shown (52) and thus necessitates a suitable mechanistic explanation. Mosberg and colleagues (47) put forth an elegant description of gap repair using the beta model. In this scenario, the linear subcloning plasmid is processed to generate the ssDNA recombination intermediate, which binds to the lagging strand with the terminal homology ends facing each other. Replication via Okazaki fragment synthesis fills in the intervening plasmid gap leading to the release of a closed circular ssDNA plasmid, which is then subsequently converted to a dsDNA form. Indeed, it has been suggested that replication fork progression in *E. coli* is not affected by events in regions preceding the fork (48) and that the stability of β annealed DNA can allow the ends of the linear DNA to anneal to homologous regions that are kilobases (> 50 kb) apart (26, 48). However, an important consideration of the beta model is the large insert sizes (> 80 kb) that can be subcloned by gap repair (61). How are such large ssDNA gaps in the template DNA tolerated or repaired after the subcloning plasmid is released? The endogenous *E. coli* gap repair pathways are known only to repair short tracts of exposed ssDNA (1). Replication fork backtracking over such large genomic distances has not been demonstrated and a failure to repair leads to replication catastrophe and cell death (62). A detailed mechanism of gap repair is lacking and the role of lagging strand recombination in gap repair needs further investigation.

Findings of this thesis

The first published work of this thesis describes a series of experiments that compared gap repair and insertion processes. The assay system utilized asymmetric terminal modified linear DNA cassettes to generate ssDNA in vivo and examined the strand directionality of recombination using a BAC plasmid. Insertion of a linear cassette showed the expected lagging strand bias. In contrast, gap repair was equally efficient on the leading and lagging strands. This suggested that gap repair did not involve lagging strand recombination. To verify that gap repair did not invoke a lagging strand recombination mechanism, insertion of a cassette was performed simultaneously during gap repair in a process termed SPI. The absence of lagging strand bias of insertion during gap repair again confirmed a different mechanistic basis of gap repair. Quantitative analysis of the SPI process revealed a predominantly concerted mode of insertion and gap repair. Further experiments with ssDNA and dsDNA subcloning plasmids and oligo mediated correction of a defective antibiotic marker in DNA Polymerase I defective cells directly showed the absence of lagging strand recombination mediated gap repair. Accumulating evidence suggests that lagging strand recombination is the major pathway of Red recombination (46-48). However, an alternate and less efficient pathway of 'ends-in' recombination of DNA has also been demonstrated (48). SPI performed with ends-out cassettes was more efficient than with ends-in cassettes suggesting that gap repair, also a form of ends-in recombination, is different from this alternate pathway. In this regard, another disfavoured Red recombination process involving dual resected DNA has also been observed (47). However, SPI recombination was found to be equally efficient with ssDNA and

dsDNA oligos. Overall, these results demonstrated that Red mediated gap repair is a novel homologous recombination pathway in *E. coli*.

The claim of a novel gap repair pathway requires consideration of alternative possibilities. Could a variation on the DBSR model or a derivative thereof like modified synthesis-dependent strand annealing (SDSA) perhaps be involved? The DBSR pathway requires the formation of a Holliday junction. However, Red recombination pathways are unaffected by deletion of the Holliday junction *ruvABC* genes (48). Indeed, Sawitzke *et al*., (63) knocked out nearly 14 *E. coli* host recombination proteins but still failed to abolish lagging strand bias of Red mediated oligo recombination. In addition, DSBR is also much more efficient with an end-in configuration than an ends-out configuration (64). The lower efficiency of SPI with ends-in cassettes relative to ends-out cassettes argues against a DBSR like mechanism. Similar to DBSR, SDSA requires strand invasion as the initiating step. However, our gap repair experimental set-up favors annealing type reactions. Additionally, DBSR and SDSA models are not compatible with multiplexed recombination. Template switching has been previously hypothesized for Red mediated insertion (45) but may be refuted since it would require a high level of co-ordination of the replication machinery to mediate multiple template switches and the argument is not parsimonious for the insertion of multiple cassettes. In fact, the addition of each insertion cassette with SPI resulted in only a 10-fold decrease in recombination frequency though co-selection (65, 66) could have been a contributing factor. Other models of dsDNA recombination based on the RecG pathway and bidirectional replication forks as proposed by Court and colleagues (44) also cannot explain SPI recombination without significant modification. The observation of modification of both DNA strands with only an ssDNA under our experimental conditions supports a new model of homologous recombination.

The lagging strand bias of insertion is caused by the replication fork structure and is conserved across all domains (41, 48, 67-70). Lagging strand recombination also provides a mechanistic understanding of the high efficiency of insertion. Gap repair does not involve a lagging strand recombination mechanism yet quantitative analysis of the SPI process show that it is a more efficient process than insertion. Whether this is due to a more efficient formation of the gap repair intermediate relative to insertion or a higher efficiency of insertion in the gap repair intermediate than at the replication fork is not known. The observation of efficient leading strand targeting, similar insertion efficiencies at different genomic sites and multiplexing (on the same contagious DNA) with ends-in cassettes, which is not observed with ends-in insertion at the replication fork suggests that Red mediated gap repair could utilize a different replication structure than the canonical replication fork. Bacteriophage recombination systems like Red recombination have evolved to take advantage of host replication processes (71). The gap repair intermediate could represent a phage specific process than a naturally occurring host structure, though its significance to the lambda phage life cycle is uncertain. Single molecule imaging of the *E. coli* replisome from a SPI reaction combined with immunoprecipitation of the replisome proteins should shed more light on the nature of the gap repair replisome architecture.

Overall, the comparison of lagging strand recombination and gap repair pathways in this study provided the basis for the characterization of lambda Red mediated gap repair, which is a poorly understood process. Elucidating the gap repair mechanism is important to developing a more comprehensive understanding of *E. coli* replication and its interaction with phage recombination systems. However, the finding of a novel gap repair pathway also has practical utility. We capitalized on the SPI phenomenon to select against intramolecular non-gap repair recombinant clones. Preliminary tests of SPI cloning with different number of insertion cassettes and plasmid backbones demonstrated its utility in plasmid vector construction. The SPI application is discussed in greater detail in the next published work.

2nd Published Work: Subcloning Plus Insertion

Background

Gene targeting is a widely used method to construct novel cell lines and transgenic models with defined genetic modifications to interrogate gene function (72). A key stage in the gene targeting process is the construction of a gene targeting vector. Gene targeting vectors are plasmids that contain the modified allele with flanking genomic regions to promote homologous recombination in the target cells. The construction of a gene targeting vector is often a complex process requiring the error-free cloning of large genomic regions and the precise placement of selection markers, recombination sites like LoxP and FRT and minigene cassettes. The cloning process represents a bottleneck in the gene targeting workflow. Recently, the technique of recombineering has been used to construct gene targeting vectors more efficiently than previous 'cut and paste' conventional cloning methods (37, 73-76), which suffer from the limitations of restriction site availability and low throughput. However, typical recombineering vector construction strategies involve multiple steps, use different strains and require intermediate purification steps (38, 77, 78). In the previous published work on gap repair recombination, the technique of subcloning plus insertion (SPI) was introduced.

Findings of this thesis

The second published work shows some examples demonstrating its efficacy in efficient and rapid assembly of vectors for different applications. SPI cloning was tested to construct gene targeting vectors that were used to tag mouse genes with tags like FLAG, and calmodulin binding protein (CBP) and fluorescent labels like enhanced yellow fluorescent protein (eYFP). A rapid method of producing tagged BAC plasmids was shown using a novel SPI BAC trimming strategy. Conditional knockout vector construction was also demonstrated with the SPI technique. A key finding was that SPI generated a large proportion of correct recombinants even with suboptimal cloning conditions like use of short homology lengths, lack of protection of the insertion cassettes, large gap repair sizes and the cloning of 3'UTR regions containing strong secondary structures.

The high recombination frequencies obtained with SPI cloning suggest that this technique could be useful in construction of different advanced vector designs. Overall, the SPI procedure proved to be an efficient and robust method to rapidly construct the desired vector.

This study shows that SPI offers a simple, rapid and flexible method to construct a variety of gene targeting vectors and is a major improvement on existing methods of vector construction. The SPI method is important to accelerate the use of gene targeting to rapidly develop newer gene models. The unique video format of this paper is also useful to make the multiplex protocol more accessible to non-recombineering specialists and increase its acceptance as a standard genetic tool.

3rd Published Work: Multiplex Recombineering of Double-Stranded DNA

Background

Standard protocols for different recombineering applications typically utilize a singleplex strategy whereby a linear DNA cassette or oligo is inserted into the *E .coli* genome or a plasmid or alternatively a region of DNA is gap repaired into a subcloning plasmid (21). A selection scheme (13) is usually employed or a PCR screening method is used to identify the correct recombinant (22, 26). This process is repeated to introduce other changes in the DNA sequence to generate the final desired recombinant clone. However, singleplex strategies limit the scope of recombineering applications. Large-scale genome recoding for rapid, combinatorial engineering of strains and the high throughput construction of gene targeting vector libraries require a multiplex approach. However, the low efficiency of singleplex recombineering methods prevented a multiplex design in the established protocols.

Church and colleagues set about to improve oligo recombineering efficiencies using a combination of optimized oligo length and homology sequence, improved oligo secondary structure modeling, phosphorothoate protection of oligos, use of methyl mismatch repair deficient strains (*ΔmutS*) and optimized growth conditions and achieved over a 100 fold increase in oligo recombineering efficiency (26, 79). The increased recombineering efficiency allowed them to develop the MAGE technique (26). MAGE is a multiplex recombineering method that involves insertion of multiple oligos simultaneously at different genomic sites. The MAGE process is repeated for a number of cycles to generate a population containing a large proportion of modified cells. An additional benefit of the MAGE system is the chemostat like nature of the MAGE culture, which allows a greater number of mutations to be introduced than the population size. The resulting combinatorial library of modified cells can be phenotyped for the desired property. MAGE and a variant technique termed Multiplexed Iterative Plasmid Engineering (MIPE) (80) have allowed large scale tagging of genes (81), optimization of metabolite pathways (25, 26, 65, 82) and the development of novel biocontainment strategies (83, 84). A recent improvement in MAGE

technology utilized small molecule sensors to select the desired producer strain thus obviating the need for exhaustive screening (85).

In the previous two papers of this thesis, the concept of multiplex recombineering using large DNA constructs was outlined. Double-stranded DNA multiplex recombineering allows the simultaneous insertion of whole genes at different genomic targets in the same cell. Similar to MAGE, this novel multiplex gene insertion methodology relies on the use of lagging strand protected cassettes to achieve the increase in recombineering efficiency necessary for multiplexing. Lagging strand recombination and SPI represent the two different recombination pathways of multiplex recombineering and provide the appropriate experimental set-up to compare singleplex and multiplex methodology and to identify parameters specific and useful for dsDNA multiplex recombineering. A greater understanding of dsDNA multiplex recombineering is required for efficient generation of producer strains containing complete heterologous metabolic pathways and plasmid libraries for bioprocess applications.

Findings of this thesis

The third published work of this thesis tested several different recombineering parameters of the Red system using singleplex and multiplex assays. The homology length of the linear DNA cassettes was found to be a key determinant of multiplex recombination efficiency. Although, maximal multiplex recombination was observed with longer homology lengths (> 180 bp), 60 bp of homology generated efficient recombination. Notably, long oligos suffer from a greater percentage of truncations and oligo synthesis errors and are more expensive to synthesize. The ability to perform multiplex recombineering using shorter oligos is useful. The greater homology requirement of 60 bp with multiplex recombineering compared to 35 bp with singleplex recombineering is not surprising considering that recombination of the different DNA fragments needs to occur on the same contiguous DNA. Indeed, applying the co-selection principle enhanced multiplexing efficiency. The amount of DNA available for recombination was another factor that greatly affected multiplex recombination frequency. Lower electroporation efficiencies using multiple cassettes and the degradation of the linear DNA by endogenous *E. coli* exonucleases likely limited multiplex recombination. All three proteins of the Red recombination were required for efficient multiplex recombination suggesting that the Exo-Beta synergy was also an important element of multiplexing using dsDNA. Cross-validation of these findings in an ExoVII deletion strain suggested a strategy to increase multiplex recombination rates.

This study represents the first systematic attempt at delineating the parameters that affect multiplex recombineering of gene-sized dsDNA cassettes. Comparison of multiplex recombineering using oligos with dsDNA cassettes identified common determinants between the two systems that may be key general parameters of multiplex recombineering. Our findings are important for the application of the multiplex recombineering in different genome engineering exercises and form the basis to further improve multiplex recombineering

efficiencies. Improved recombineering methodologies could expand the spectrum of in vivo bacterial genetic engineering applications. In addition to multiplex recombination, this study also provided primary data on the homology requirement of gap repair. Maximal recombination was observed with 120 bp of homology similar to lagging strand recombination. Notably, the process of linear-linear homology recombination (also a form of gap repair) does not reach maximal recombination at 120 bp (39) and shows a linear relationship with increasing homology.

Concluding remarks

The published works presented in this thesis demonstrate a novel gap repair pathway in *E. coli* cells expressing a phage annealing protein. The gap repair pathway was also used to develop a novel recombineering application termed SPI for more efficient gene targeting vector construction. Furthermore, comparison of singleplex and multiplex recombineering using the gap repair and lagging strand recombination pathways identified several key multiplex recombination parameters. A better understanding of the mechanism of gap repair is required to identify the role of alternative replication and recombination structures in the maintenance of genomic integrity and evolution. Mechanistic knowledge of the recombineering pathways is also useful for the development of novel recombineering applications. The possibility of generating artificial bacteria de novo through large-scale genome recoding and gene insertion represents an exciting future venture.

References

1. Kuzminov A. Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage λ. Microbiol Mol Biol Rev. 1999;63: 751-813.
2. Dillingham MS, Kowalczykowski SC. RecBCD enzyme and the repair of double-stranded DNA breaks. Microbiol Mol Biol Rev. 2008;72: 642-671.
3. Spies M, Kowalczykowski SC. Homologous recombination by RecBCD and RecF pathways. The bacterial chromosome. In Higgins NP, editor. ASM Press, Washington, DC. 2005: p 389-403.
4. Motamedi MR, Szigety SK, Rosenberg SM. Double-strand-break repair recombination in *Escherichia coli*: physical evidence for a DNA replication mechanism in vivo. Genes Develop. 1999;13: 2889-2903.
5. Szczepanska AK. Bacteriophage-encoded functions engaged in initiation of homologous recombination events. Crit Rev Microbiol. 2009;35: 197-220.
6. Signer ER, Weil J. Recombination in bacteriophage lambda. I. Mutants deficient in general recombination. J Mol Biol. 1968;34: 261-271.
7. Subramanian K, Rutvisuttinunt W, Scott W, Myers RS. The enzymatic basis of processivity in λ exonuclease. Nucleic Acids Res. 2003;31: 1585-1596.
8. Muniyappa K, Radding CM. The homologous recombination system of phage lambda. Pairing activities of beta protein. J Biol Chem. 1986;261: 7472-7478.

9. Karakousis G, Ye N, Li Z, Chiu S, Reddy G, Radding C. The beta protein of phage lambda binds preferentially to an intermediate in DNA renaturation. J Mol Biol. 1998;276: 721-731.
10. Li Z, Karakousis G, Chiu SK, Reddy G, Radding CM. The beta protein of phage lambda promotes strand exchange. J Mol Biol. 1998;276: 733-744.
11. Murphy KC. The λ Gam Protein Inhibits RecBCD Binding to dsDNA Ends. J Mol Biol. 2007;371: 19-24.
12. Murphy KC. λ Gam protein inhibits the helicase and χ-stimulated recombination activities of *Escherichia coli* RecBCD enzyme. J Bacteriol. 1991;173: 5808-5821.
13. Zhang Y, Buchholz F, Muyrers J, Stewart A. A new logic for DNA engineering using recombination in *Escherichia coli*. Nat Genet. 1998;20: 123-128.
14. Muyrers JPP, Zhang Y, Buchholz F, Stewart AF. RecE/RecT and Redα/Redβ initiate double-stranded break repair by specifically interacting with their respective partners. Genes Develop. 2000;14: 1971-1982.
15. Hall SD, Kolodner RD. Homologous pairing and strand exchange promoted by the *Escherichia coli* RecT protein. Proc Natl Acad Sci USA. 1994;91: 3205-3229.
16. Murphy KC. Use of bacteriophage λ recombination functions to promote gene replacement in *Escherichia coli*. J Bacteriol. 1998;180: 2063-2071.
17. Datta S, Costantino N, Court DL. A set of recombineering plasmids for gram-negative bacteria. Gene. 2006;379: 109-115.
18. Yu D, Ellis H, Lee E, Jenkins N, Copeland N, Court D. An efficient recombination system for chromosome engineering in *Escherichia coli*. Proc Natl Acad Sci USA. 2000;97: 5978-5983.
19. Wang J, Sarov M, Rientjes J, Fu J, Hollak H, Kranz H, et al. An improved recombineering approach by adding RecA to λ red recombination. Mol Biotechnol. 2006;32: 43-53.
20. Copeland N, Jenkins N, Court D. Recombineering: a powerful new tool for mouse functional genomics. Nat Rev Genet. 2001;2: 769-779.
21. Sharan SK, Thomason LC, Kuznetsov SG, Court DL. Recombineering: a homologous recombination-based method of genetic engineering. Nat Protoc. 2009;4: 206-223.
22. Swaminathan S, Ellis HM, Waters LS, Yu D, Lee EC, Court DL, et al. Rapid engineering of bacterial artificial chromosomes using oligonucleotides. Genesis. 2001;29: 14-21.
23. Costantino N, Court D. Enhanced levels of lambda Red-mediated recombinants in mismatch repair mutants. Proc Natl Acad Sci USA. 2003;100: 15748-15753.
24. Nyerges Á, Csörgő B, Nagy I, Latinovics D, Szamecz B, Pósfai G, et al. Conditional DNA repair mutants enable highly precise genome engineering. Nucleic Acids Res. 2014;42: e62.
25. Ryu YS, Biswas RK, Shin K, Parisutham V, Kim SM, Lee SK. A simple and effective method for construction of *Escherichia coli* strains proficient for genome engineering. PLoS ONE. 2014;9: e94266.
26. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, et al. Programming cells by multiplex genome engineering and accelerated evolution. Nature. 2009;460: 894-898.

27. Warming S, Costantino N, Court DL, Jenkins NA, Copeland NG. Simple and highly efficient BAC recombineering using *galK* selection. Nucleic Acids Res. 2005;33: e36.
28. Wang H, Bian X, Xia L, Ding X, Müller R, Zhang Y, et al. Improved seamless mutagenesis by recombineering using ccdB for counterselection. Nucleic Acids Res. 2014;42: e37.
29. Gregg CJ, Lajoie MJ, Napolitano MG, Mosberg JA, Goodman DB, Aach J, et al. Rational optimization of tolC as a powerful dual selectable marker for genome engineering. Nucleic Acids Res. 2014;42: 4779-4790.
30. Sarov M, Murray John I, Schanze K, Pozniakovski A, Niu W, Angermann K, et al. A genome-scale resource for in vivo tag-based protein function exploration in *C. elegans*. Cell. 2012;150: 855-866.
31. Sarov M, Schneider S, Pozniakovski A, Roguev A, Ernst S, Zhang Y, et al. A recombineering pipeline for functional genomics applied to *Caenorhabditis elegans*. Nat Methods. 2006;3: 839-844.
32. Poser I, Sarov M, Hutchins J, Heriche J, Toyoda Y, Pozniakovsky A, et al. BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals. Nat Methods. 2008;5: 409-415.
33. Bird AW, Erler A, Fu J, Heriche J-K, Maresca M, Zhang Y, et al. High-efficiency counterselection recombineering for site-directed mutagenesis in bacterial artificial chromosomes. Nat Methods. 2012;9: 103-109.
34. Zhang Y, Buchholz F, Muyrers JPP, Stewart AF. A new logic for DNA engineering using recombination in *Escherichia coli*. Nat Genet. 1998;20: 123-128.
35. Fu J, Wenzel SC, Perlova O, Wang J, Gross F, Tang Z, et al. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition. Nucleic Acids Res. 2008;36: e113.
36. Perlova O, Fu J, Kuhlmann S, Krug D, Stewart AF, Zhang Y, et al. Reconstitution of the myxothiazol biosynthetic gene cluster by Red/ET recombination and heterologous expression in *Myxococcus xanthus*. App Env Microbiol. 2006;72: 7485-7494.
37. Zhang Y, Muyrers JPP, Testa G, Stewart AF. DNA cloning by homologous recombination in *Escherichia coli*. Nat Biotechnol. 2000;18: 1314-1317.
38. Skarnes WC, Rosen B, West AP, Koutsourakis M, Bushell W, Iyer V, et al. A conditional knockout resource for the genome-wide study of mouse gene function. Nature. 2011;474: 337-342.
39. Fu J, Bian X, Hu S, Wang H, Huang F, Seibert PM, et al. Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. Nat Biotechnol. 2012;30: 440-446.
40. Kuzminov A. Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. Microbiol Mol Biol Rev. 1999;63: 751-813.
41. Ellis H, Yu D, DiTizio T, Court D. High efficiency mutagenesis repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. Proc Natl Acad Sci USA. 2001;98: 6742-6746.
42. Zhang Y, Muyrers J, Rientjes J, Stewart A. Phage annealing proteins promote oligonucleotide-directed mutagenesis in *Escherichia coli* and mouse ES cells. BMC Mol Biol. 2003;4: 1.

43. Yu D, Sawitzke J, Ellis H, Court D. Recombineering with overlapping single-stranded DNA oligonucleotides: testing a recombination intermediate. Proc Natl Acad Sci USA. 2003;100: 7207-7212.
44. Court D, Sawitzke J, Thomason L. Genetic engineering using homologous recombination. Annu Rev Genet. 2002;36: 361-388.
45. Poteete A. Involvement of DNA replication in phage lambda Red-mediated homologous recombination. Mol Microbiol. 2008;68: 66-74.
46. Lim S, Min B, Jung G. Lagging strand-biased initiation of red recombination by linear double-stranded DNAs. J Mol Biol. 2008;384: 1098-1105.
47. Mosberg JA, Lajoie MJ, Church GM. Lambda Red recombineering in *Escherichia coli* occurs through a fully single-stranded intermediate. Genetics. 2010;186: 791-799.
48. Maresca M, Erler A, Fu J, Friedrich A, Zhang Y, Stewart AF. Single-stranded heteroduplex intermediates in lambda Red homologous recombination. BMC Mol Biol. 2010;11: 54.
49. Mosberg JA, Gregg CJ, Lajoie MJ, Wang HH, Church GM. Improving lambda Red genome engineering in *Escherichia coli* via rational removal of endogenous nucleases. PLoS ONE. 2012;7: e44638.
50. Lajoie MJ, Gregg CJ, Mosberg JA, Washington GC, Church GM. Manipulating replisome dynamics to enhance lambda Red-mediated multiplex genome engineering. Nucleic Acids Res. 2012;40: e170.
51. Murphy K. Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. J Bacteriol. 1998;180: 2063-2071.
52. Maresca M, Erler, A.S., Fu, F., Seibert, P.M., Stewart, A.F. & Zhang, Y., inventor; Gene Bridges GmbH, assignee. Method of nucleic acid recombination. USA: US-Pat. 8728820 patent US 8728820 B2. 2014.
53. Muyrers J, Zhang Y, Buchholz F, Stewart A. RecE/RecT and Redalpha/Redbeta initiate double-stranded break repair by specifically interacting with their respective partners. Genes Dev. 2000;14: 1971-1982.
54. Kobayashi I, Takahashi N. Double-stranded gap repair of DNA by gene conversion in *Escherichia coli*. Genetics. 1988;119: 751-757.
55. Kobayashi I. Mechanisms for gene conversion and homologous recombination: The double-strand break repair model and the successive half crossing-over model. Adv Biophy. 1992;28: 81-133.
56. Takahashi NK, Kusano K, Yokochi T, Kitamura Y, Yoshikura H, Kobayashi I. Genetic analysis of double-strand break repair in *Escherichia coli*. J Bacteriol. 1993;175: 5176-5185.
57. Takahashi N, Kobayashi I. Evidence for the double-strand break repair model of bacteriophage lambda recombination. Proc Natl Acad Sci USA. 1990;87: 2790-2794.
58. Liu P, Jenkins NA, Copeland NG. A highly efficient recombineering-based method for generating conditional knockout mutations. Genome Res. 2003;13: 476-484.
59. Stahl MM, Thomason L, Poteete AR, Tarkowski T, Kuzminov A, Stahl FW. Annealing vs. invasion in phage λ recombination. Genetics. 1997;147: 961-977.
60. Rybalchenko N, Golub E, Bi B, Radding C. Strand invasion promoted by recombination protein beta of coliphage lambda. Proc Natl Acad Sci USA. 2004;101: 17056-17060.

61. Lee EC, Yu D, Martinez de Velasco J, Tessarollo L, Swing DA, Court DL, et al. A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. Genomics. 2001;73: 56-65.
62. Michel B, Grompone G, Florès M-J, Bidnenko V. Multiple pathways process stalled replication forks. Proc Natl Acad Sci USA. 2004;101: 12783-12788.
63. Sawitzke JA, Costantino N, Li XT, Thomason LC, Bubunenko M, Court C, et al. Probing cellular processes with oligo-mediated recombination and using the knowledge gained to optimize recombineering. J Mol Biol. 2011;407: 45-59.
64. Hastings PJ, McGill C, Shafer B, Strathern JN. Ends-in vs. ends-out recombination in yeast. Genetics. 1993;135: 973-980.
65. Wang HH, Kim H, Cong L, Jeong J, Bang D, Church GM. Genome-scale promoter engineering by coselection MAGE. Nat Methods. 2012;9: 591-593.
66. Carr PA, Wang HH, Sterling B, Isaacs FJ, Lajoie MJ, Xu G, et al. Enhanced multiplex genome engineering through co-operative oligonucleotide co-selection. Nucleic Acids Res. 2012;40: e132.
67. Yamamoto T, Moerschell RP, Wakem LP, Komar-Panicucci S, Sherman F. Strand-specificity in the transformation of yeast with synthetic oligonucleotides. Genetics. 1992;131: 811-819.
68. Wu X-S, Xin L, Yin W-X, Shang X-Y, Lu L, Watt RM, et al. Increased efficiency of oligonucleotide-mediated gene repair through slowing replication fork progression. Proc Natl Acad Sci USA. 2005;102: 2508-2513.
69. van Pijkeren J-P, Britton RA. High efficiency recombineering in lactic acid bacteria. Nucleic Acids Res. 2012;40: e76.
70. Van Kessel JC, Hatfull GF. Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering: characterization of antimycobacterial drug targets. Mol Microbiol. 2008;67: 1094-1107.
71. Fricker AD, Peters JE. Vulnerabilities on the lagging-strand template: opportunities for mobile elements. Annu Rev Genet. 2014;48: 167-186.
72. Capecchi MR. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat Rev Genet. 2005;6: 507-512.
73. Nedelkova M, Maresca M, Fu J, Rostovskaya M, Chenna R, Thiede C, et al. Targeted isolation of cloned genomic regions by recombineering for haplotype phasing and isogenic targeting. Nucleic Acids Res. 2011;39: e137.
74. Valenzuela DM, Murphy AJ, Frendewey D, Gale NW, Economides AN, Auerbach W, et al. High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. Nat Biotechnol. 2003;21: 652-659.
75. Testa G, Zhang Y, Vintersten K, Benes V, Pijnappel WWMP, Chambers I, et al. Engineering the mouse genome with bacterial artificial chromosomes to create multipurpose alleles. Nat Biotechnol. 2003;21: 443-447.
76. Muyrers J, Zhang Y, Testa G, Stewart A. Rapid modification of bacterial artificial chromosomes by ET-recombination. Nucleic Acids Res. 1999;27: 1555-1557.
77. Chan W, Costantino N, Li R, Lee SC, Su Q, Melvin D, et al. A recombineering based approach for high-throughput conditional knockout targeting vector construction. Nucleic Acids Res. 2007;35: e64.
78. Fu J, Teucher M, Anastassiadis K, Skarnes W, Stewart AF. A recombineering pipeline to make conditional targeting constructs. In: Paul MW,

Philippe MS, editors. Methods in Enzymology. Volume 477: Academic Press; 2010. p. 125-144.
79. Wang HH, Church GM. Multiplexed genome engineering and genotyping methods applications for synthetic biology and metabolic engineering. In Voigt C, editor. Methods in Enzymology. Volume 498: Academic Press; 2011. p. 409-426.
80. Li Y, Gu Q, Lin Z, Wang Z, Chen T, Zhao X. Multiplex iterative plasmid engineering for combinatorial optimization of metabolic pathways and diversification of protein coding sequences. ACS Synth Biol. 2013;2: 651-661.
81. Wang HH, Huang P-Y, Xu G, Haas W, Marblestone A, Li J, et al. Multiplexed in vivo his-tagging of enzyme pathways for in vitro single-pot multienzyme catalysis. ACS Synth Biol. 2012;1: 43-52.
82. Li Y, Gu Q, Lin Z, Wang Z, Chen T, Zhao X. Multiplex iterative plasmid engineering for combinatorial optimization of metabolic pathways and diversification of protein coding sequences. ACS Synth Biol. 2013;2: 651-661.
83. Mandell DJ, Lajoie MJ, Mee MT, Takeuchi R, Kuznetsov G, Norville JE, et al. Biocontainment of genetically modified organisms by synthetic protein design. Nature. 2015;518: 55-60.
84. Rovner AJ, Haimovich AD, Katz SR, Li Z, Grome MW, Gassaway BM, et al. Recoded organisms engineered to depend on synthetic amino acids. Nature. 2015;518: 89-93.
85. Raman S, Rogers JK, Taylor ND, Church GM. Evolution-guided optimization of biosynthetic pathways. Proc Natl Acad Sci USA. 2014;111: 17803-17808.

Table of Published Works

1: **Reddy T.R.**, Fevat, L.M.S., Munson, S.E., Stewart, A.F., Cowley, S.M. Lambda Red mediated gap repair utilizes a novel replicative intermediate in *Escherichia coli.* PLoS ONE. 2015; 10: e0120681. http://dx.doi.org/10.1371/journal.pone.0120681

2: **Reddy T.R.**, Kelsall, E.J., Fevat, L.M.S., Munson, S.E., Cowley, S.M. Subcloning plus insertion (SPI) - a novel recombineering method for the rapid construction of gene targeting vectors. J Vis Exp. 2015; 95: e52155. http://dx.doi.org/10.3791/52155

3: **Reddy T.R.**, Kelsall, E.J., Fevat, L.M.S., Munson, S.E., Cowley, S.M. Differential requirements of singleplex and multiplex recombineering of large DNA constructs. PLoS ONE. 2015; 10: e0125533. http://dx.doi.org/10.1371/journal.pone.0125533