Towards the Development of a ϕC31Int-Competent Sf9 Cell Line

by

Paul Lachance Brogee

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Applied Science

in

Chemical Engineering

Waterloo, Ontario, Canada, 2018

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The Sf9 cell line, as a component of the Baculovirus Expression Vector System (BEVS), is a popular platform for industrial-scale production of protein products. However, the natural course of the BEVS results in host-cell lysis before many down-stream protein modifications can be made, making it suboptimal for production of complex products such as membrane or glycosylated proteins. Stably transforming a Sf9 cell line for constitutive expression of a gene product can effectively circumvent this issue, however the process of developing these cell lines remains inefficient and time-consuming.

Synthetic biology is a field of research that approaches biological problems through the lens of engineering design. With this in mind, a design for a Sf9 line was developed that leverages ϕC31 Integrase technology to facilitate Recombinase Mediated Cassette Exchange (RMCE) *in vivo* in the Sf9 host cell. RMCE is a type of site-specific DNA recombination that results in the exchange of a DNA cassette from, in this case, a donor plasmid to a genomic locus within the host cell.

A pair of fluorescent reporter proteins, mAzamiGreen and mKOk, were tested to determine their suitability for use in confirming successful cassette exchange in Sf9. These were found to be easily differentiable in both mixed culture and co-transfection scenarios.

Additionally, some insight was gleaned with respect to media formulation when performing a limiting dilution protocol on Sf9 cells. Culturing cells in a 50/50 Fresh/Conditioned media formulation showed increased cell viability when compared with 100% conditioned media.

The addition of FBS drastically improved cell viability after 1 week. However, the difference between 2% and 10% FBS supplementation was not significant.

The genetic components of the design were assembled via Splice Overlap Extension PCR and Gibson Assembly protocols. However, complications during the construction of the genetic architecture meant that the design was not seen to completion. Nevertheless, the lessons learned throughout this work will aid in future attempts to develop the proposed system.

Acknowledgements

Firstly, I would like to extend my deepest gratitude to Dr. Marc Aucoin. Thank you for taking me on in your lab, for allowing me to develop a project I was interested in, and for your patience along the way. Most importantly, thank you for a level of support and understanding over the past few years that was above and beyond the call of duty.

Thank you to Dr. Trevor Charles and Dr. Valerie Ward for generously taking the time to review this work, and providing your guidance and feedback.

I am sincerely appreciative of all of my lab mates during my tenure in the Aucoin Lab - thank you for your input and insight both in the lab and in the lunch room. In particular, I'd like to acknowledge Mark Bruder for his guidance in all things cloning, and Marco Quattrociocchi & Sadru Walji for sharing their tips on maintaining Sf9 cells. Thanks to Julia Manalil and Megan Logan for sharing both an office space and your friendship during my time here.

Dedication

To my always supportive family,

John, Julie, Mike, and Karen.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

αE	E-Helix
aa	Amino Acid
AcMNPV	*Autographa californica* Nucleopolyhedrovirus
BAC	Bacterial Artificial Chromosome
bp	Base Pairs
BEVS	Baculvirus Expression Vector System
CC	Coiled-Coil
CDS	Coding Sequence
CHO	Chinese Hamster Ovary
CM	Conditioned Media
CSSR	Conservative Site Specific Recombination
CTD	C-Terminal Domain
DICE	Dual Integrase Cassette Exchange
FACS	Fluorescence-Activated Cell Sorting
FBS	Fetal Bovine Serum
FLP	Flippase
FM	Fresh Media
FRT	Flippase Recombination Target
GFP	Green Fluorescent Protein
GOI	Gene Of Interest
GPCR	G-Protein Coupled Receptor
HR	Homologous Recombination
HTH	Helix-Turn-Helix
IHF	Integration Host Factor
Kb	Kilobase Pairs

LSR	Large Serine Recombinase
NTD	N-Terminal Domain
OD	Optical Density
OpMNPV	*Origyia pseudotsugata* Nucleopolyhedrovirus
ORF	Open Reading Frame
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
RD	Recombinase Domain
RDF	Recombination Directionality Factor
RFP	Red Flourescent Protein
RMCE	Recombinase Mediated Cassette Exchange
SOE-PCR	Splicing Overlap Extension Polymerase Chain Reaction
SR	Serine Recombinase
SSR	Site Specific Recombinase
TAE	Tris base, Acetic acid, Ethylenediaminetetraacetic acid
TCC	Total Cell Count
TPA	Tissue Plasminogen Activator
TR	Tyrosine Recombinase
TRD	Terminal Repeat Domain
UTR	Untranslated Region
VLP	Virus-Like Particle
wt	Wild Type
ZD	Zinc Domain

Chapter 1

Introduction

1.1 Production of Recombinant Proteins Through Animal Cell Culture

The field of biotechnology might easily be thought of as a modern one, but its roots extend back all the way to 1907. It was in that year that Ross Harrison first demonstrated of cell growth *in vitro*, by maintaining frog cells in a hanging drop for as many as four weeks (Harrison, Greenman, Mall, & Jackson, 1907). It would take several decades for this discovery's ramifications to cement, but the field would then proceed rapidly. The 1949 discovery that poliovirus could be cultured *in vitro* quickly led to the first industrial use of animal cell culture – the production of the Salk poliomyelitis vaccine in 1954 (Enders, Weller, & Robbins, 1949; Kretzmer, 2002). Soon, there began a wave of vaccines produced by way of animal cell culture, including Foot-and-Mouth, Measles, Rabies, Mumps, and Rubella (Kretzmer, 2002). But these pharmaceuticals, while landmarks in their own right, were all "native" biologics. What about more complex or recombinant products?

Beginning in the 1980's, recombinant protein production in mammalian cell culture became a reality. Human Tissue Plasminogen Activator was produced by inserting the tPA gene into Chinese Hamster Ovary (CHO) cells, and it became the first licenced recombinant protein produced in animal cell culture (Lubinieki 1989). While CHO cells are still the most frequently used host in the mammalian cell culture space, others such as human embryonic kidney (HEK-293), human-retina-derived (PER-C6), or baby hamster kidney (BHK) cells are common as well (Butler, 2005). With respect to producing pharmaceuticals for human

consumption, these mammalian platforms offer advantages such as precise glycosylation and post-translational modifications of the product.

However, many recombinant protein products can be accurately reproduced outside of a mammalian cell. In such cases, insect cell culture is a more economically attractive option. Additionally, insect cell culture provides access to high-output recombinant protein production through a virally mediated system called the BEVS.

1.2 The Baculovirus Expression Vector System

BEVS is a popular platform for recombinant protein expression. The BEVS consists of two main components – the *Autographa californica* Multicapsid Polyhedrovirus (AcMNPV) and a eukaryotic (often lepidopteran-derived) host cell.

AcMNPV is a large enveloped dsDNA virus, belonging to the *Baculoviridae* family (Jehle et al., 2006). The virus's native hosts are invertebrates from orders such as *Diptera*, *Hymenoptera*, and *Lepidoptera* (Jehle et al., 2006). The Baculovirus lifecycle is split into 3 distinct phases: early phase (0-6 hpi), late phase (6-24 hpi), and very-late phase (24-72 hpi). During the very-late phase, AcMNPV is known to produce distinct occlusion bodies. Two genes in particular, polh and p10, contribute heavily to the development of this phenotype by contributing the lion's share of the matrix in the occlusion bodies (van Oers, Malarme, Jore, & Vlak, 1992). The polh and p10 promoters are considered to be some of the strongest promoters currently known. The BEVS derives its utility by leveraging the power of these two promoters, and the fact that the genes they drive are not obligatory for replication in cell culture (Kost,

Condreay, & Jarvis, 2005; van Oers, Pijlman, & Vlak, 2015). By replacing the coding sequences (CDS) of either the *polh* and/or *p10* genes with a gene of interest (GOI), extraordinary amounts of GOI-coded protein can be obtained upon infection of a host cell with the recombinant baculovirus (Smith, Summers, & Fraser, 1983).

Since the inception of the BEVS, significant changes have been made to the AcMNPV genome in an effort to improve the system's ease-of-use, viral stability, as well as the quality and quantity of the protein products (Kitts, Ayres, & Possee, 1990; Pijlman, van Schinjndel, & Vlak, 2003). These improvements have led to increased use of the BEVS in research applications, as well as industrially to produce proteins, Virus-Like Particles (VLPs), and viral vectors for both human and veterinary applications.

While AcMNPV has a variety of potential hosts, there are two in particular that are commonly used in industrial applications. Ovarian tissue samples of *Spodoptera frugiperda* and *Trichoplusia ni* gave rise to the cell lines Sf9 and Hi-5, respectively (Hink, 1970; Vaughn, Goodwin, Tompkins, & Mccawley, 1977). The Sf9 cell line is a clonal isolate derived from the original *S. frugiperda* line Sf21, and is adapted to grow in serum-free media. The Sf9 line can be grown both adherently or in suspension, and is suitable for both transfection and recombinant protein expression. However, despite the significant advances made to the BEVS through engineering the AcMNPV genome, less attention has been paid to how the Sf9 genome might be optimized.

1.3 Genome Engineering via Site Specific Recombination

DNA recombination refers to any reorganization of two or more strands of DNA into a new configuration. This can occur naturally during the routine life-cycle of an organism, such as instances of homologous recombination, or can be performed in a laboratory setting where the process is termed "cloning."

Conservative site-specific recombination (CSSR) is a particular method wherein regions of DNA can be excised, inverted, integrated, or otherwise translocated based on sequence-specific criteria. This form of recombination is typified by three important factors: i) the recombination event is catalyzed by a specialized enzyme, a site-specific recombinase (SSR); ii) recombination takes place only at unique sequence-specific loci; and iii) recombination does not require DNA synthesis in order to proceed (Grindley, Whiteson, & Rice, 2006).

There are two main classes of SSRs, the tyrosine recombinases (TR) and the serine recombinases (SR), so identified by their catalytic amino acid residue (Brown, Lee, Xu, & Smith, 2011). Each of these can be further partitioned into two subgroupings; TRs are divided broadly by their directionality, whereas SRs are done so by size (Wang, Kitney, Joly, & Buck, 2011). While these classes of SSRs are evolutionarily and mechanistically distinct, they follow analogous principles of operation. The SSR(s) recognizes and binds to two sequence-specific regions of DNA, forming the synaptic complex. While in the synaptic complex, the SSR cleaves, exchanges, and recombines the segments of DNA. Once the recombination is complete, the synaptic complex dissolves (Grindley et al., 2006).

Genome editing is more challenging than *in vitro* cloning. In most cases, genome editing harnesses the natural DNA repair mechanisms of a cell to make site specific changes

via homologous recombination (HR) (Camerini-Otero & Hsieh, 1995). This technique, while serviceable, is extremely inefficient when compared with SSR techniques (Groth, Olivares, Thyagarajan, & Calos, 2000; Mumm, Landy, & Gelles, 2006).

In the past thirty years, SSRs have become a useful tool in engineering plant species. They can be used in conjunction with HR (such as to remove selectable markers after HR), or to directly introduce transgenes into plants such as thale cress, tobacco, and rice (Choi, Begum, Koshinsky, Ow, & Wing, 2000; Dale & Ow, 1991; Louwerse et al., 2007; Srivastava & Ow, 2001).

1.4 Hypothesis

Incorporating CSSR capabilities from Phage C31 into the Sf9 genome will allow for rapid and stable gene integration at an individual genomic locus via ϕC31-Integrase mediated cassette exchange.

1.5 Research Objectives

1.5.1 Design

The design phase sought to identify key features that a successful SSR-competent Sf9 cell line should exhibit. Potential genetic components were evaluated based on previously reported studies and/or current experimentation. A gene architecture was then designed to accommodate as many system features as possible, such as orthogonality and irreversibility.

1.5.2 Construction

A series of *in vitro* DNA cloning techniques, such as Splicing Overlap Extension Polymerase Chain Reaction (SOE-PCR) and Gibson Assembly, were used in an effort to construct the gene architecture.

1.5.3 Implementation

The completed gene architecture would then be integrated into the Sf9 genome via piggyBac transposon technology (a transposon derived from *T. ni)*. A limiting dilution protocol will be used to isolate monoclonal cell lines from the resulting culture of SSR-competent cells.

Chapter 2

Literature Review

2.1 Site Specific Recombination

There are an estimated $2.9x10^{29}$ bacteria currently living on planet Earth (Kallmeyer, Pockalny, Adhikari, Smith, & D'Hondt, 2012). While this is an astounding number, an even larger number is required to describe the Earth's population of viruses. Viruses are the most numerous genetic beings in the world, with over 10^{30} in the oceans alone (Suttle, 2007). Given the size of the world's bacterial population, it is unsurprising that the majority of the planet's viruses are bacteriophages; they outnumber bacteria 10 to 1 (Hambly & Suttle, 2005). Two types of bacteriophage lifecycles exist: lytic and lysogenic. The lytic cycle, involves a relatively straightforward itinerary of infection, replication, and lysis of the host cell. Under some circumstances, bacteriophage can engage in a lysogenic cycle. During this process, the viral genome is integrated directly into the host cell's DNA. The virus, now called a prophage, can lay dormant until triggered to activate by environmental conditions, at which time the prophage DNA is excised from the host and the lytic cycle begins again (Ptashne, 2004).

This integration and excision of the phage genome is mediated by CSSR. The phenomenon of CSSR was initially described by Allan Campbell in 1962. He observed that the integration of *Escherichia coli* phage λ required a cross-over event between a specific DNA attachment site in the bacteria (attachment Bacteria, or attB), and a corresponding site in the phage (attachment Phage, or attP) (Campbell, 1963). The enzyme responsible for orchestrating

this cross-over, λ Integrase (λInt), was isolated in 1974, and would become the prototypical specimen for CSSR (Nash, 1974).

CSSR can also be used by mobile DNA elements throughout the bacterial and archaeal kingdoms, on plasmids and transposons (Kostrewa et al., 1991; Phoebe A. Rice, Yang, Mizuuchi, & Nash, 1996). But despite their near ubiquity in the "lower" kingdoms, comparatively few instances of native CSSR are found in eukaryotes.

The products of SSR-mediated reactions are often dependant on the orientation of the DNA substrates. Linear recombination occurs in the event of a single cross-over between two independent linear DNA molecules, but in the event that one or both attachment sites lie on a circular DNA molecule, SSRs will mediate an integration reaction. Should two attachment sites be found on the same DNA molecule, a cognate SSR will facilitate either an excision or inversion reaction depending on whether the attachment sites are oriented head-to-tail or head-to-head respectively (Nash, 1996).

Cognate attachment sequences can be identical or non-identical depending on the particular SSR. Systems utilizing non-identical attachment sites (ex. λInt, ϕC31Int) are typically differentiated by using the terms attB (attachment-Bacteria) & attP (attachment-Phage) for the substrate sequences, and attL (attachment-Left) & attR (attachment-Right) for the product sequences. attL and attR refer to the standard orientation of the post-recombination sequences, one on the left (attL, upstream) of a GOI, the other on the right (attR, downstream).

In the years since Campbell's foundational 1962 paper, a plethora of Site-Specific Recombinases have since been discovered and characterized. These enzymes are nearly all partitioned into two classes based on their catalytic amino-acid residue: the Tyrosine

Recombinases (in the past referred to as the λ Integrase Family), and the Serine Recombinases (Grindley et al., 2006; Nunes-Düby, Kwon, Tirumalai, Ellenberger, & Landy, 1998).

2.2 Tyrosine Recombinases

The Tyrosine Recombinases (TRs) are a diverse group of enzymes in terms of both origin and function. It was originally believed TRs were nearly exclusively found in prokaryotes, with the only eukaryotic example being from the species *Saccharomyces cerevisiae* (Chan, Liu, Ma, Jayaram, & Sau, 2013). However, this was refuted by the discovery of TR-active retrotransposons in *Dictyostelium discoideum, Panagrellus redivivus,* and *Phycomyces blakesleeanus* (Poulter & Goodwin, 2005). TRs perform a wide array of functions including controlling plasmid copy-number in *S. cerevisiae,* regulating the lysogenic vs lytic cycles of various phage, and resolving hairpin telomeres in *E. coli* phage N15 (Futcher, 1986; Ravin, 2011; Volkert & Broach, 1986). Over 1300 TRs from bacterial genomes have been discovered. However, only a few well-characterized TRs will be discussed here (Houdt, Leplae, Lima-mendez, Mergeay, & Toussaint, 2012).

The first, and most thoroughly characterized, TR is λ Integrase (λInt). However, two other key TRs, Cre Recombinase from phage P1, and Flippase (FLP) from *S. cerevisiae*, have since been developed and are nearly, if not equally, as well characterized (Y. Chen & Rice, 2003a; Grindley et al., 2006).

TRs are classified as such due to their use of a core tyrosine residue for catalysis. They also typically contain a highly conserved catalytic fold (Esposito & Scocca, 1997; Nunes-Düby et

al., 1998). This fold consists of five residues: Arginine, Lysine, Histidine, Arginine, and either Histidine or Tryptophan (Gibb et al., 2010). There is also evidence that a sixth conserved residue, Aspartic Acid/Glutamic Acid, indirectly stabilizes the catalytic site via additional hydrogen bonding (Gibb et al., 2010). Some small TRs, such as FimE and FimB, consist of only this catalytic domain (Bryan et al., 2006). However, small TRs such as these are uncommon. Most TRs contain at least one, if not multiple, auxiliary N-terminal domains to facilitate DNA binding.

Formation of the synaptic complex begins with the binding of recombinase monomers to their cognate recognition sequences. They do so on each side of a central region in the attachment sequence, delineated by the scissile phosphates – approximately 6-8 bp apart. A set of four bound TRs bring the two DNA molecules together and form a tetrameric synaptic complex (Grindley et al., 2006).

Strand exchange begins with single-strand cleavage of each DNA substrate, at their opposing scissile phosphates. This proceeds according to a type IB Topoisomerase mechanism, wherein the catalytic tyrosine performs a nucleophilic attack on a scissile phosphate, resulting in a 3'-phosphotyrosyl intermediate and a 5'-hydroxyl (Krogh & Shuman, 2000). Strands are exchanged within the synapse via the newly freed 5' hydroxyls attacking their opposing 3' phosphotyrosines, to form a Holliday junction intermediate. Of the four TRs in the tetrameric complex, only two are catalytically active at a time (Chen et al., 2000). This necessitates the two-step strand-exchange process typical of TRs. The Holliday junction isomerizes to activate the inactive recombinase monomers, and the process repeats (Guo, Gopaul, & Duyne, 1997). While full scale recombination only occurs in the context of a tetrameric synapse, there is

evidence that TR dimers can in rare cases initiate strand cleavage on their own (Ghosh, Lau, Gupta, & Duyne, 2005).

Highly conserved amino acid residues are expected to have important functional characteristics. Structural and mutational studies have elucidated some of these functions for the amino acids surrounding the nucleophilic tyrosine in the TR active site. The pair of immutable arginines (indexed 173 & 292 in Cre, 191 & 308 in FLP) are responsible for counteracting the negatively-charged non-bonding oxygens of the scissile DNA phosphate (Biswas et al., 2005; Gibb et al., 2010). The active-site lysine stabilizes the 5'-hydroxyl moiety produced during strand cleavage (Ghosh, Lau, et al., 2005). It has been observed in FLP that conserved histidine-305 operates in a basic fashion, drawing a proton from the core tyrosine, thereby priming it for the strand-cleavage reaction (Whiteson, Chen, Chopra, Raymond, & Rice, 2007). The semi-conserved His/Trp residue may be involved in either hydrogen bonding with the cleaved 5'-hydroxyl, or positioning of the catalytic tyrosine via hydrophobic interactions, and these functions may vacillate based on the particular TR (Chen & Rice, 2003b; Gibb et al., 2010; Ma, Kwiatek, Bolusani, Voziyanov, & Jayaram, 2007).

An interesting point of diversity in the TRs emerges when studying the assembly of the enzymes' active site. For most TRs (including λInt and Cre) , the active site is formed in a self-contained fashion, wherein all key components of the catalytic fold originate from an individual recombinase monomer. However, in other TRs (such as FLP and R), the active site is assembled at the confluence of two recombinase molecules. Observations of these *trans* active sites show the inactive member of a recombinase pair donating a catalytic tyrosine to the active member (Gibb et al., 2010; Yang & Jayaram, 1994). But despite these distinct

differences, the orientation of the catalytic tyrosine and sex associated auxiliary amino acids relative to one another is remarkably similar regardless of *cis* or *trans* assembly (Jayaram, 1997).

While some TRs, such as Cre and FLP, recombine two identical DNA binding regions, others such as λInt act on two distinct attachment sites. Nevertheless, it has long been understood that the core sequences directly recombined through the Holliday junction intermediate must be perfectly homologous. This homology promotes a stable intermediate, and in its absence the Holliday junction resolves to a prior stage or simply does not form at all (Joonsoo Lee & Jayaram, 1995; Nunes-Düby, Azaro, & Landy, 1995). Recently, however, some examples of TRs bending the homology rule have been unearthed. The conjugative transposon CTnDOT, from *Bacteroides*, codes for a TR named IntDOT. IntDOT has a mechanism of action that results in crossover between 5 bp non-homologies, which are not resolved until further DNA replication takes place (Malanowska, Salyers, & Gardner, 2006). Filamentous bacteriophage CTXϕ eschews this convention in a slightly different manner. During integration into *V. cholera*, the phage's ssDNA genome forms a forked hairpin wherein a 3 bp non-homology region occurs between two active sites, XerC and XerD. This non-homology region prevents typical resolution of the Holliday junction, and the entire phage genome is then resolved into the host genome during DNA replication (Das, Martinez, Midonet, & Barre, 2013).

2.2.1 λ Integrase

λ Integrase is notable in the TR family for facilitating recombination between two distinct DNA attachment sites. The size and complexity of the attP site, 240 bp, is large when compared to other TRs. This large size is due to the presence of additional binding sites for both λInt as well as auxiliary proteins. In contrast, λ attB is only 21bp (Fogg, Colloms, Rosser, Stark, & Smith, 2014).

λInt binds not only the regions immediately flanking the 7 bp site of recombination (the Core-type binding sites which interact with the λInt C-terminal domain (CTD)), but also to 5 arm-type binding sites that bind with the λInt N-terminal domain (NTD). Interspersed between these Arm-type sites are loci for binding of 3 additional proteins: integrating host factor (IHF), Fis, and Xis (Landy, 1989) **(Figure 1)**. Binding of these co-factors are crucial in determining the efficacy and direction of recombination by λInt.

No matter whether recombination is integrative or excisive, λInt is central to the reaction. λInt consists of two domains, C-terminal and N-terminal, with each domain having a distinct structure and role. The λInt CTD contains the catalytic tyrosine residue, and is responsible for binding to core-type regions of both attP and attB sites (Tirumalai, Healey, & Landy, 1997). This domain strongly resembles the structure of other TRs, such as Cre and Flp, but is not capable of facilitating recombination in the absence of the λInt NTD. It is, however capable of resolving previously constructed Holliday junctions (Dibyendu Sarkar et al., 2002; Tirumalai et al., 1997).

The λInt CTD can be further partitioned into a catalytic domain (responsible for providing the nucleophilic tyrosine 342) and a central binding domain that allows greater sequence specificity. These two sub-domains are connected via a 17 amino acid (aa) linker (Aihara, Kwon, Nunes-Düby, Landy, & Ellenberger, 2003). Within the catalytic domain, the Tyr342 is positioned away from the catalytic fold when unbound. However, during the formation of a Holliday junction, the Tyr342 moves 20 angstroms, into alignment with the catalytic fold (Aihara et al., 2003; Kwon, Tirumalai, Landy, & Ellenberger, 1997). This shift coincides with an interprotomer conformational change wherein the β7 strand in the λInt CTD extends to interact with adjacent λInt CTD domains. Mutating or eliminating this region destroys recombination activity (Tekle et al., 2002).

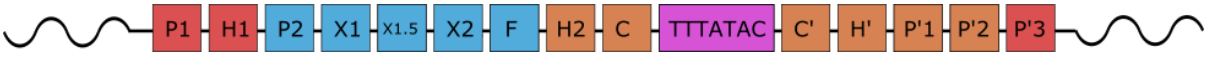

Figure 1 - λ attP Diagram

Gene diagram of the λ attP region. P sites (arm-type sites) interact with λInt NTD binding domain, whereas C sites (core-type sites) are bound by the λInt core binding domain. H sites are bound by IHF, X sites by Xis, and F by Fis. Red indicates sites necessary for integration, Blue for excision, and Orange for both. The magenta region contains base pairs involved in the formation of the Holliday junction. Figure adapted from (Seah et al., 2014). NTD, N-Terminal Domain; IHF, Integrase Host Factor.

The λInt NTD binds arm-type DNA regions via interactions between a three-stranded β-sheet structure and the major groove of the arm-type region (Wojciak, Sarkar, Landy, & Clubb, 2002). This type of interaction bears strong similarity to other DNA-binding domains such as in Tn916 integrase, or GCC-box Binding Domain in *Arabidopsis thaliana* (Allen, Yamasaki,

Ohme-Takagi, Tateno, & Suzuki, 1998). An 11 aa chain that precedes the β1 strand further stabilizes binding by interacting with the arm-type site minor groove, though is unstructured when not bound (Fadeev, Sam, & Clubb, 2009).

Given that the presence of the λInt NTD is the most striking difference between λInt and its monovalent counterparts, its role in λ recombination is curious. A 2001 study found that, in an extremely counterintuitive result, λInt CTD exhibited increased levels of DNA binding and topoisomerase activity in the absence of the NTD (Sarkar, Radman-Livaja, & Landy, 2001). The λInt NTD appears to counteract this shortcoming by providing λInt with its trait of directionality. Linking the λInt NTD to a Cre Recombinase was found to convert Cre from a bidirectional recombinase to a unidirectional one. Moreover, it allowed Cre functionality to be directly influenced by λ accessory proteins Integration Host Factor (IHF), Xis, and Fis (Warren, Laxmikanthan, & Landy, 2008).

IHF is a host-coded heterodimeric protein required for successful λ recombination. While it serves a variety of roles *in vivo*, including contributions to gene regulation and DNA replication, its primary role in λ recombination is to control the 3-dimensional arrangement of the attP DNA (Ryan, Grimwade, Camara, Crooke, & Leonard, 2004; Sze, Laurie, & Shingler, 2001). Bending of the attP DNA allows λInt to properly bind to both core-type and arm-type binding regions, and even small changes in accuracy of this IHF-facilitated bending can drastically reduce recombination efficiency (Nunes-Düby, Smith-Mungo, & Landy, 1995).

Xis is a phage-coded protein responsible for excessive recombination by λInt. Xis binds in triplicate to 3 closely spaced regions of attR titled X1, X1.5, and X2. Xis binding induces a

DNA bend of 72°. The distinct DNA topology produced by Xis binding, allows for tetrameric λInt formation despite the difference in DNA substrate sequences (attB/P vs attL/R) (Abbani et al., 2007).

Fis, like IHF, is a host-encoded protein that helps control DNA orientation during λ recombination. Notably, it increases the efficacy of excisive recombination by up to 20 fold in situations where Xis concentration is low (Thompson, Vargas, Koch, Kahmann, & Landy, 1987). Conversely, excision drops by as much as 100 fold when Fis is absent (Papagiannis et al., 2007).

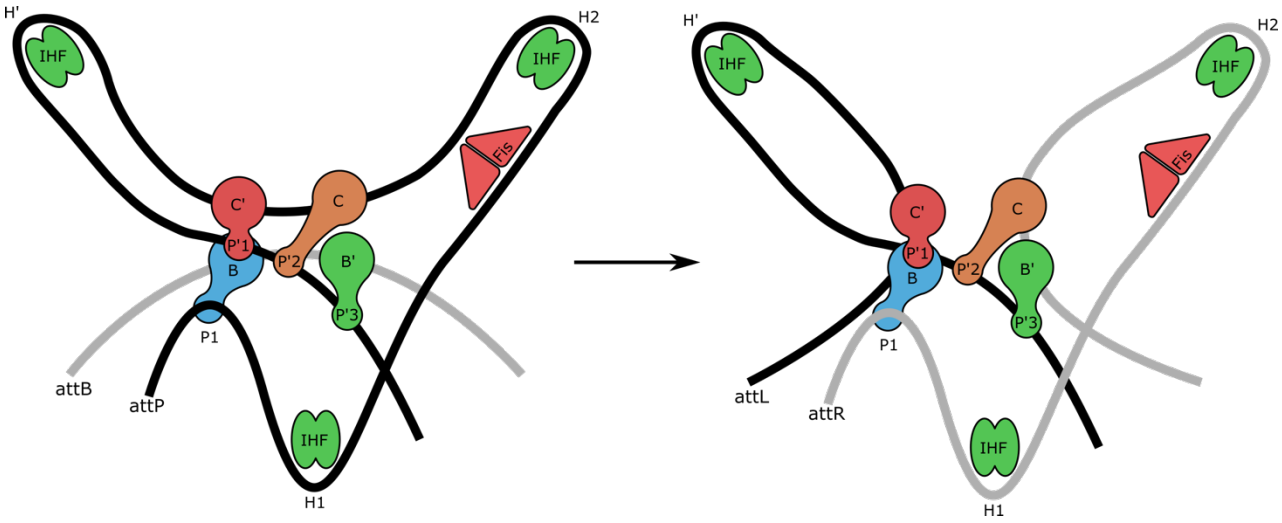

Figure 2 – λ Recombination Topology

λ integration takes place under precise topological constraints. Proteins such as IHF play a crucial role in mediating this DNA topology so that integrative recombination can take place. This diagram shows the DNA topology before and after recombination. The Holliday Junction (not shown here), occurs central to the 4 λInt monomers. Figure adapted from (Seah et al., 2014). IHF, Integrase Host Factor.

The interactions between these auxiliary proteins and λInt result in a complex and distinct topology of DNA and protein that facilitates recombination **(Figure 2)**.

The landmark application of λInt from a biotechnology or synthetic biology standpoint was the development of Gateway cloning. Gateway technology was first described in 2000, as a major advancement for *in vitro* cloning (Hartley, Temple, & Brasch, 2000). The first component of this system is the Entry Vector, with a GOI flanked by orthogonal attL sites and a backbone coding for kanamycin resistance. The second component, the Destination Vector, contains the gene ccdB flanked by orthogonal attR sites and a backbone coding for ampicillin resistance (Hartley et al., 2000) **(Figure 3)**. The application of a solution containing λInt, IHF, and Xis, instigates the movement of the GOI from the Entry Vector to the Destination Vector via λ Recombination. The selection parameters, with differential backbone selection as well as counter selection by ccdB (a gene that results in *E. coli* death via double-stranded DNA breakage) allow for excellent cloning efficiency (Bernard & Couturier, 1992).

This technology has since been leveraged to produce instances of multisite cloning, by chaining together orthogonal recombination sites, as well as instances where multiparallel expression libraries can be kept for both *E. coli* and *S. cerevisiae* (Cheo et al., 2004; Freuler, Stettler, Meyerhofer, Leder, & Mayr, 2008; Giuraniuc, MacPherson, & Saka, 2013).

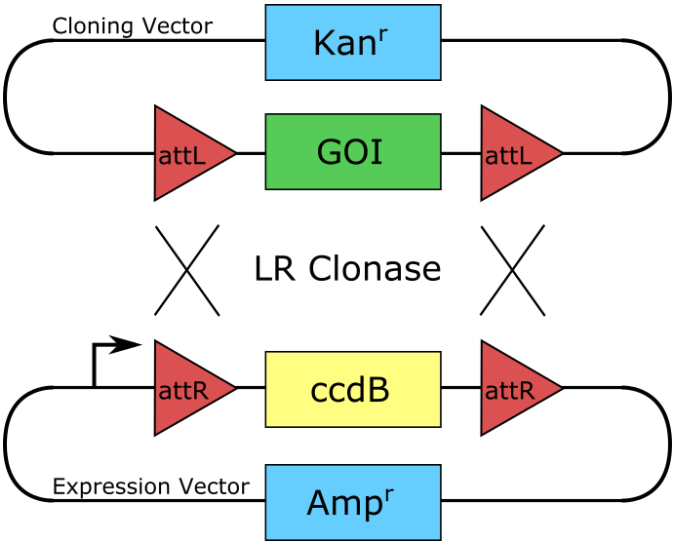

Figure 3 – Gateway Cloning Diagram

Gateway Cloning was one of the first industrially relevant uses for Site-Specific Recombinase technology. It allowed genes in a cloning library to be easily transferred to an expression library via the application of LR Clonase – a mixture containing λInt as well as all necessary auxiliary proteins. Figure adapted from (Hartley et al., 2000).

2.2.2 Cre Recombinase

Over a decade after the discovery of λ Integrase, a new phage-derived recombinase was purified and characterized in phage P1 (Abremski & Hoess, 1984). This recombinase was dubbed Cre, for **C**auses **re**combination.

In contrast with phage λ, the genome of phage P1 does not integrate into the *E. coli* genome during a lysogenic phase. Instead, the linear P1 DNA circularizes via Cre-recombination and maintains itself as a low-copy plasmid within the host cell. Additionally, Cre is used to resolve topologically linked plasmids caused by replication of the viral DNA.

Cre binds to a single DNA locus, titled *loxP*. *loxP* is a 34 bp sequence consisting of a unique 8 bp recombination region flanked by two 13 bp perfect inverted repeat regions **(Table 1)**. Like all Tyrosine Recombinases, recombination occurs via a Holliday junction intermediate, with cleavage occurring at indices 14 and 20 of the *loxP* region (Hoess, Abremski, & Sternberg, 1984).

The structure of Cre has been elucidated through several crystal structures examined at various points through the Cre-*loxP* interaction process (Gibb et al., 2010; Guo et al., 1997; Martin, Pulido, Chu, Lechner, & Baldwin, 2002). Cre can be split into two domains, a 130 aa N-terminal domain, and a 211 aa CTD. The former consists of five alpha-helices (A-E) involved in DNA binding, but not catalytic activity (Ghosh, Chi, Guo, Segall, & Van Duyne, 2005). In fact, Cre recombinase can maintain its function even when as many as 20 amino acids are eliminated from the NTD (Rongrong, Lixia, & Zhongping, 2005). The C-terminal domain contains the entire conserved catalytic fold, including the central catalytic tyrosine residue which in Cre is indexed to amino acid 324.

loxP binding initiates a C-clamp formation in Cre wherein the NTD binds to the major groove with helices B and D, and the CTD binds both major and minor grooves on the opposing side. The interactions between the Cre protein and the *loxP* DNA are, in some cases, augmented by hydration effects (Baldwin et al., 2003).

Table 1 – Variations of the Cre-Lox System

Directed evolution experiments have resulted in the development of several variants of the Cre-Lox pairing, in which the Cre amino-acid sequence and Lox base-pair sequences have been altered. Regions highlighted in red indicate involvement in a Holliday junction.

Cre Variant	Site Name	Site Sequence
Cre	*loxP*	ATAACTTCGTATA**ATGTATGC**TATACGAAGTTAT
Fre	*loxH*	ATATATACGTATA**TAGACATA**TATACGTATATAT
$Cre_{2(+/-)\#4}$	*loxM7*	ATAACTCTATATA**ATGTATGC**TATATAGAGTTAT
$CRE\text{-}_{R3M3}$	*loxK2*	GATACAACGTATA**TACCTTTC**TATACGTTGTTTA
Brec1	*loxBTR*	AACCCACTGCTTA**AGCCTCAA**TAAAGCTTGCCTT

Two Cre monomers will bind cooperatively to a *loxP* region, causing a 100° bend in the recombination region of *loxP* (Ennifar, Meyer, Buchholz, Stewart, & Suck, 2003). Subsequently, a tetrameric synapse is formed, bringing together 4 Cre and 2 *loxP* sites. The ensuing strand-exchange via Holliday junction takes place in a manner typical of the TR family.

The Cre-Lox system began to grow in popularity as a tool for genetic engineering in mammalian cells, due to its efficiency and thermostability. It has been suggested that this efficiency is in part due to the highly stable synaptic conformation formed between Cre and *loxP*. The bound Cre-*loxP* conformation has a dissociation constant of 10 nM, though this stability is pH dependant and falls precipitously at pH values above 8.5 (Ghosh, Guo, & Van Duyne, 2007). Cre recombinase shows optimal excision activity at 37°C, compared to a

recombinase like FLP which shows optimal excision at 30°C (Buchholz, Ringrose, Angrand, Rossi, & Stewart, 1996).

In particular, the Cre-Lox system has been used extensively for *in vivo* gene-knockout experiments in mice. This protocol uses Cre recombinase in an excision reaction to eliminate genomic loci flanked by *loxP* sites, and has been used to study a wide variety of model biological functions including renal and cardiopulmonary disorders (Li et al., 2000; Rojek, Füchtbauer, Kwon, Frøkiaer, & Nielsen, 2006). More recently, Cre-Lox technology has seen preliminary usage in pigs (Li et al., 2014).

In an effort to expand the efficacy and functionality of the Cre-Lox system, numerous variants of the Cre-*loxP* pairing have been developed **(Table 1)**. The first attempts at expanding the Cre toolbox used directed evolution to develop novel recombinase-DNA pairs. This resulted in the first variant, Fre-*loxH*, which differs from the wild-type by 4 nucleotides in the inverted repeats, as well as significant changes to the recombination region (Buchholz & Stewart, 2001). Shortly after the development of Fre-*loxH*, came a third variant titled $Cre_{2(+/-)\#4}$-*loxM7* (Santoro & Schultz, 2002). More recently, computational biology has made concerted engineering of recombinases attainable. By identifying 4 mutationally active regions a new Cre descendant was developed, titled Brec1-*loxBTR*. This particular recombinase may have implications for excising HIV-1 provirus in humans (Karpinski et al., 2016).

2.2.3 FLP Recombinase

The majority of industrially or academically useful recombinases are derived from phages, such as λ, P1, and ϕC31. Flippase (FLP) is one of the primary exceptions to this pattern, and furthermore is of the rare recombinases to be derived from a eukaryotic source: *Saccharomyces cerevisiae.* Wildtype FLP interacts with a region of DNA dubbed Flippase Recombination Target (FRT). The Flippase Recognition Target is a 34 bp sequence, consisting of a unique 8bp recombination sequence, flanked by a set of 13 bp inverted repeats. This FRT sequence is sufficient for excision and inversion reactions, but it has been shown that integration protocols benefit greatly from duplicating one of the inverted repeats upstream of the recombination sequence (Lyznik, Rao, & Hodges, 1996). Being a monovalent Tyrosine recombinase, FLP largely operates on the same mechanistic principles as Cre recombinase.

The gene coding for Flippase was first identified as one of four genes found on the 2μ Plasmid in *S. cerevisiae* (Andrews, Proteau, Beatty, & Sadowski, 1985). The 2μ Plasmid is a selfish genetic element, wherein none of its genes have been shown to provide a benefit to the yeast organism. Rather, the genes of the 2μ element are strictly concerned with replicating and partitioning the selfish plasmid during S phase of mitosis, and maintaining a non-disruptive plasmid copy number during *S. cerevisiae's* resting stage (Chan et al., 2013). Copy number of this plasmid typically holds between 40-60; increases above that can be deleterious and even lethal to the host cell (Chen, Reindle, & Johnson, 2005).

The 2μ Plasmid replicates via a modified rolling circle mechanism, whereby part of the plasmid is inverted after DNA replication has begun resulting in replication forks that move in

the same direction **(Figure 4)**. The product of this is a large molecule that contains two full copies of the 2μ DNA that is then resolved into two distinct copies. Both of these reactions, the inversion and the resolution, are mediated by FLP and a corresponding set of parallel FRT sites on the 2μ Plasmid (Futcher, 1986).

The FLP monomer, like Cre, consists of two discreet domains. The NTD (residues 1-107), shows no apparent structural similarities to the Cre NTD, but seems to provide a similar function. The C-Terminal Domain contains the catalytic fold, including the archetypic Tyrosine residue (Y343). The two domains are connected by Helix D, a 28 aa amphipathic region that serves to not only bridge the NTD and CTD, but also to help stabilize the synaptic complex during recombination by interacting with adjacent NTDs in the tetrameric form of FLP (Chen et al., 2000).

FLP induces recombination with half-of-the-sites activity, wherein only two catalytic tyrosine residues are active at a time. A dual nucleophilic attack occurs via a type IB Topoisomerase mechanism, a Holliday Junction is formed, the junction isomerizes, a second dual nucleophilic attack occurs, and the junction is resolved via strand exchange. While this process seems identical to Cre Recombinase, there is one key aspect that differentiates Flippase from both Cre as well as all other well-characterized Tyrosine Recombinases. Within the synaptic complex, the catalytic Y343 residue is active in *trans*. Helix M (the location of Y343), is donated to an adjacent FLP monomer such that Y343 joins the catalytic fold of the adjacent monomer. Y343 residues that are inactive are seen to be disordered, and reside 10 angstroms away from their respective catalytic folds (Chen et al., 2000). FLP appears to show no preference for order of strand exchange once recombination has commenced (Jehee Lee,

Tribble, & Jayaram, 2000). The domain swapping of helices D and M helps ensure that the full synaptic complex is formed before any accidental nucleophilic attacks take place. It may also be beneficial should one FLP monomer become degraded partway through a reaction – another FLP monomer can help resolve it (Chan et al., 2013).

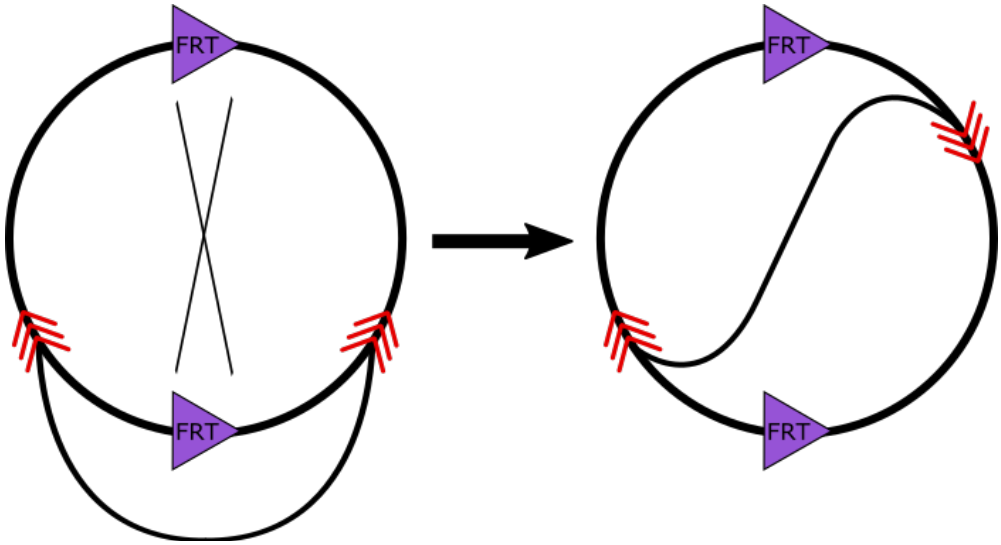

Figure 4 - 2μ Plasmid Replication

In its native state, FLP-FRT helps mediate replication of the 2μ plasmid. Typically, during DNA replication, the replication forks (red chevrons) move in opposite directions. By inducing site-specific recombination mid-replication, the 2μ plasmid is able to undergo rolling-circle replication wherein the replication forks move in concert with one another. Figure adapted from (Chan et al., 2013).

FLP in its native form operates optimally at 30°C (Buchholz et al., 1996). However, interest in using FLP in mammalian model organisms was high, and directed evolution of FLP soon led to a more thermostable protein called FLPe (e for **e**ighth generation). Relative to wildtype FLP, FLPe shows 4-fold increased recombination efficiency at 37°C, and 10-fold increased

recombination efficiency at 40°C (Buchholz, Angrand, & Stewart, 1998). A new variant, FLPo, has been developed that is codon-optimized for expression in mice. FLPo shows an order of magnitude increased efficiency relative to FLPe, and appears to be comparable to Cre Recombinase in mouse models (Raymond & Soriano, 2007). The development of FLPo has allowed the FLP-FRT system to be utilized in gene knock-out/in experiments in not only mouse models, but porcine species as well (Kranz et al., 2010; Lin et al., 2015).

Successful efforts have been made to orthogonalize the FLP-FRT system through directed evolution. These FLP variants bind to altered FRT sites, generically titled mFRT sequences. The initial development saw a single amino acid change in FLP (K82Y) result in preferential binding to sequence mFRT11 (Voziyanov, Stewart, & Jayaram, 2002). Subsequent work would see a variety of mFRT sequences developed with substitutions ranging from 1-5bp (Voziyanov, Konieczka, Stewart, & Jayaram, 2003). This allows multiple sites of recombination to be utilized by implementing combinations of FLP variants, reminiscent to the development of Fre-*loxH* from Cre-*loxP* (Konieczka, Paek, Jayaram, & Voziyanov, 2004). With the potential of developing novel FLP variants by directed evolution in mind, studies have begun to examine regions in the human genome that might be close enough to the native FRT sequence such that new Flippases could be developed to edit human DNA. In a preliminary study, over 600,000 regions show promise for such an endeavour (Shultz, Voziyanova, Konieczka, & Voziyanov, 2011).

2.3 Serine Recombinases

Five years after the isolation of λ Integrase, the study of bacterial transposon Tn3 uncovered a new protein with apparently similar function. Coded for by the *tnpR* gene, it took little time before it was realized that not only did this new recombinase have very distinct features relative to λInt, but that there were a large group of similarly featured enzymes (Heffron & Mccarthy, 1979; Kostriken, Morita, & Heffron, 1981). This family of enzymes would be named Resolvases, and the product of *tnpR* was dubbed Tn3 Resolvase (Plasterk, Brinkman, & Putte, 1983). A closely related recombinase from the γδ transposon (γδ Resolvase) became the archetypic member of this class when the mechanical distinctions between it and λInt were outlined in two 1981 papers by Reed and Grindley (Reed, 1981; Reed & Grindley, 1981). Shortly thereafter, it was observed that the enzymatic properties of this family were mediated principally by a conserved serine residue (Hatfull & Grindley, 1986; Reed & Moser, 1984). By the late 1990's, it became common to divide the recombinases into two categories, the Tyrosine Recombinases, and the Serine Recombinases. It is now understood that within the Serine Recombinases are two subclasses: The Small Serine Recombinases (also referred to as the Resolvase/Invertase family), and the Large Serine Recombinases (LSRs; sometimes referred to as Serine Integrases).

All Serine Recombinases (SRs) share a catalytically active core serine residue. Much like the TRs, there are a set of highly conserved amino acids that function within the SR active site (Grindley et al., 2006). However, unlike the TRs, the entirety of the SR recombination domain retains a consistent size throughout the family. SRs are multi-domain proteins, and while all tautologically contain the SR catalytic domain, they vary drastically in their ancillary domains

(Smith & Thorpe, 2002). It is, in fact, the variety and functional properties of these added domains that necessitates the distinction between the small and large families of SRs. The catalytic SR domain is nearly exclusively oriented at the N-terminus of the protein, though there is at least one example of a transposon coding an SR with a C-terminal catalytic domain (Kersulyte, Mukhopadhyay, Shirai, Nakazawa, & Berg, 2000).

For the purposes of this review, and because much of the pioneering research on SRs was performed on γδ and Tn3 Resolvases, the SR catalytic domain will be discussed primarily through the lens of the Small Serine Recombinases. A later discussion of the LSRs and their unique structures and properties will follow.

2.3.1 Small Serine Recombinases (Resolvases/Invertases)

The first SRs to be characterized (γδ and Tn3) functioned in the context of replicative transposition. The Tn3 transposon naturally forms a single cointegrant molecule during its replication. This cointegrant now contains two *res* recombination sites (the cognate binding sequence of Tn3 Resolvase). These *res* sequences allow for CSSR mediated by the Tn3 resolvase, thus "resolving" the cointengrant into two separate molecules.

Another well characterized member of the Small Serine Recombinases is the *Salmonella enterica* enzyme Hin. Hin, coded for by the eponymous *hin*, is a 190 aa invertase responsible for flagellar phase variation in *S. enterica.* By recombining two inverted imperfect palindromes (*hixL* and *hixR*), and thereby inverting the orientation of the SIGMA28 promoter, it effectively

controls expression of two alternate flagellar proteins (Osuna, Lienau, & Hughes, 1995; Silverman & Simon, 1980).

The γδ crystal structure was first published in 1990, and was further studied and refined in the following years (Rice & Steitz, 1994; Sanderson et al., 1990). Each γδ monomer is 183 amino acids long. The SR domain lies at the N-terminus, with a Helix-Turn-Helix (HTH) domain at the C-terminus. The two domains are connected by a short flexible linker peptide (Yang & Steitz, 1995). The SR domain consists of a set of α-helices encircling a core β-sheet. An extended 36aa α-helix, the E-helix (αE), protrudes from the C-terminal side (Yang & Steitz, 1995). The N-terminal region of this E-helix allows for dimerization of γδ resolvase in solution. The 10 aa linker peptide connects the αE to the HTH domain, passing through the DNA minor groove when bound to its cognate substrate (Nollmann, He, Byron, & Stark, 2004). This characteristic SR domain structure is conserved across the SR family.

The formation of the synaptic complex begins with γδ resolvase dimers binding to their associated *res* sequences. The C-terminal HTH domains bind along the major groove, each roughly 10 bp from the scissile phosphate. These domains are very likely uni-functional, given that an experiment replacing them with zinc-finger binding domains led to no change in recombinase function (Akopian, He, Boocock, & Stark, 2003). Meanwhile, the C-terminal region of the E-helix helps position the catalytic region through minor-groove contact (Yang & Steitz, 1995).

The synapse itself consists of two DNA *res* sites, and four SR units in a tetrameric formation. While these ingredients sound reminiscent of the TR synaptic complex, that is where the similarities end. The SR synapse is arranged with the DNA crossover regions on the outside,

with the SR tetramer bound together at the synapse core (Li, Kamtekar, Xiong, & Sarkis, 2005). While the full tetrameric crystal structure is now available, this synaptic formation had been suggested earlier when X-Ray diffraction data suggested significant separation of the DNA strands (Nollmann et al., 2004).

Upon completion of the synaptic complex, each of the four SR units performs a nucleophilic attack via catalytic Serine-10, thereby covalently linking the recombinases to the 5' ends of what are now four simultaneously broken strands (Reed & Grindley, 1981; Reed & Moser, 1984). This leaves each *res* sequence with a 2bp, 3' overhang. From this arrangement, a 180˚ rotation is engaged between each half of the synapse (Stark, Sherratt, Soocock, & Gil, 1989; Wasserman, Dungan, & Cozzarelli, 1985). This aligns the 5'-phosphoseryl moieties with the opposing 3'-hydroxyl groups, which then supplant the serine-10 residues, completing the recombination reaction. It is worth noting that the 2 bp overhangs are under strict conditions of homology; in the event that Watson-Crick pairing is unable to proceed, the synaptic complex performs a secondary rotation to realign the overhangs in their parental form. While this process prevents unwanted cross-over reactions, it interestingly imposes a new topology on the two DNA strands (Heichman, Moskowitz, & Johnson, 1991; Stark, Grindley, Hatfull, & Boocock, 1991).

2.3.2 Large Serine Recombinases

The Large Serine Recombinases (LSR) have a number of unique characteristics relative to the Resolvases, most notably their size. Where the smaller class of SRs are typically 180-200

amino acids long, the LSRs vary anywhere from 400-700 aa (Smith & Thorpe, 2002). Whereas the resolvases are hosted equally by both gram-negative and gram-positive bacteria, the LSRs are found disproportionately more often in gram-positives (Mullany, Roberts, & Wang, 2002; Smith & Thorpe, 2002). With respect to their role in the phage life cycle, LSRs function similarly to phage λ in *E. coli*. The recombinase protein allows for controlled integration and excision of the prophage into the host genome, allowing for strict regulation between the viruses lytic and lysogenic phases (Campbell, 1963). However, in contrast with phage λ, the LSRs require no auxiliary binding proteins to facilitate integration and require only a single Recombination Directionality Factor (RDF) to catalyze the reverse reaction (Rutherford & Duyne, 2014).

LSRs utilize the asymmetric attB/attP site specificity found in phage λ. In the wild, LSRs commonly, though not exclusively, integrate into intragenic attB sites. This often results in the silencing of the host gene, though in some cases the integrating DNA will contain translational signals to allow for continued gene expression such as during *SSCmec* mobilization in *Staphylococcus* (Misiura et al., 2013). **Table 2** provides a list of some of the best characterized LSRs and their wild-type locus of recombination.

While many serine recombinases are produced by active bacteriophages, others are coded for by domesticated prophages. Examples of these include SpoIVCA in *B. subtilis*, A118 in *L. monocytogenes,* and LI Integrase in *L. innocua* (Kunkel, Losick, & Stragier, 1990; Loessner, Inman, & Lauer, 2000; Rutherford, Yuan, Perry, Sharp, & Duyne, 2013). LI Integrase is of particular interest, as it was the first LSR to have a properly elucidated crystal structure (Rutherford et al., 2013).

The N-terminal catalytic domain of the LSRs shares close homology with that of the Resolvase family. They exist as protein dimers in solution, but come together to form a tetrameric synaptic complex in the presence of their DNA binding sites (Ghosh, Pannunzio, & Hatfull, 2005). And, just like the resolvases, elimination of the catalytic serine residue prevents recombination completely (Ghosh, Kim, & Hatfull, 2003; Smith, Till, & Smith, 2004).

Table 2 - Origins and Hosts of Select LSRs

The Large Serine Recombinases are derived from a wide variety of bacterial origins. A selection of these are listed, along with their native locus of integration.

Host	Phage/Prophage	Locus of Integration	Reference
M. smegmatis	Bxb1	*groEL1*	(Kim et al., 2003)
S. lividans	ϕC31	*SCO3798*	(Kuhstoss & Rao, 1991)
S. parvulus	R4	Acyl Co-A synthetase	(Shirai, Nara, Sato, Aida, & Takahashi, 1991)
M. tuberculosis	ϕRv1	*Rv1587c*	(Bibb & Hatfull, 2002)
S. coelicolor	ϕBT1	*SCO4848*	(Gregory, Till, & Smith, 2003)
L. innocua	LI prophage	*comK*	(Rutherford et al., 2013)

What makes the LSRs clearly distinct from their smaller cousins is the presence of a large CTD of roughly 300-500 aa in size (Rutherford & Duyne, 2014). The CTD is responsible for the high degree of sequence specificity exhibited by the LSRs, and is attached to the serine-containing NTD by the αE (Ghosh et al., 2005; Singh, Ghosh, & Hatfull, 2013). There are three discrete sub-domains within the CTD: a recombinase domain (RD) and a zinc-ribbon domain (ZD) which are connected by a short linker peptide, and a coiled-coil motif (CC) which emanates from the ZD (Rutherford et al., 2013). While crystal structures are lacking for many LSRs, sequence comparisons between LI Integrase and other LSRs suggests the insights gleaned there are transferable to the LSR class in general (Rutherford et al., 2013).

When in contact with an appropriate attP site, each CTD subunit exhibits specific binding behaviours. The RD binds to the 13 bp immediately emanating from the core dinucleotide, whereas the ZD interacts with a smaller 9 bp motif, at a distance of 3bp outside the RD binding region. While the RD subunit will bind to the attP site in absence of the ZD, it does so with significantly less affinity (Ghosh et al., 2005; Mcewan, Rowley, & Smith, 2009). Experiments using CC-deficient versions of LI Int have shown that the CC motif is not necessary for effective DNA binding (Rutherford et al., 2013).

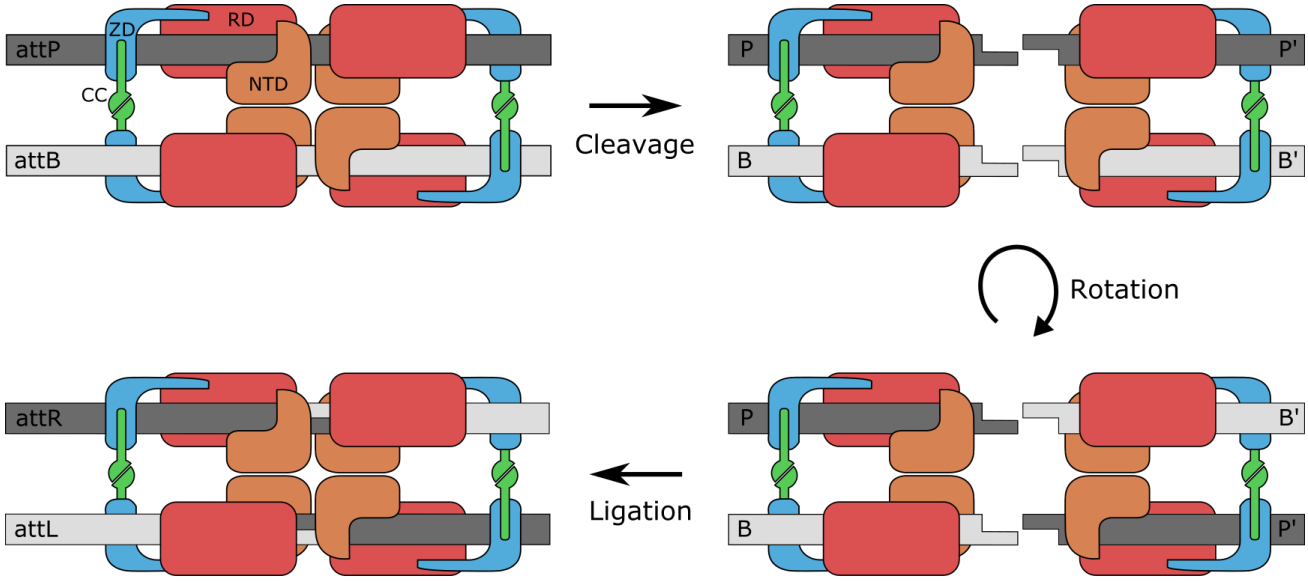

Figure 5 - LSR Recombination

The Large Serene Recombinases eschew the Holliday junction found in the Tyrosine Recombinases. Instead, the tetrameric complex causes a staggered break in each DNA molecule and then rotates to create a new alignment. These fragments are re-ligated, to form the resultant attL and attR regions. Figure adapted from (Rutherford & Duyne, 2014).

This paradigm shifts slightly during LSR-attB binding. The ZD motif appears to be shifted 5 bp towards the core dinucleotide (Rutherford et al., 2013). This has two important ramifications: the ZD binding motif is spatially shifted by one half-turn, and now overlaps the RD binding region through a span of 2 bp. This shift appears to be present in other LSRs, including ϕBxB1 and ϕC31, and provides a strong explanation of why attB sequences are typically shorter than their attP counterparts (Gupta, Till, & Smith, 2007; Singh et al., 2013).

2.3.3 ϕC31 Recombinase

Streptomycetes have been medically interesting for more than 80 years, and have produced some of the most useful antibacterial and antifungal agents ever developed. But in the 1980s, the field of investigation expanded to include the viruses that are associated with the Streptomyces genus.

Phage ϕC31 is a temperate virus that targets various members of *Streptomyces*, including *S. lividans, S. ambifaciens,* and *S. coelicolor*. ϕC31 is a double-stranded DNA virus. Its morphology consists of a polyhedral capsid with an associated elongated tail (Lomovskaya, Chater, & Mkrtumian, 1980). The genome is 4.5 kb in size, with 63% GC content, and is linear-cohesive (Sinclair & Bibb, 1988). Like all temperate phage, ϕC31 is capable of maintaining a lysogenic phase through prophage integration. While no crystal structure has yet been developed for ϕC31 Integrase, its close sequence similarity to LI Integrase suggest a nearly identical structure to its *Listeria*-sourced cousin discussed in section 2.3.2 (Rutherford et al., 2013).

Early biotechnological applications focussed on using engineered ϕC31 as a cloning tool for the *Streptomyces* platform. The characterization of the attP and attB sequences, as well as the first isolation of the ϕC31 Integrase would kick off the development of serine integrases as biotechnological tool for genetic engineering (Kuhstoss & Rao, 1991).

The first demonstration of ϕC31Int as a tool for recombinant genetics came in 1998, when it was observed to facilitate cointegration between two plasmids both *in vitro* and in *E. coli*. Additionally, it was shown that ϕC31Int functionality was not constrained by the topology of its DNA substrates (Thorpe & Smith, 1998).

After its performance *in vitro*, it was believed that ϕC31Int required no additional factors to facilitate attP/attB recombination. While this remains true of protein co-factors, it has been shown that trace levels of zinc are necessary for successful recombination. Pre-incubation of ϕC31Int with chelating agent EDTA prevented DNA binding by the protein. However, this phenotype was quickly restored by the addition of exogenous Zn^{2+} (McEwan, Raab, Kelly, Feldmann, & Smith, 2011). Additionally, zinc has also been shown to contribute to proper binding of ϕC31 Recombination Directionality Factor gp3 (Fogg et al., 2018).

ϕC31 Integrase has become popular for a variety of synthetic biology applications involving gene integration and cassette exchange. Recombinase Mediated Cassette Exchange (RMCE) is a process wherein paired sets of attB and attP sites recombine in such a way that results in the interchange of the intergenic regions between them. RMCE via ϕC31 Integrase has been successfully demonstrated in bacterial platforms, such as *Sinorhizobium meliloti*, *Ochrobactrum anthropi*, and *Agrobacterium tumefaciens* (Heil, Cheng, & Charles, 2012). In

insect cells, ϕC31 Integrase has been shown to mediate both standard integration events as well as RMCE events (Bateman, Lee, & Wu, 2006; Groth, Fish, Nusse, & Calos, 2004). Notably, ϕC31 Integrase was used to perform *in vivo* RMCE in *Bombyx mori* (a close cousin of *Spodoptera frugiperda)*, resulting in the exchange of the reporter DsRed with EGFP in the *B. mori* genome (Long et al., 2013).

2.4 *Spodoptera frugiperda* and the Sf9 Cell Line

Spodoptera frugiperda, colloquially known as the Fall Army Worm, is a small moth in the order Lepidoptera. It is common throughout eastern and central North America and is known predominantly for its role as a crop-destroying pest in the agricultural industry. However, in the biotechnology industry the species is most famous as the progenitor of the Sf21 and Sf9 cell lines.

By the mid-1970's, a variety of insect cell lines had been established from both Lepidopteran sources (such as *Trichoplusia ni)* as well as Dipteran (such as *Aedes aegypti)* (Greene, Charney, Nichols, & Coriell, 1972; Hink, 1970). In an effort to expand the repertoire of lepidopteran cultures, two preliminary cell-lines were established from *S. frugiperda*, initially titled IPLB-Sf-21, and IPLB-Sf-1254 (Vaughn et al., 1977). These strains initially required a highly undefined media for growth, but IPLB-Sf-21 was soon modified through progressive adaptation to subsist on a medium devoid of Bovine Serum Albumin or *Bombyx mori* hemolymph (Gardiner & Stockdale, 1975). This newly modified strain of *S. frugiperda* was titled IPLB-Sf-21-AE, or more commonly referred to as simply Sf21.

Tissue-derived lepidopteran cell lines are heterogenous in nature. This principle extends even to ploidy number, with the modal number of chromosomes sitting around 100 (Ennis & Sohi, 1976). Individual cells in the Sf21 line vary in size, but average at roughly 19.5 um in diameter (Vaughn et al., 1977). Progress made by Summers and Smith would help alleviate some of this heterogeneity by producing a clonal isolate of IPLB-Sf-21-AE, titled Sf9 (Summers & Smith, 1987). Along with Sf21, Sf9 would become a standard commercial offering in the field of insect cell culture, with Sf9 providing smaller more uniform cells and Sf21 providing slightly larger more variable cells (Invitrogen, 2002).

Sf9 cells provide a versatile platform for biotechnological applications. They can be grown both adherently as well as in suspension and can be grown both in serum-free media as well as serum-supplemented media. They can be transitioned seamlessly between adherent and suspension settings. However, movement from serum to serum-free media requires an adaptation period of up to 7 days (Invitrogen, 2002). Sf9 cells are grown optimally at 27°C, but will replicate comfortably in the range of 25°-30°C (Summers & Smith, 1987).

The primary application of the Sf9 cell line is to provide a host for the Baculovirus Expression Vector System. While the BEVS is capable of providing extremely high levels of recombinant protein production, it suffers from the inherent drawback of killing the host Sf9 cells. However, as early as 1990, there began research exploring the efficacy of using Sf9 as a platform for protein expression outside of the BEVS context (Jarvis, Fleming, Kovacs, Summers, & Guarino, 1990). Results showed that under certain circumstances, transformed Sf9 cultures outperform infected Sf9 cultures. While β-galactosidase production was inferior

by a factor of 1000, production of human tissue-plasminogen activator (TPA) was on par with the BEVS alternative. Furthermore, TPA was processed more efficiently than in BEVS due to the Sf9 secretory pathway being uninhibited by Baculovirus infection (McCarroll & King, 1997).

Proteins that require specific post-translational modifications to achieve optimal activity, such as membrane receptors, may benefit from this expression environment. Expressed Chick Nicotinic Acytelcoline Receptor proteins, though present in lower quantity, were shown to have increased binding efficiency when expressed constitutively in Sf9 cells (Atkinson, Henderson, Hawes, & King, 1996). In a more extreme example, Human β_2-Adrenergic Receptor was found to be non-functional when expressed through BEVS, but was expressed successfully in a stably transformed Sf9 cell line (Kleymann et al., 1993). Recombinant Factor C, natively produced in *Limulus* horseshoe crabs, was shown to be successfully expressed in Sf9 through stable transformation. The protein product was found to be fully functional, with 84% being secreted into the culture supernatant (Wang, Ho, & Ding, 2001).

G-Protein Coupled Receptors (GPCRs) are a class of proteins that exhibit a complex trans-membrane folding pattern. This makes them difficult to produce accurately in heterologous expression systems. However, transformed Sf9 cells have shown promise in this area. Human μ Opiod Receptor was successfully expressed under constitutive OpIE2 control in Sf9 cells, producing 11,000-15,000 active receptors per cell (Kempf et al., 2002). This result drastically outperformed production in yeast cells, and while receptors were produced in much higher

peak quantities (20-fold higher) in HEK293 cells, the expression was not stable and waned quickly (Stanasila, Pattus, & Massotte, 1998).

Stably transformed Sf9 cells are commonly pursued as an alternative to BEVS. However, they can also be used to compliment this system. Engineered strains of Baculovirus do not form occlusion bodies due to the elimination of their polh gene (they are genotypically occ-). Engineering of a host Sf9 line to independently produce polh during baculovirus infection successfully resulted in the production of phenotypically occ+, but genotypically occ-, baculoviruses (López, Alfonso, Carrillo, & Taboga, 2010). These virions improve safety and increase the potential uses for *in vivo* implementations of Baculovirus expression systems where oral infection of larvae is required.

Other host engineering efforts have focussed on extending the lifespan of an Sf9 cell through deferring apoptosis during Baculovirus infection. By the early 2000s, *Campoletis sonorensis* inchnovirus protein P-vank-1 had been shown to inhibit apoptotic processes in a variety of insect cells. In 2009 this technology was expanded to the Sf9 cell line. A recombinant line constitutively expressing P-vank-1 was shown to exhibit 5-fold increased viability relative to engineered Sf9, 5 days after infection with Baculovirus (Fath-Goodin, Kroemer, & Webb, 2009). Additionally, this line (titled VE-CLp1) was shown to have increased longevity under duress by UV radiation and camptothecin. The delay in cellular breakdown allows for the host cells to maintain their secretory pathway longer into the infection cycle. This results in improvements in both quantity and quality of protein produced in the BEVS-Sf9 system. VE-

CL02 (a clonal isolate of VE-CLp1) was shown to produce 3x the protein yield vs control, as well as improved accuracy in post-translational processing (Steele et al., 2017).

Of the promoters utilized by AcMNPV, only the IE1 promoter is active in uninfected *Spodoptera* cells (Jarvis et al., 1990). However, efforts have been made to incorporate the later, stronger, AcMNPV promoters into the Sf9 genome. While not constitutively active, genomic implementations of the 39K, p6.9, and polh promoters can allow for host-expression of proteins in synchronicity with the Baculovirus infection cycle (Lin & Jarvis, 2013).

2.4.1 Single Cell Isolation by Limiting Dilution

Given that locations of genomic integrations are difficult control, attempts to transform a monoclonal cell-culture for transgene expression typically result in a heterogeneous culture. In the case of insect cell culture, this eliminates many of the advantages of using Sf9 cells instead Sf21 cells. However, the clonal nature of a cell culture can be recouped through a procedure known as limiting dilution, which involves mechanically isolating individual cells from which a monoclonal culture is re-established. At its inception, limiting dilution used a statistical measure, Poisson's distribution, to estimate the likelihood of a derived culture being clonally isolated; though with improvements in phase-contrast microscopy the statistical estimate is no longer necessary (Staszewski, 1984). Cells are aliquoted to a concentration of <1 cell/aliquot, and typically plated in 96-well tissue culture plates. (Gross et al., 2015) This extreme dilution inhibits growth by also resulting in a dearth of secreted growth factors for many cell species,

and often requires the application of specialized or conditioned media collected from cultures in exponential growth phase. (Gross et al., 2015)

Isolating cells causes extreme stress, and results in high rates of cell death; it is reported that 50-75% of isolated insect cells will die in a successful limiting dilution protocol (Rodrigues, Costa, Henriques, Azeredo, & Oliveira, 2014). While not all insect cell-lines are capable of surviving the limiting dilution protocol, examples of both the Sf9 and High-Five cell lines have been shown to be successful. (Steele et al., 2017; Vidigal, Fernandes, Coroadinha, Teixeira, & Alves, 2014) The timeline for successful producing clonal isolates of a transformed Sf9 culture through limiting dilution has been clocked at between 2-4 months. (Rodrigues et al., 2014)

Chapter 3

Design Considerations

As the field of molecular biology has progressed, the research community has developed an interest in building biological systems from an *a priori* standpoint. This has resulted in an exciting new field of synthetic biology. In an effort to further this type of approach, this project aimed to bring an element of engineering design to the Sf9 cell line.

From the outset, the goal was to balance three primary elements: function, simplicity, and flexibility. For the final product, an RMCE-competent Sf9 cell line, to be most useful, it should provide efficient recombination with the least required number of steps (ex. transformations, selections) and be capable of facilitating as many types of experimental protocols as possible. The following section will address various design aspects and how they are balanced in the final genetic architecture.

3.1 Considerations on Recombinase Systems

The first issue to be considered is which of the Site-Specific Recombinases should be central to this design. While dozens of recombinases/integrases have been identified, significantly fewer have been well-characterized and demonstrated as useful from a biotechnology standpoint. “Degree-of-characterization” is not a quantitative measure, and thus a somewhat arbitrary cut-off was necessary when determining which recombinases would be considered in the design phase; the initial list of recombinases consisted of λ Integrase, Cre

Recombinase, Flippase, and ϕC31 Integrase. After evaluating these recombinase systems on the criteria discussed in sections 3.1.1-3.1.3, ϕC31 Integrase was deemed to be the most promising for use in the Sf9 platform **(Table 3)**.

Table 3 - Evaluation of SSR Systems

SSR	Simplicity	Orthogonality	Compatibility	Stability/Reversibility
Lambda	+	+	+	+++
Cre	+++	++	+	+
FLP	+++	++	+++	+
ϕC31	+++	+++	+++	+++

3.1.1 Stability and Reversibility

When genomic integration is the goal, the stability of integration is a key factor to consider. This hinges heavily on whether a recombinase system is reversible. In the cases of the Cre-*loxP* and FLP-FRT systems, the product sequence formed upon recombination is identical, or near-identical, to the substrate sequence. This stems from the fact that the regions of recombination include inverted repeats. These systems are deemed to be reversible.

Both the λ Integrase and ϕC31 systems operate by recombining two non-identical DNA sequences. Though the exact sequences differ for each system, these regions are termed attP and attB in both. The resultant regions of DNA are distinct from their precursors. These product sequences, attL and attR, are not products for further recombination unless influenced

by a Recombination Directionality Factor (RDF). ϕC31 Integrase has a well-characterized RDF, gp3, that allosterically regulates the direction of recombination (Khaleel, Younger, Mcewan, Varghese, & Smith, 2011). Because of this, they are deemed to be non-reversible recombination systems.

Non-reversible recombination systems provide added stability when compared with reversible systems because they are not subject to secondary recombination reactions while the recombinase is active in the cell. Through the lens of reversibility, both the λ and ϕC31 integrases provide an advantage over their compatriots.

3.1.2 Orthogonality

Recombinase Mediated Cassette Exchange (RMCE) is built on the principle of engaging multiple recombination reactions concurrently. One of the risks of this approach is the potential for crosstalk between recombination sites. This can result in a variety of unwanted results **(Figure 6)**.

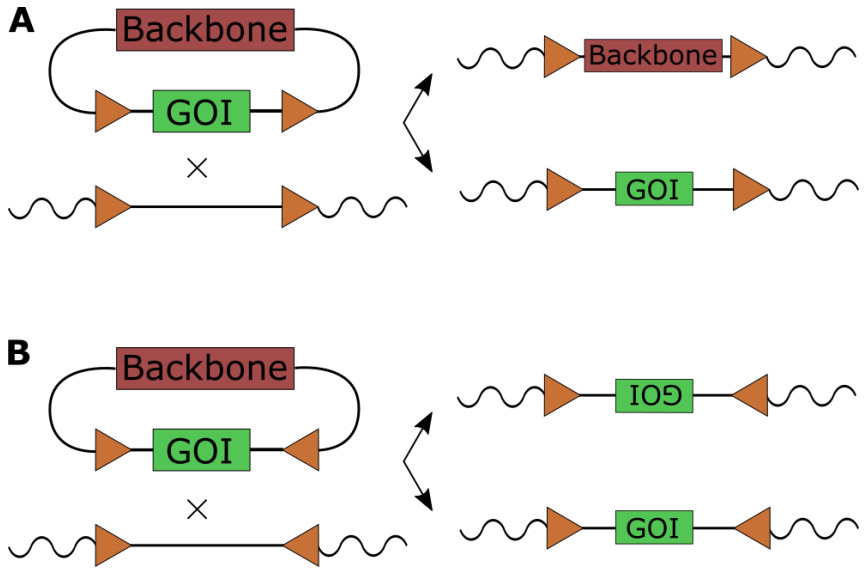

Figure 6 - Unwanted Results of RMCE

A) In parallel orientation, both the GOI and the vector backbone can be feasibly integrated through RMCE. B) When att-sites are in reverse orientation, the GOI is guaranteed to be integrated, but the orientation of the integrated cassette cannot be controlled. GOI, Gene of Interest.

In cases where the recombination sequences are identical, such as in Cre and FLP systems, genomic regions can react with themselves causing unwanted excision or inversion of the baseline region. If the system is one that requires non-identical recombination sites, the orientation of these sites can create ambiguity in the orientation of the integrated cassette or alternately result in risk of vector backbone integration.

An orthogonal recombinase system is one in which the regions of recombination are incapable of engaging in unwanted crosstalk (i.e. Sequence A will only recombine with Sequence B, while Sequence C will only recombine with Sequence D). Orthogonality has been pursued with the development of recombinase variants (ex. Fre-*loxH*) as discussed previously,

but also with an approach called dual RMCE (often listed variously in the literature as dRMCE, or DICE for Dual Integrase Cassette Exchange). This approach sees the use of two separate recombinases simultaneously to achieve orthogonality. This technique was first developed by combining the use of Cre and FLPe in embryonic stem cells, and showed a recombination efficiency of 14%, 7-fold higher than a previously reported RMCE experiment using a heterospecific Cre-Lox pairing (Lauth, 2002). This DICE approach appears to work in various formats provided the recombinases function in the host cell. It has been demonstrated successfully with Cre/ϕC31int pairing in mammalian DT40 cells, as well as with Cre/R pairing in tobacco plants (Dafhnis-Calas et al., 2005; Nanto & Ebinuma, 2008).

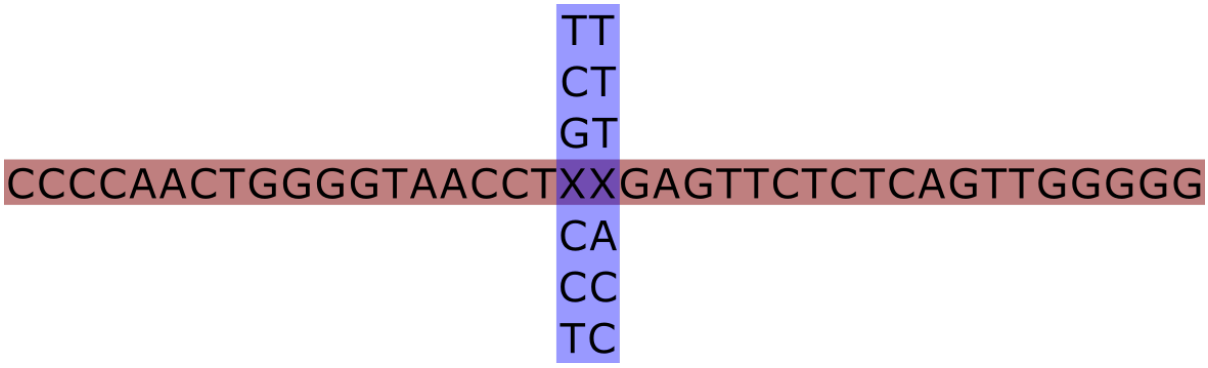

Figure 7 - ϕC31 attP Core-Dinucleotide Variants

In the ϕC31 attachment sequences, there is a two dinucleotide catalytic region. This results in 6 possible non-palindromic variations. Because proper Watson-Crick pairing is required for successful recombination, ϕC31-mediated RMCE can be made effectively orthogonal by utilizing different variations of these core-dinucleotides. The standard attP sequence regions are highlighted in red, while the di-nucleotide variants are listed orthogonally in blue. RMCE, Recombinase Mediated Cassette Exchange.

While the DICE method provides orthogonality, much like evolved heterospecific recombinase pairs, they require an unfortunate increase in complexity from utilizing multiple recombinases simultaneously. The ϕC31 may have an advantage in this regard. Variants of the ϕC31 attP and attB sequences have been engineered that provide orthogonality without the need for concurrently evolved integrase proteins (Colloms et al., 2014). Because the core dinucleotide involved in recombination is not responsible for binding specificity, but must nevertheless be identical for successful recombination, replacing this dinucleotide with any of the six possible variations creates an orthogonal system **(Figure 7)**. This provides a simple method of producing orthogonality within the confines of a single recombinase system, and shows no decrease in recombination efficiency with basic RMCE. The development by Colloms et al. makes the ϕC31 recombinase system particularly attractive.

3.1.3 Compatibility with Sf9 Platform

Regardless of any mechanistic advantages or disadvantages provided by a recombinase system, it must be operational in the context of the Sf9 cell line, which is cultured at 27°C. A crucial factor in this respect is the thermostability/active temperatures of these recombinases.

Both Cre and λ Integrases are derived from *E. coli* phages, and so unsurprisingly they exhibit optimal activity at 37°C. Wildtype Flippase, in contrast, operates best at 30°C, and will completely denature by 38°C (Buchholz et al., 1996). While successful efforts have been made to engineer FLP variants with peak activity at 37°C (FLPe, discussed previously), developing

variants of the *E. coli* phage integrases with low temperature activity has not been incentivized (Raymond & Soriano, 2007). While there does not appear to have been any study specifically directed to determining optimal temperatures for ϕC31 Integrase activity, it is derived from *Streptomyces* phages which are cultured from 28-30°C (Shepherd, Kharel, Bosserman, & Rohr, 2010).

Neither the λ nor the Cre recombinase systems have shown to be popular tools in insect platforms. In fact, Cre exhibited 0% transgene integration efficiency in both *D. melanogaster* and *A. aegypti* cells (Nimmo, Alphey, Meredith, & Eggleston, 2006). In contrast, the same study found integration efficiencies of up to 31.8% in the same settings using ϕC31 Integrase. There do not appear to be any concerted studies examining the *E. coli*-phage-derived recombinases in the Sf9 platform specifically.

Flippase has been utilized successfully in a variety of insect cell platforms, including the Sf9 cell line. A strain of Sf9 was developed to facilitate transgene integrations using a FLP landing pad. While this study showed positive results, recombination efficiency was extremely low; the most successful trial showed ~100 successful recombination events per 10^6 transfected cells, while others showed as few as 5 per 10^6 cells (Fernandes et al., 2012). Nevertheless, this technique was used to produce a stably transformed Sf9 culture producing rotavirus core-like particles (Fernandes et al., 2014). Despite its low rate of efficiency for transgene integration in Sf9, FLP has shown stellar results when used for excision reactions in the closely related lepidoptera *Bombyx mori*. *In vivo* excision of a previously integrated (via piggyBAC) eGFP cassette flanked with FRT sites was shown to be as high as 97.1% (Long, Lu, & Hao, 2016).

The ϕC31 recombinase system has also been validated for use in the Sf9 platform. Chompoosri et al. used a lacZ assay to demonstrate plasmid-to-plasmid ϕC31 RMCE in the Sf9 cell line (Chompoosri et al., 2009). This resulted in as many as 4100 successful exchange events per 10^6 cells, over an order of magnitude higher than the most successful trial in the Fernandes FLP cell-line. Nevertheless, it should be noted that this was a transient plasmid-to-plasmid experiment and it is not guaranteed that this degree of efficiency would be seen in a genomic integration.

While there has not been any genomic integration/RMCE events in Sf9 cells, there have been in *B. mori*. Single-site genomic integration was performed in a ϕC31-competent strain where *in vivo* integration efficiency was found to be 5.2% (Yonemura et al., 2013). Additionally, a full ϕC31-based RMCE competent line of *B. mori* showed successful *in vivo* exchange at a rate of 4.55% (Long et al., 2013).

Most Tyrosine recombinases are not compatible with the culture conditions of a lepidopteran cell line. Although Flippase orchestrated RMCE has been developed for Sf9, the degree of success is questionable. The ϕC31 Integrase system, while less proven, shows obvious potential to out-perform FLP in the Sf9 platform. It also provides stable cassette exchange with structurally simple orthogonality.

3.2 Mistargeting and Pseudo-attP Sites

One of the inherent risks involved with DNA editing is the potential for off-target effects. This occurs when an active recombinase binds to regions outside of its cognate binding site. Off-target effects can potentially have drastic deleterious effects to a cell, should an essential gene in the host cell be accidentally disrupted or excised. Locations of off-target integration are dubbed "pseudo-attP sites."

Off-target effects due to ϕC31Int implementation have been observed in S2 cells, though at very low levels (Groth et al., 2004). Only 3 instances of pseudo-attP integration were reported *in vitro*, though precise efficiencies were not reported. These results were not replicated *in vivo*. A trial of ϕC31 Integrase mediated gene integration in human cells showed that pseudo-attP sites within the human genome, though present and active, exhibited 8-17 fold fewer successful recombination events than a native attP site (Thyagarajan, Olivares, Hollis, Ginsburg, & Calos, 2001).

While no similar studies have been performed in Sf9 cells, an alignment search of ϕC31 attP against the current published Sf9 genome suggests no regions bear more than 41% sequence identity **(Figure E 1)**. This figure coincides with the identity levels reported in Groth et al. 2004 mentioned previously, wherein low-level *in vitro* recombination occurred at pseudo-attP sites with 41% identity relative to a native attP site, and thus significant off-target effects by ϕC31Int are not expected to be problematic in Sf9. A trial of ϕC31 RMCE into the closely related lepidopteran species *B. mori* resulted in no pseudo-attP integration events (Long et al.,

2013). However, a promoter trap could be used to fully select against off-target integration **(Figure 8)**.

3.3 Selection Markers and Reporters

In the construction and operation of the proposed genetic system, selection markers are needed to ensure that the genomic cassette has successfully integrated into the Sf9 platform, as well as to isolate cells that have successfully exchanged cassettes after RMCE.

Upon integration of the full construct in the Sf9 genome via piggyBAC (as discussed later), culturing in the presence of an antibiotic will provide a convenient avenue for selecting successful integrants. A variety of antibiotic selection markers have been shown to operate well in the Sf9 platform for the purposes of stable genomic integration, such as Hygromycin B, G418, Puromycin, and Zeocin (Douris et al., 2006). While all of these options are usable, a hygromycin resistance gene (Hygromycin phosphotransferase) CDS was included alongside the piggyBAC materials from Dr. Jarvis (University of Wyoming). Due to its availability, Hygromycin B was selected for use in this project.

Promoter traps have been used to great effect in previous instances of RMCE (Fernandes et al., 2012). A methionine deficient selection marker can be implemented immediately following the cassette exchange region. This gene will remain inactive until a supplementary promoter and methionine are introduced, via cassette exchange, to activate antibiotic resistance **(Figure 8)**. To streamline the RMCE process, it is desirable to have a strong, rapid-acting selection scheme associated with this promoter trap. Puromycin and Blasticidin both allow for selection

protocols of one week or less, but Blasticidin is cost prohibitive; Puromycin was therefore selected for the promoter trap selection scheme (Thermo Fisher Scientific Inc., Mississauga ON).

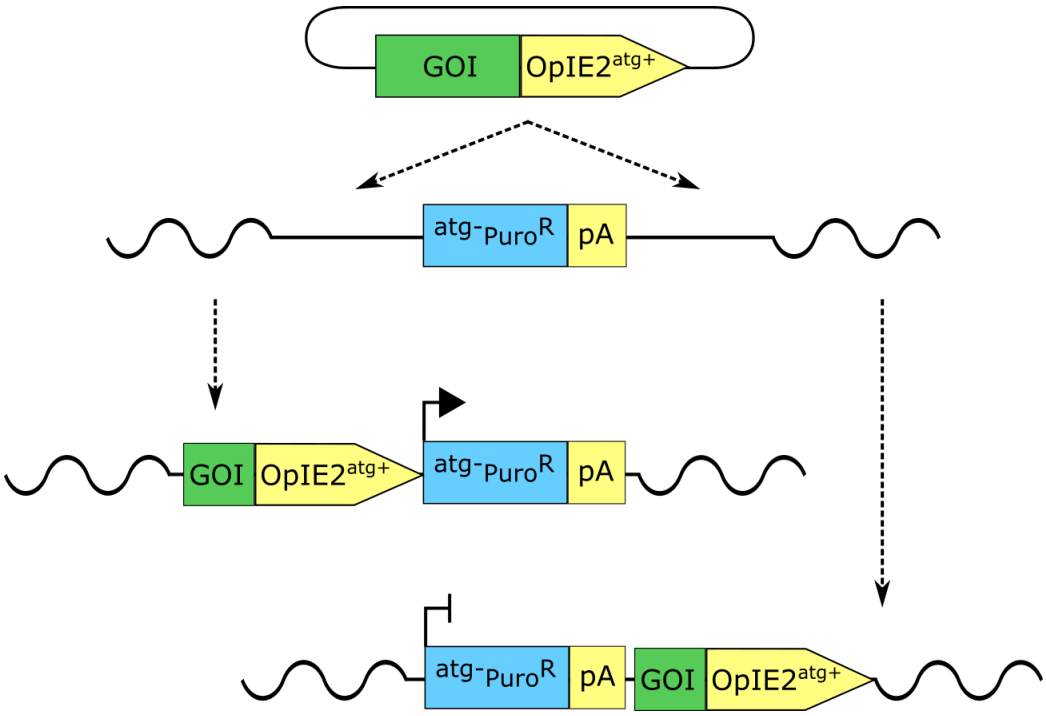

Figure 8 – Example Schematic of Promoter Trap

A promoter trap allows selection for integration events at a specific pre-determined locus. In the donor plasmid, the GOI is flanked by a promoter and subsequent methionine codon. A methionine deficient selection marker within the genome (Puromycin in this case), is activated only when the integrating cassette lands immediately upstream of it. Puromycin resistance is not activated when the GOI lands elsewhere. GOI, Gene of Interest; pA, poly A sequence.

To track the success of the Recombinase Mediated Cassette Exchange reaction, a reporter gene system can be used. The ϕC31 recombinase system was initially validated in Sf9 cells

using a lacZ reporter, but this required transferring the donor plasmid back in to *E. coli* for verification (Chompoosri et al., 2009). Fluorescent reporter proteins, such as those used in Fernandes et al. (2012) or Long et al. (2013), provide a convenient *in vivo* approach to tracking cassette exchange via flow cytometry (Fernandes et al., 2012; Long et al., 2013).

While these fluorescent reporters would ideally never be co-expressed, a mixed culture of successful and unsuccessful recombinant Sf9 cells is a likely outcome post-treatment. Therefore, two fluorescent reporter genes with clearly differentiated emission peaks should be chosen. Additionally, the fluorescence intensity of each reporter should be balanced such that in a sample expressing both genes concurrently, both are easily detectable. A pool of monomeric fluorescent proteins already undergoing characterization in the Aucoin Lab was used as initial candidates **(Table 4)**. mAzami-Green and mKOk were selected as candidates to be validated through a co-transfection trial *in vivo* (Chapter 5).

Table 4 - Flourescent Reporter Protein Candidates

Flourescent reporter proteins considered for use in RMCE tracking. All proteins listed were being characterized in Aucoin Lab during the design phase of this thesis.

Protein	Emmision Peak (nm)	AddGene Reference
mAzami-Green	505	54798
eGFP	507	13031
mKOk	563	53617
mKate2	633	37132
LSSmKate2	605	31867
mIFP	704	54620

3.4 Promoters

Driving strong gene expression in the Sf9 cell line commonly relies on late-activated promoters from the *AcMNPV*, such as polh or p10. However, these viral promoters, central to the Baculovirus Expression Vector System, are only recognizable by their cognate viral polymerase. In the context of constitutive Sf9 expression, the early *AcMNPV* promoter IE1 is the BEVS-derived promoter of choice. This promoter has been in use extensively since it was first demonstrated to drive constitutive expression of tissue plasminogen activator in Sf9 cells (Jarvis et al., 1990). The Immediate Early 1 promoter has since been central to various insect expression systems - most notably the InsectDirect system, a series of expression vectors that were shown to produce stable yields of protein kinase, heat shock protein, and phospholipase in Sf9 cells (Loomis et al., 2005). Activity of the IE1 promoter can be improved with the upstream addition of an hr5 enhancer region (Ren et al., 2011).

AcMNPV is not the only *Baculoviridae* member with promoters that function in lepidoptera. The *Orgyia pseudotsugata Multicapsid Nucelopolyhedrosis Virus* (*OpMNPV)* contains two early stage promoters, listed as OpIE1 and OpIE2, that natively perform similar roles to their *AcMNPV* counterparts (Theilmann & Stewart, 1992). These promoters were first utilized in a recombinant setting during testing of the antibiotic Zeocin. During these early tests, the OpIE2 promoter appeared to show 5-10 fold increased activity relative to OpIE1 (Pfeifer, Hegedus, Grigliatti, & Theilmann, 1997).

The first direct comparison of these promoters showed that the enhanced IE1 promoter hr5IE1 provided a 20% advantage over OpIE2 during transient expression in Sf9 cells (Ren et

al., 2011). However, a more recent study determined that over a longer timeframe (96h vs 24h), OpIE2 showed a 3.3 fold increase in eGFP activity when compared to hr5IE1. This difference was even more pronounced when compared with other endogenous *Spodoptera* promoters (Bleckmann et al., 2015).

Based on these reports, any of the *Baculoviridae*-derived promoters could be reasonably considered for use in a recombinant system based in the Sf9 cell line. For driving a key reporter gene the OpIE2 promoter appears to show the highest sustained output, and as such is the first choice for this design.

3.5 Gene Architecture

After considering a variety of potential design choices, a gene architecture fulfilling all of the proposed guidelines was developed **(Figure 9)**. This arrangement is composed of two major components. The first is a selection region for developing the target cell-line. This is composed of a hygromycin resistance gene flanked by FRT regions. The hygromycin resistance allows for selection for successful integrants while developing the cell-line. This gene is driven by an OpIE2 promoter, and aborted by the associated OpIE2 transcriptional terminator. The FRT regions provide a prospective laboratory with the option to remove this selection marker from the cell-line should it be so desired. FLP Recombinase's natural predilection for excision makes it an optimal choice for this purpose.

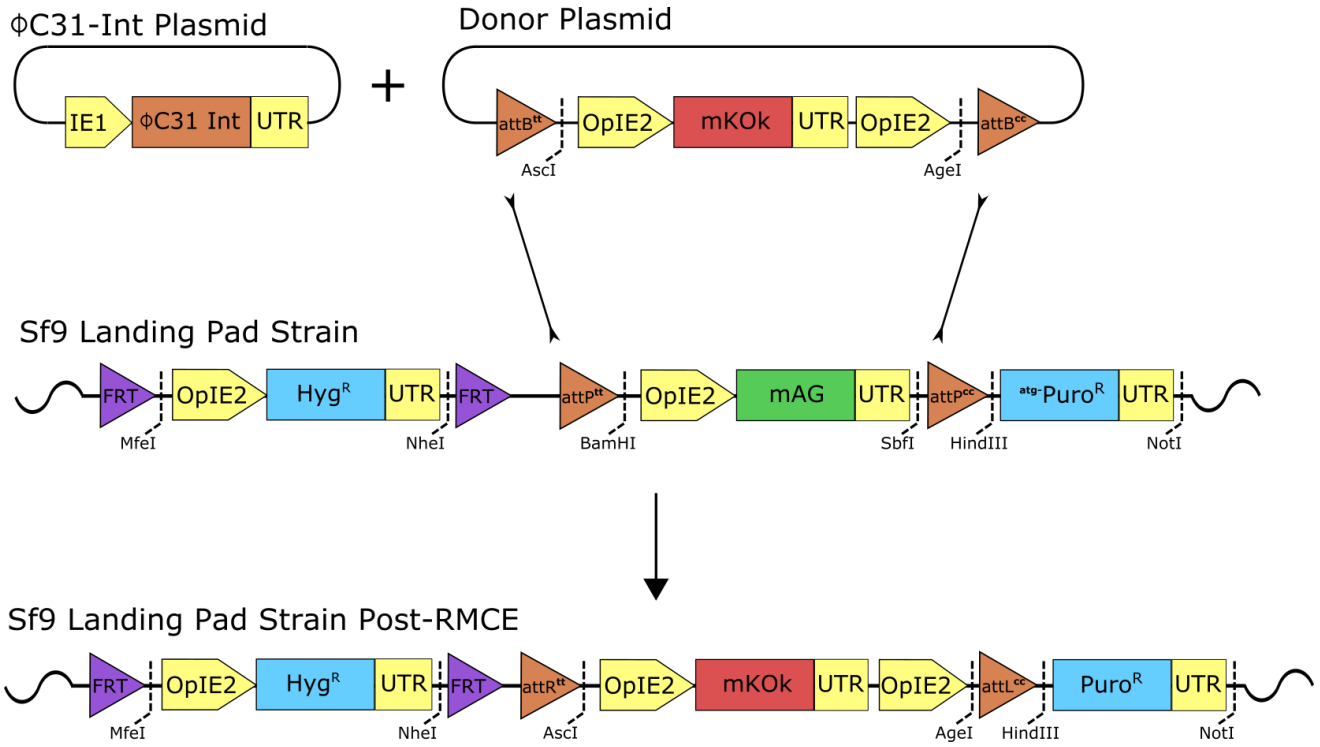

Figure 9 - Gene Architecture

Full diagram of the proposed recombinogenic gene architecture. Within the genomic landing pad region, there are three sub-cassettes with distinct functions. A Hygromycin resistance gene (Hyg^R) is placed between two FRT loci. This selection marker allows the genomic cassette to be successfully integrated in the Sf9 cell, while the FRT loci provide the option for its convenient removal if necessary. The fluorescent reporter mAG is placed between two orthogonal ϕC31 attP sites which form an RMCE landing pad. An ATG-deficient puromycin resistance gene is used as a promoter trap, and is placed immediately following the 3' attP site. When recombination occurs between the donor plasmid (expressing mKOk) and the genomic cassette, the result will be an Sf9 cell expression mKOk as well as expressing a now-active Puromycin resistance gene. This recombination event is mediated by a ϕC31 Integrase that is introduced via an expression vector that is co-transfected alongside the donor plasmid. Orthogonal att-sites are delineated by superscripts indicating their core-dinucleotides, tt and cc.

The second region is the area providing ϕC31-Recombinase Mediated Cassette Exchange functionality. A pair of ϕC31 attP sites are placed around a reporter gene. Placing attP sites in the genome is counterintuitive given that the natural lifecycle of Phage C31 sees the virus integrate into genomic attB sites. However, it has been shown in *B. mori* that RMCE between donor-attB and genomic-attP sites have a near 2-fold increase in success when compared to the reverse arrangement (Long et al., 2013). The particular orthogonal attP sites are denoted attP-cc and attP-tt, based on their core dinucleotide. These particular core sequences were chosen to reduce any chance for misrecognition between sites.

These ϕC31 attP sites flank a green fluorescent reporter gene, mAzamiGreen. Like the hygromycin resistance gene, the mAG CDS is driven by an OpIE2 promoter, and terminated by the OpIE2 3' untranslated region (UTR).

Immediately subsequent to the 3' attP site, there sits a methionine deficient puromycin gene, followed by an OpIE2 transcriptional terminator. This promoter trap will allow for puromycin selection upon completion of RMCE by including a promoter and in-frame ATG at the end of the donor cassette.

While Gibson Assembly has become the standard approach for bench-scale *in vitro* gene assembly, many labs still utilize restriction endonuclease techniques. In anticipating that this construct could potentially be shared with other research groups, restriction sites have been included flanking each functional gene such that different selection markers or reporters could be swapped in to facilitate another laboratory's objectives. These sites were chosen to all be 6-cutters or 8-cutters, and all are compatible with FastDigest^{tm} restriction endonucleases. The

HindIII that is 5' adjacent to the ATG^- puromycin resistance gene was placed specifically to keep the $Puro^r$ gene in frame after RMCE.

The donor cassette consists of a red fluorescent reporter gene, mKOK, driven by an OpIE2 promoter and terminated by the associated OpIE2 3' UTR. Downstream of this 3' UTR is another OpIE2 promoter with followed by a single methionine codon. Following RMCE, this region will activate the ATG^- puromycin resistance present in the genomic cassette. These sequences are flanked by ϕC31 attB-cc and attB-tt sites, corresponding to the orthogonal attP sites on the genomic cassette.

3.6 Implementation

Developing the proposed strain of Sf9 first requires integrating the target cassette into the Sf9 genome. This can be facilitated by the use of piggyBAC transposon technology. The piggyBac transposon is a cut-and-paste transposon that is mediated by the enzyme piggyBac Transposase. The genetic component of piggyBac consists of two 13bp terminal inverted repeats and two 19bp subterminal inverted repeats, separated by 3bp on the 5' side and 31bp on the 3' side. These reside within less-stringently conserved regions of 311bp and 235bp respectively **(Figure 10)**. It inserts randomly into any TTAA tetranucleotide sequence (Elick, Bauser, & Fraser, 1996).

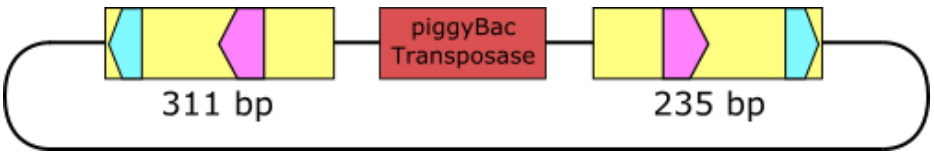

Figure 10 – piggyBac Transposon

In its native form within *T. ni*, the piggyBac transposase gene (red) is flanked by two perfect terminal inverted repeats (blue, magenta), within larger semi-conserved regions (yellow). Shown here are the minimal forms of these TRDs. Figure adapted from (X. Li et al., 2005). TRD, Terminal Repeat Domain.

PiggyBac technology is derived from the lepidopteran species *T. ni*, and has a long history of use in *Spodoptera* cell lines. After expression levels of genes introduced by piggyBac were seen to be subject to position effects, a set of altered piggyBac terminal repeat domains were developed to eliminate these effects (Shi et al., 2007). This iteration of piggyBac technology has been successfully implemented in Sf9 cells to produce a variety of glycogene-expressing cell-line variants (Aumiller, Mabashi-Asazuma, Hillar, Shi, & Jarvis, 2012). Developing the target strain of ϕC31-competent Sf9 cells is an ideal application of this piggyBac transposase. The target cassette will be placed in a piggyBac vector, and cotransfected with a piggyBac transposase into a culture of Sf9 cells. Successful integrants can be selected for via hygromycin.

To favour integration events in transcriptionally active regions of the genome, a sterile run through a Flourescence-Activated Cell Sorter (FACS) could be used to filter out cells exhibiting low-levels of mAG expression. The resultant subculture of high-expressing integrants can be converted to a series of clonal isolates through a limiting dilution protocol.

Limiting dilution involves diluting cells and aliquoting them into tissue culture plates such that each plate-well contains a single cell. Adherent cultures grown up from each of these isolated cells can be considered genetically identical (Harrison & Jarvis, 2010).

The final version of the proposed ϕC31-competent Sf9 cell line should exhibit a single locus of recombination. Thus, candidate strains produced through limiting dilution can be screened via qPCR to eliminate any isolates in which the piggyBac transposase caused multiple integration events. For single-integrant isolates, the location of integration within the Sf9 genome can be determined through sequencing an inverse PCR.

The remaining clonal isolate with the highest degree of mAG expression can then be tested as a ϕC31-competent Sf9 cell line. A co-transfection experiment involving the donor cassette and the ϕC31 Integrase on suicide vectors (plasmid vectors incapable of replication within the host cell) will confirm the functionality of this new cell-line. A successful trial will see post-RMCE passages of the cell-line exhibit consistent expression of mKOK, as well as resistance to puromycin.

Chapter 4

General Materials & Methods

4.1 Plasmids Used and Constructed

Table 5 - Plasmids Used & Constructed

Name	Source	Description/Contents
pJC2	Charles Lab, Waterloo, ON	ϕC31 Integrase CDS
pBSII-IE1-orf	Jarvis Lab, Laramie, WY	piggyBac CDS, IE1 promoter
pIE1Hygro	Jarvis Lab, Laramie, WY	Hygromycin Resistance CDS
pXLBacII-TetOnGG1-DsRed1	Jarvis Lab, Laramie, WY	piggyBac TRD forward
pXLBacII-TetOnGG57-DsRed2	Jarvis Lab, Laramie, WY	piggyBac TRD reverse
pMTL85141	Aucoin Lab, Waterloo, ON	Cloning vector
pOpIE2-Csy4	Aucoin Lab, Waterloo, ON	OpIE2 promoter/UTR
pBsynbackbone001	Synthesized, Thermo	Recombinatorial backbone
pMTL-synbackbone	Constructed	Rec backbone in cloning vector
pMTL-synbackbone-Hygro	Constructed	Rec backbone + Hyg^r cassette in cloning vector
pMTL-synbackbone-Hygro-mAG	Constructed	Rec backbone + Hyg^r + mAG cassettes in cloning vector

4.2 Bacterial Maintenance Protocols

4.2.1 Overnight Culture and Strain Storage

E. coli containing plasmids of interest were grown overnight in 25ml of LB broth at 37°C, shaken at 300rpm. The resultant overnight culture was then stored in 1ml aliquots, consisting of 85% culture and 15% glycerol, at -80°C. All *E. coli* used for cloning was strain DH5α.

4.2.2 Chemically Competent Cells

The following was the standard procedure for making chemically competent *E. coli* DH5α cells. Volumes can be adjusted proportionally if needed.

A 300ml flask of LB broth was inoculated with 3ml of overnight *E. coli* DH5α culture, at 37°C, 235rpm, and incubated until cell density reached OD_{600}=0.35. Once an optimal cell density was achieved the culture flask was placed on ice for 30 minutes, with occasional agitation. The culture was aliquoted evenly into 6 pre-chilled Falcon^{tm} 50ml Conical Centrifuge Tubes (Fisher Scientific Ltd., Ottawa, ON), and pelleted at 3000g for 15 minutes in a 4°C centrifuge. Supernatant was then decanted, and cell pellets were gently resuspended in in 30ml of ice-cold 100mM $MgCl_2$.

Cultures were consolidated down from 6 tubes to 4 and pelleted at 2000g for 15 minutes at 2°C. The supernatant was decanted, and cell pellets were gently resuspended in 50ml of ice-cold 100mM $CaCl_2$. Cultures were incubated on ice for 45-60 minutes. After incubation, cultures were once again pelleted at 2000g for 15 minutes in a 2°C centrifuge. The supernatant was decanted and cell pellets were gently resuspended in 20ml of ice-cold 85mM $CaCl_2$, 15% glycerol.

Cultures were consolidated down from 4 tubes to 2, and pelleted at 2000g for 15 minutes in a 2°C centrifuge. Supernatant was decanted, and cell pellets were gently resuspended in 1ml of 85mM $CaCl_2$, 15% glycerol. Cultures were pipetted in 50µl aliquots into pre-chilled 1.5ml microfuge tubes, and stored at -80°C.

4.2.3 Bacterial Transformations

Chemically competent *E. coli* cell preps were placed on ice. 5µl of plasmid-buffer solution was added to each competent cell prep, gently mixed, and incubated on ice for 40 minutes. Cells were heat-shocked in a 42°C water bath for 45 seconds, and then immediately incubated on ice for 2 minutes. 950µl of SOC media was then added into each transformation prep. Transformations were incubated for 1 hour at 37°C, 275rpm. 50µl, 100µl, and 500µl volumes were plated on LB agar (imbued with relevant antibiotic as necessary). Plates were incubated overnight at 37°C, and then inspected for colony growth.

4.2.4 Miniprepping Transformed Plasmids

Minipreps were performed using a GeneJET^{tm} MiniPrep kit (Thermo Fisher Scientific Inc., Mississauga ON). All actions taken during miniprepping were done in accordance with the manufacturer's protocols.

4.3 DNA Manipulation Protocols

4.3.1 PCR & Primer Design

Primers for PCR were designed with Benchling's Primer Wizard (Benchling Inc., www.benchling.com) algorithm when possible. If the Primer Wizard algorithm was unsuccessful in selecting primers, primers would be designed manually within the Benchling sequence editor by analyzing the properties of surrounding DNA relative to the GOI. Primers

were always designed such that their Tm values were within 5°C of one another. Whenever possible, primers were designed with a 3' terminal G or C clamp.

Primers were ordered from Integrated DNA Technologies Inc. Primers were received as individual lyophilized samples. Samples were resuspended in dH_2O at a concentration of 200μM, and stored at -40°C.

PCR reaction were carried out at 50μl volumes, using Phusion Polymerase and Phusion HF Buffer (Thermo Fisher Scientific Inc., Mississauga ON) **(Table D 1)**. Annealing temperatures were calculated using the NEB Tm Calculator v.1.9.10 (New England Biolabs Inc., tmcalculator.neb.com). Elongation times were based on a 45sec/kb heuristic, with a minimum time of 1 minute. Template plasmids were diluted 1/20 when necessary. PCR reactions were run for a standard 35 cycles. Products were analyzed on and extracted from an agarose gel.

4.3.2 Splicing Overlap Extension – Polymerase Chain Reaction

Splicing Overlap Extension – Polymerase Chain Reaction (SOE-PCR) protocols were performed in two 25μl mixes, for a total of 50μl **(Table D 2)**. Annealing temperatures for both the overlap region and the outer primers were calculated using the NEB Tm Calculator v.1.9.10. Template fragments were included in roughly equimolar amounts; volumes were adjusted based on concentration.

Mix A would be run for 10 cycles, and then halted at 10°C. Mix B would be added into Mix A, gently mixed via pipette, and the combination would be run for a full 35 cycle PCR schedule. Products would be analyzed on and extracted from an agarose gel.

4.3.3 Gibson Assembly

Fragments of interest were prepared by via PCR with Gibson primers. Gibson primers were designed using Benchling's Gibson Assembly Wizard algorithm. Gibson reactions were performed in 10µl volumes using NEB Gibson Assembly Master Mix (New England Biolabs Ltd., Whitby ON). A total of 0.1pmol of DNA was used, with a 2:1 ratio of insert fragment to backbone fragment **(Table D 3)**. Volumes were adjusted based on size and concentration of DNA fragments. Components were combined and incubated for 1 hour at 50°C. Product was transformed according to section 4.1.3.

4.3.4 Restriction Cloning

Restriction reactions utilized Thermo FastDigest enzymes, and the associated FastDigest Green buffer (Thermo Fisher Scientific Inc., Mississauga ON). ~1µg of DNA was used for each reaction **(Table D 4)**. Components were combined, gently mixed via pipette, and allowed to incubate at 37°C for 5 minutes. Resultant fragments were analyzed and extracted on an agarose gel.

Ligation reactions were performed at 20µl volumes, using T4 DNA Ligase in Thermo 10x T4 Buffer (Thermo Fisher Scientific Inc., Mississauga ON). Insert and backbone fragments were included at a 3:1 molar ratio **(Table D 5)**. Components were combined, mixed via pipette, and allowed to incubate overnight at 4°C. Products were transformed according to section 4.1.3.

4.3.5 Agarose Gel Electrophoresis

Agarose gels were made at volumes of 50ml. For most applications, an agarose content of 0.8% w/v was used. However, sometimes small fragments would be run on a 1.4% w/v gel. iNtRON RedSafe Dye (iNtRON Biotechnology Inc., Burlington MA) was utilized for DNA visualization **(**

Table D 6**)**. Electrophoresis was run for 1 hour, at 90V/300W/3A. Samples were prepared in 50µl volumes, with the addition of 8µl Thermo 6x Loading Dye (Thermo Fisher Scientific Inc., Mississauga ON), and run against a GeneRuler^{tm} 1Kb DNA Ladder or a GeneRuler^{tm} 100bp DNA Ladder (Thermo Fisher Scientific Inc., Mississauga ON) **(Figure 11)**. When extraction was necessary, samples would be run in parallel lanes, with a diagnostic lane containing 8µl, and an extraction lane with 50µl. Gels were lit with UV-B radiation using a Life Technologies TFX-35M Illuminator.

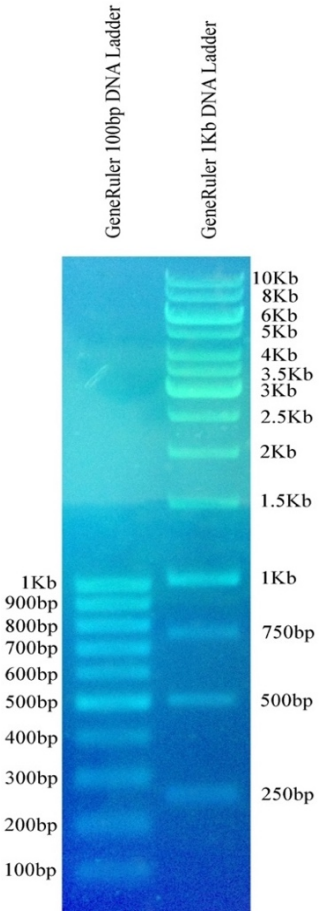

Figure 11 - DNA Ladders

GeneRuler^{tm} 100bp DNA Ladder and GeneRuler^{tm} 1Kb DNA Ladder were used for DNA fragment size comparison during electrophoresis.

DNA bands of interest were cut out of the gel and chemically extracted. Gel extractions were performed using a Thermo Scientific GeneJET Gel Extraction Kit (Thermo Fisher Scientific Inc., Mississauga ON). All actions taken during miniprepping were done in accordance with the manufacturer's protocols.

4.3.6 DNA Sequencing

DNA sequencing was performed at The Centre for Applied Genomics (SickKids, Toronto ON). Sequencing primers were designed in Benchling (Benchling Inc., www.benchling.com). Samples were prepared in 7µl aliquots at a concentration of 0.05pmol-0.15pmol. 0.7µl of a sequencing primer, at a concentration of ~5pmol was added, for a total sample volume of 7.7µl. Samples were labeled and shipped in PCR tubes, further sealed in a Ziploc bag.

4.4 Maintanance of Sf9 Cells

Sf9 cells are grown in SF900-III media (Thermo Fisher Scientific Inc., Mississauga ON). They are maintained in shake flasks at 27°C and agitated at 130rpm. Cells were passaged when cell density reached 3-5x10^6 cells/ml, and were seeded down to 0.5x10^6 cells/ml.

4.5 Flow Cytometry

To prepare samples for flow cytometry analysis, cells were mechanically detached from plates via gentle repeated aspiration. Samples were then spun down for 5 minutes at 150xG and the supernatant removed. The cell pellet was resuspended in 500µl of a 2% formaldehyde solution and set aside for 1 hour. Samples were subsequently transferred into 5ml round bottom test tubes and diluted with 2ml of PBS buffer.

Samples were run on a FACScalibur flow cytometer, with 10,000 events collected per sample. Data was collected via the FL1-H (530nm) and FL2-H (585nm) emission filters.

Chapter 5

Validating mAG and mKOk as Reporter Genes

Based on a previous examination of emission frequency and intensity of several fluorescent proteins, it was believed that mAzamiGreen and mKOk would function well as a GFP/RFP pairing for testing the ϕC31-competent Sf9 cell line. To confirm that these reporters would function concurrently in the Sf9 line, a series of transfections were performed and analyzed via flow cytometry. Two plasmids, expressing mAG and mKOk respectively, were transfected individually, as well as co-transfected, into preparations of Sf9 cells. Data for each of these treatments, along with a 4th treatment consisting of a mixture of individually transfected cells, were collected using a FACSCalibur Flow Cytometer (BD Biosciences Inc., Mississauga ON). The FACSCalibur possesses a single 15 milliwatt argon-ion excitation laser with a wavelength of 488nm. It has 3 bandpass fluorescence detectors, with filter peaks at 530nm, 585nm, and 670nm (Biosciences, 2010).

5.1 Materials and Methods

5.1.1 Transfection of Sf9 Cells

$5x10^5$ Sf9 cells were taken from suspension culture and plated into each well of a 6-well tissue culture plate. Cells were incubated for 2 hours to allow for attachment.

Initial samples of mAG and mKOk plasmids were obtained from Aucoin Lab (University of Waterloo). Samples were thawed on ice from a -40°C storage freezer. They were measured via nanodrop to be at 411 ng/µl and 496 ng/µl respectively. Preparations of 1µg and 2µg of each plasmid were diluted in dH_2O to a total volume of 100µl each. Preparations of 8µl of Escort IV Transfection Reagent (Sigma-Aldrich Canada Co., Oakville ON), a liposome-forming mixture, was diluted in 92µl of SF900III media, for a total volume of 100µl. DNA solutions were mixed with transfection reagent solutions and allowed to incubate at room temperature for 60 minutes.

After incubation of cells was complete, liquid media was aspirated from each well and replaced with 800µl of fresh SF900III media. 200µl of transfection mixture was added to each well, for a total volume of 1000µl. Plates were then incubated for 48 hours at 27°C. Samples were analyzed by flow cytometry according to section 4.4.

5.2 Results

Three transfection permutations were explored:

1. mAzamiGreen (a green-fluorescent monomer)
2. mKOk (a red-flourescent monomer)
3. Co-transfection with both mAzamiGreen and mKOk

Each of these was performed via transfection with 1μg of plasmid **(**Figure 12**)**, and 2μg of plasmid (Figure 13). Additionally, a post-transfection mixture of 2μg transfections with mAG and mKOk was included in cytometric analysis **(Figure 13)**.

Data presented has been pre-gated to eliminate untransfected cells. Visual gates delineating treatment populations are labelled, mAG, mKOk, and mAG+mKOk. These labels include the percentage of data points found within the defined region. All plots are presented on log-log axes. Total Cell Count (TCC) for each plot is specified in the captions.

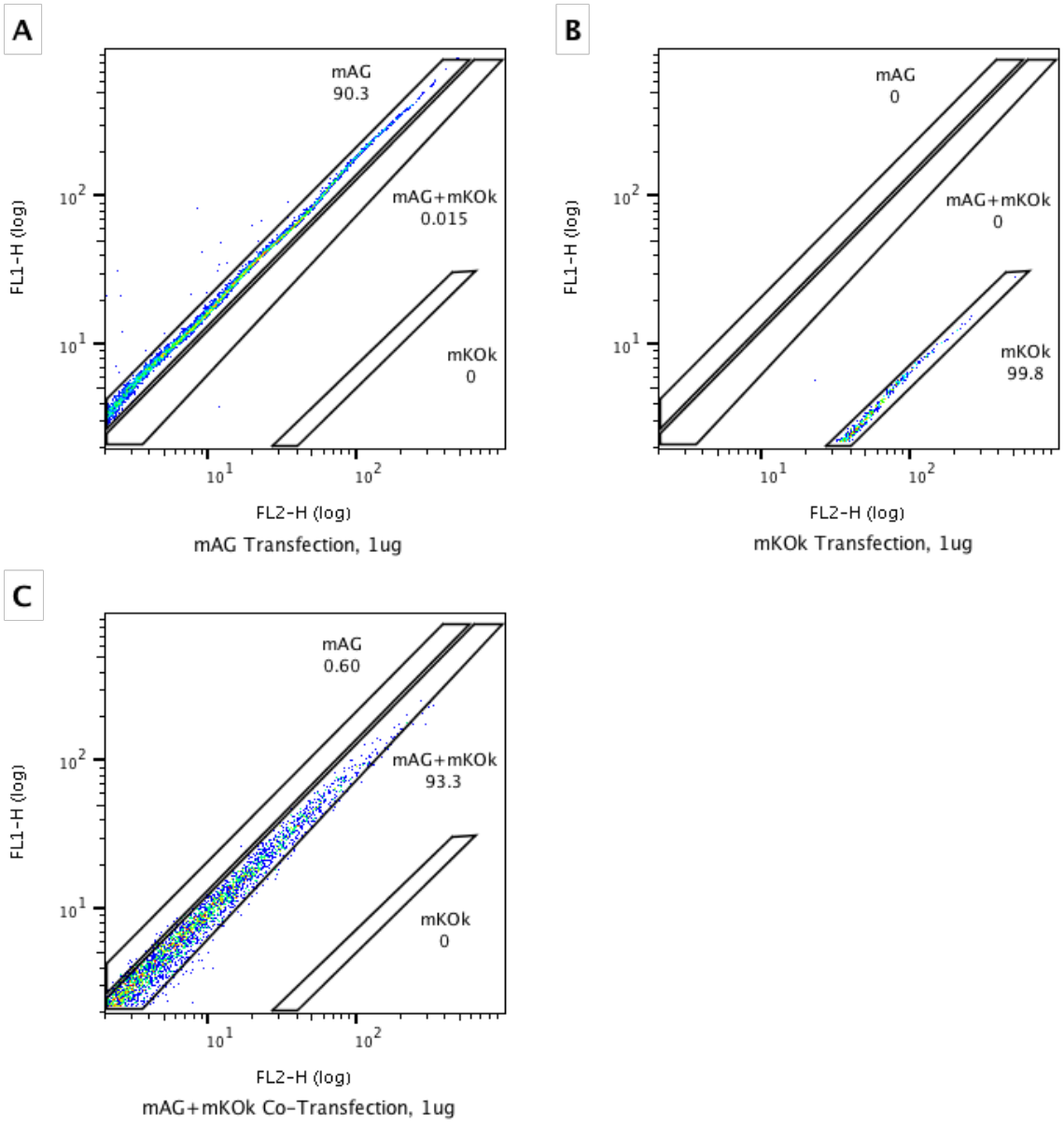

Figure 12 - Fluorescence Response of Sf9 Cells Transfected with 1µg of Plasmid

A) Individual mAG transfection. TCC = 6467. B) Individual mKOk transfection. TCC = 444.

C) Co-Transfection with 0.5µg each of a mAG and mKOk plasmid. TCC = 3313.

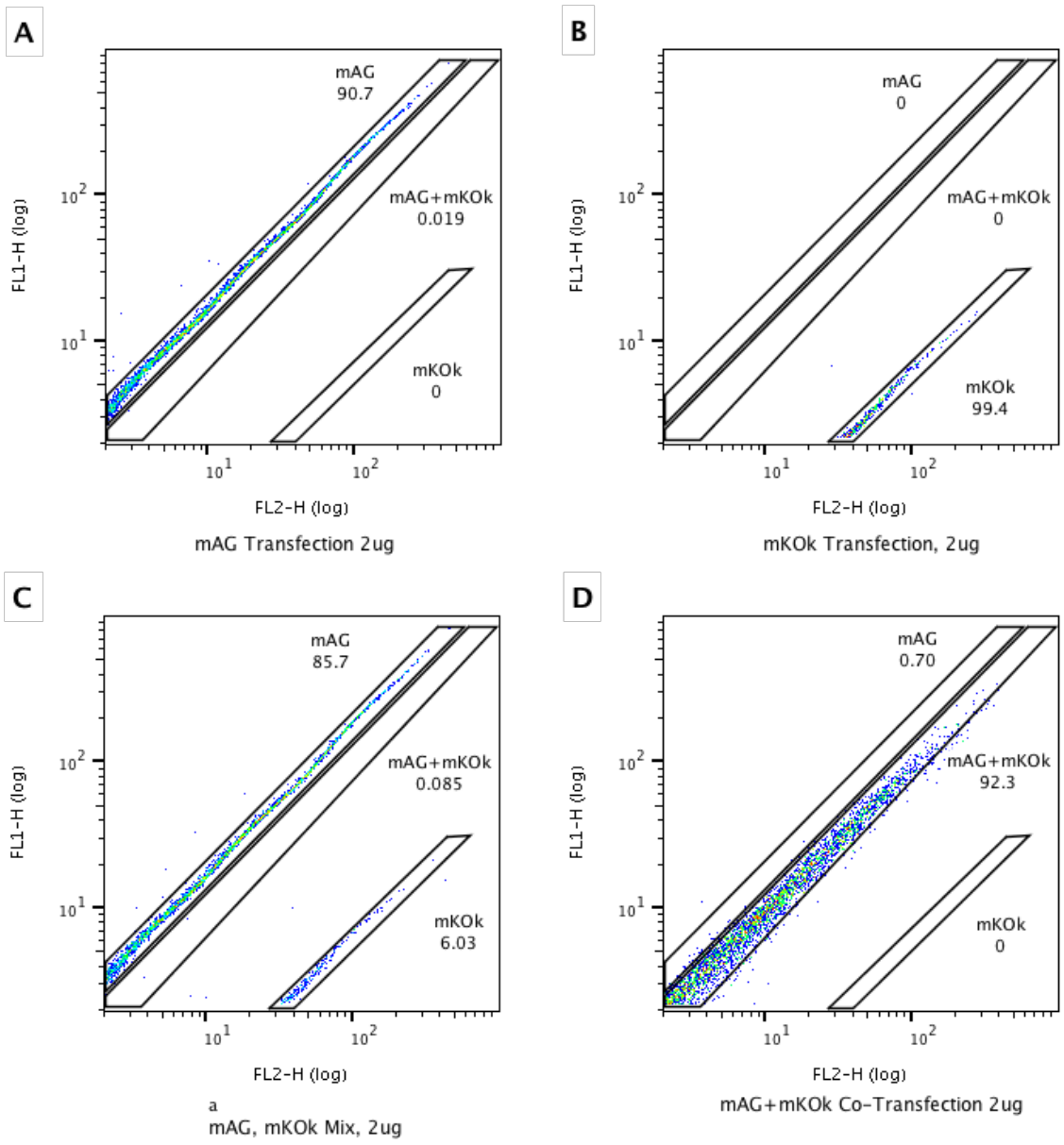

Figure 13 – Fluorescence Response of Sf9 Cells Transfected with 2µg of Plasmid

A) Individual mAG transfection. TCC = 5316. B) Individual mKOk transfection. TCC = 347. C) Mixture of cells individually transfected with mAG and mKOk. TCC = 3514. D) Co-transfection with 1µg each of mAG and mKOk. TCC = 3848

5.3 Discussion

The efficacy of a fluorescent reporter pair hinges on whether a transfected population of cells can be easily differentiated via flow cytometry or FACS. If non-overlapping flow cytometry gates can be produced whereby little-to-no data points from one treatment are present in a competing gate, the reporter pair can be deemed useful. For mAG and mKOk, this appears to be the case.

In both instances of a singular mKOk transfection **(Figure 12B, Figure 13B)**, 0% of points were present in the competing gates. Singular transfection of mAG **(Figure 12A, Figure 13A)** found no bleed into the mKOk region, and trivial bleed into the mAG+mKOk region. This is an expected finding, given the reported difference in these two protein's fluorescent emission peaks. It is worth noting, however, that significantly fewer total events showed mKOk expression relative to mAG expression. A longer incubation period may help mitigate this, by allowing greater build-up of mKOk proteins within the cell.

In a future clonally-isolated RMCE-competent cell-line, expression of one of these proteins will be exchanged for the other. While the promoter trap function of the proposed design will ideally eliminate any unsuccessful RMCE-events, it may be necessary to sort these out through FACS. **Figure 13C**, a post-transfection mixture of the singular mAG and mKOk populations, mimics the result of a partially-successful RMCE reaction and shows that successful and unsuccessful reactions will be easily differentiable.

In the proposed RMCE-competent Sf9 cell line, co-expression of these two proteins is an undesired outcome, and should also be sorted out of a post-RMCE population. While mKOk and mAG have significantly different peaks, mAG displays over twice the emission intensity

of mKOk. Prior to this experiment, it was unclear if fluorescence by mKOk would be strong enough to push the co-transfected population far enough away from the singular mAG population to be clearly distinguishable. In both **Figure 12C** and **Figure 13D**, less than 1% of events from the co-transfection were present in the mAG region. Furthermore, an overlay of all transfection data qualitatively shows that all three expression profiles are clearly distinct **(Figure 14)**. Based on these data, the use of mAzamiGreen and mKOk as a fluorescent reporter pair for use in testing an RMCE-competent cell line has been validated.

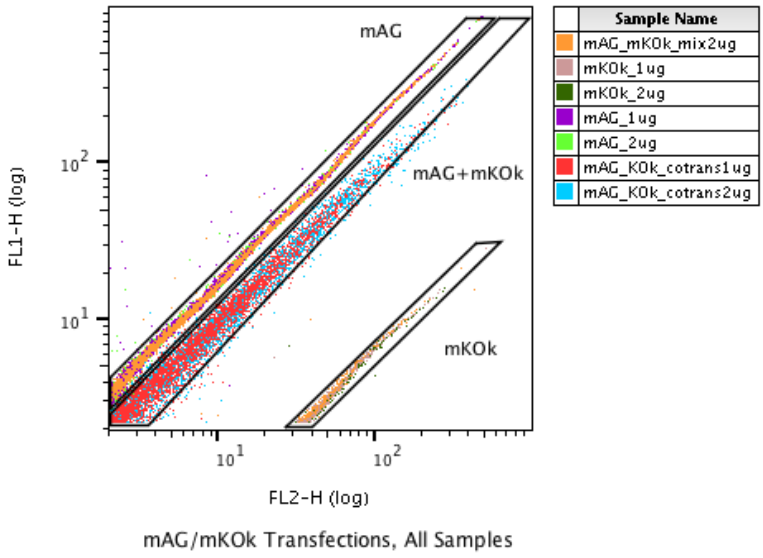

Figure 14 – Composite Fluorescence Response of All mAG, mKOk Transfections

An overlay of all cytometry data collected. Includes both 1µg and 2µg transfections. TCC = 23,249.

Chapter 6

Construction of Target Cassette

The proposed genetic system involves several distinct gene cassettes that need to work in concert. To construct the full system, sub-cassettes coding for each ORF would be constructed, and then assembled together in context with the recombinase attachment sites and piggyBAC regions **(Figure 15)**.

Firstly, a "recombinatorial backbone" was synthesized as a gblock fragment through IDT Inc. This fragment contained a pair of FRT sites, as well as orthogonal attB and attP pairs, with restriction endonuclease sites interjected in between them **(Figure 16)**. The region containing the FRT sites and attP sites were PCR'd out to be used in constructing the genomic cassette, while the region containing the attB sites was separated for the donor cassette.

Three sub-cassettes were constructed using PCR to isolate their component parts, and an SOE-PCR protocol was utilized to assemble them. These cassettes are the hygromycin resistance cassette, the mAzamiGreen cassette, and the puromycin promoter trap cassette. Once these sub-cassettes were constructed, both Gibson assembly and restriction cloning protocols were used to merge them with the recombinatorial backbone.

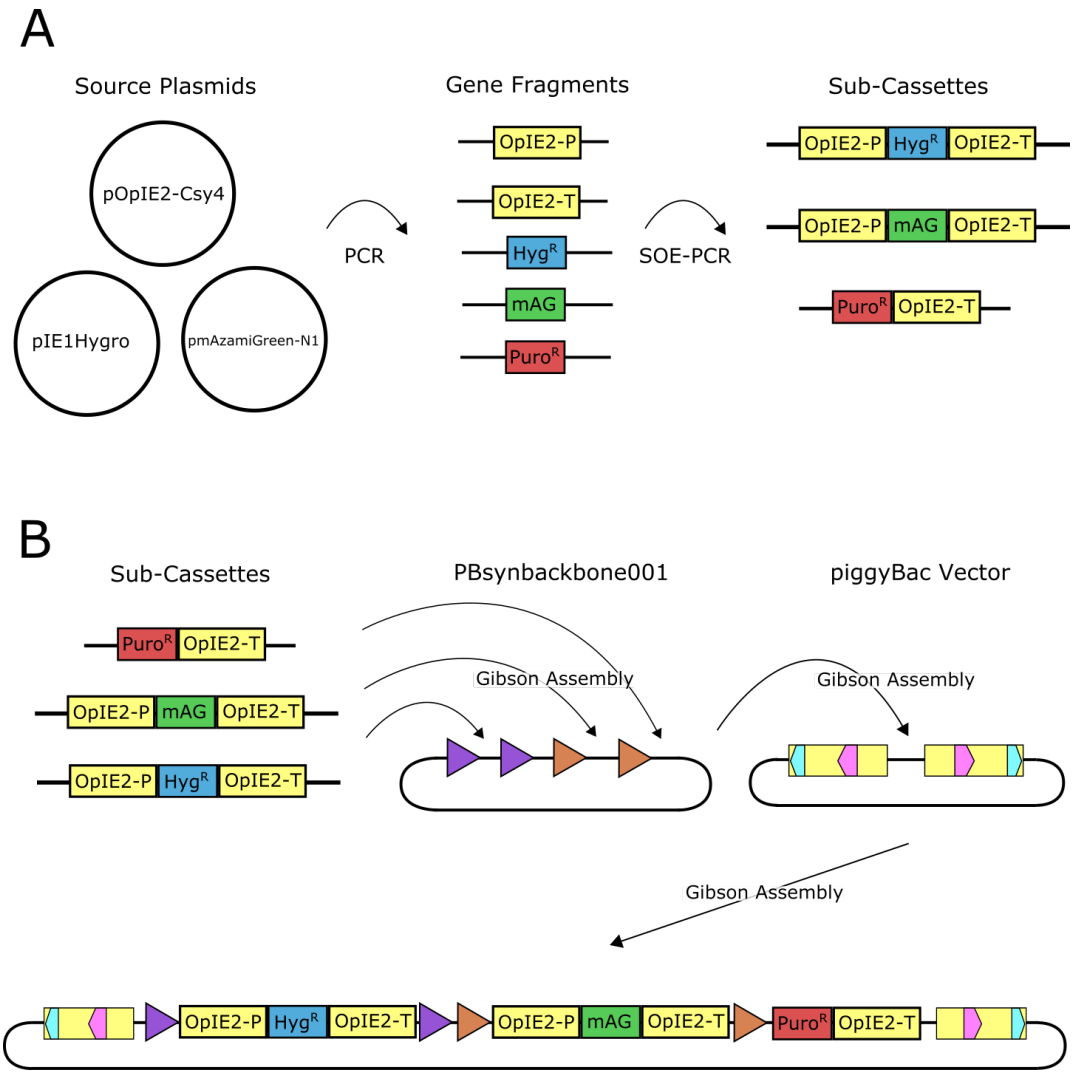

Figure 15 – Gene Construction Overview

A) Individual gene fragments (promoters, CDSs, 3' UTR sequences) are sourced and isolated via PCR. These gene fragments are then stitched together using an SOE-PCR protocol to produce sub-cassettes. B) A "recombinatorial backbone", PBsynbackbone001, containing recombinase attachment sequences is synthesized, and the sub-cassettes are integrated via Gibson assembly. The full cassette is placed between a set of piggyBac terminal repeat domain regions, for later integration into the Sf9 genome. CDS, Coding Sequences; SOE-PCR, Splicing by Overlap Extension Polymerase Chain Reaction; OpIE2-P, OpIE2 promoter; OpIE2-T, OpIE2 transcriptional terminator; Hyg^R, Hygromycin resistance coding sequence; mAG, mAzamiGreen coding sequence; $Puro^R$, Puromycin resistance coding sequence.

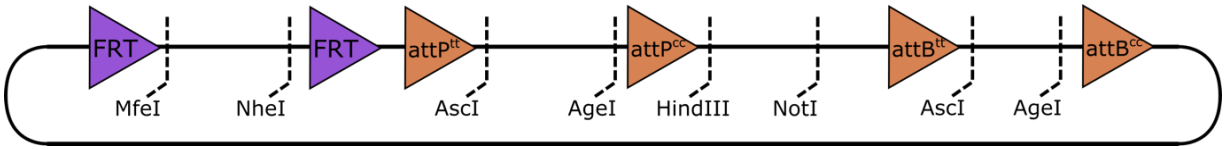

Figure 16 – Diagram of Recombinatorial Backbone

The recombinatorial backbone is a synthesized DNA fragment containing all the necessary recombinase attachment sites, along with intermittent restriction endonuclease recognition sites to provide the option restriction-cloning. This fragment is separated into the genomic component (containing the FRT and attP sites) and a donor component (containing the attB sites).

PiggyBAC Terminal Repeat Domain (TRD) sequences and sequence for the piggyBac Transposase CDS were obtained courtesy of Dr. Donald Jarvis (University of Wyoming). The piggyBac Transposase CDS was subsequently codon-optimized for expression in Sf9 cells, and synthesized under control of the AcMNPV IE1 promoter. Synthesis of this CDS was performed via Thermo Fisher GeneArt (Thermo Fisher Scientific Inc., Mississauga ON).

6.1 Methodology and Results

6.1.1 Constructing Sub-Cassettes

Promoter sequences, CDS's, and transcriptional terminators required for the genomic cassette were isolated from plasmids via PCR as described in section 3.8.1 **(Table 6)**. Primers used for fragment isolation included overlapping overhang regions to facilitate subsequent SOE-PCR reactions. The promoter and terminator fragments were amplified multiple times using different primer overhangs to allow for use in multiple different SOE-PCRs. DNA

fragments are commonly labeled using the primers used in their creation (ex. PBpr007-008 is the fragment produced using primers PBpr007 and PBpr008). Fragments referenced in this chapter can be found in **Table 7**.

Table 6 - Source Plasmids

Component	Plasmid	Source
OpIE2 Promoter	pOpIE2-Csy4	Aucoin Lab
OpIE2 Terminator	pOpIE2-Csy4	Aucoin Lab
Hyg^r CDS	pIE1Hygro	Dr. Donald Jarvis (Hollister, Shaper, & Jarvis, 1998)
$Puro^r$ CDS	pOpIE2-Csy4	Aucoin Lab
mAzamiGreen CDS	pmAzamiGreen-N1	Dr. Michael Davidson (Addgene #54798) (Karasawa, Araki, Yamamoto-Hino, & Miyawaki, 2003)

Fragments PBpr001-002 (OpIE2 Promoter), PBpr003-004 (Hyg^r), and PBpr005-006 (OpIE2 Terminator), were isolated via PCR from their respective sources. They were used in two sequential SOE-PCR reactions in an attempt to construct a full Hyg^r expression cassette. The first reaction, merging PBpr001-002 and PBpr003-004 was successful **(Figure 17)**. However, the subsequent attempt to merge this new fragment (PBpr001-004) with the 3' UTR resulted in an erroneous band size **(Figure 18, Figure B 1)**.

Table 7 - Constructed DNA Fragments

Fragment	Description
PBpr001-002	OpIE2 promoter with 5' MfeI site
PBpr003-004	HygR CDS
PBpr005-006	OpIE2 3’ UTR Transcriptional Terminator with 3' NheI site
PBpr007-008	OpIE2 promoter with 5' BamHI site
PBpr009-010	mAG CDS
PBpr011-012	OpIE2 3’ UTR Transcriptional Terminator with 3' SbfI site
PBpr013-014	ATG- PuroR CDS with 5' HindIII site
PBpr015-016	OpIE2 3’ UTR Transcriptional Terminator with 3' NotI site
PBpr001-004	OpIE2 promoter/HygR CDS with 5' MfeI site
PBpr001-006	OpIE2 promoter/HygR CDS/OpIE2 3’ UTR Transcriptional Terminator, with 5' MfeI and 3' NheI
PBpr028-012	Re-amplification of PBpr011-012 with 5' extension
PBpr007-012	OpIE2 promoter/mAG CDS/OpIE2 3’ UTR Transcriptional Terminator, with 5' BamHI and 3' SbfI
PBpr013-031	Re-amplification of PBpr013-014 with 3' extension
PBpr029-030	ATG- PuroR CDS/OpIE2 3’ UTR Transcriptional Terminator, with 5' HindIII and 3' NotI
PBpr013-032	ATG- PuroR CDS/OpIE2 3’ UTR Transcriptional Terminator, with 5' HindIII and 3' NotI
PBga001-002	pMTL85141 prepped for Gibson assembly with PBga003-004
PBga003-004	PBsynbackbone001 prepped for Gibson assembly with PBga001-002
PBga005-006	pMTL-synbackbone prepped for Gibson assembly with PBga007-008
PBga007-008	PBpr001-006 (HygR cassette) prepped for Gibson assembly with PBga005-006
PBga009-010	pMTL-synbackbone-HygR prepped for Gibson assembly with PBga011-012
PBga011-012	PBpr007-012 (mAG cassette) prepped for Gibson assembly with PBga009-010
PBpr007X-012X	Re-amplification of PBpr007-012 with 5' and 3' extensions
PBga021-022	pMTL85141 prepped for Gibson assembly with PBtransIE1

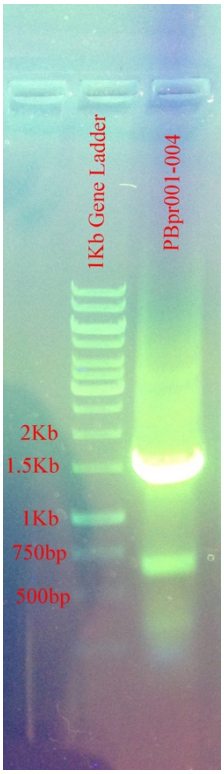

Figure 17 – SOE-PCR of PBpr001-004

Successful SOE-PCR protocol assembling PBpr001-0014. SOE-PCR, Splicing by Overlap Extension Polymerase Chain Reaction.

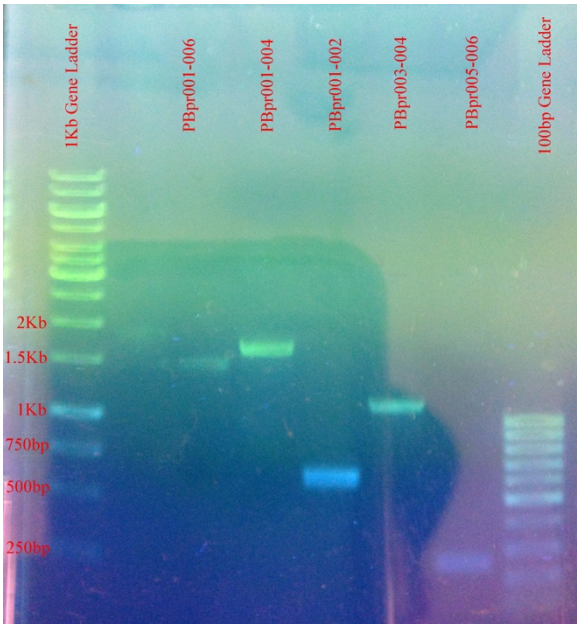

Figure 18 – Attempted SOE-PCR of Hyg^{R} Cassette (PBpr001-006)

Lanes 2-3 show failed SOE-PCR protocol of PBpr001-006. Lanes 4-7 show component pieces. SOE-PCR, Splicing by Overlap Extension Polymerase Chain Reaction.

This SOE-PCR was reattempted with a range of annealing temperatures, 68°C, 69°C, and 70°C, but was unsuccessful **(Figure B 2)**. The resulting bands from these unsuccessful SOE-PCR reactions were consistently too small. With the suspicion that the PBpr005-006 fragment might have insufficient overlap, new primers were designed to further extend the 5' overlap region of this fragment from 20bp to 35bp. The newly extended fragment, PBpr025-026, was used in an SOE-PCR with PBpr001-004 and this reaction was successful **(Figure 19)**.

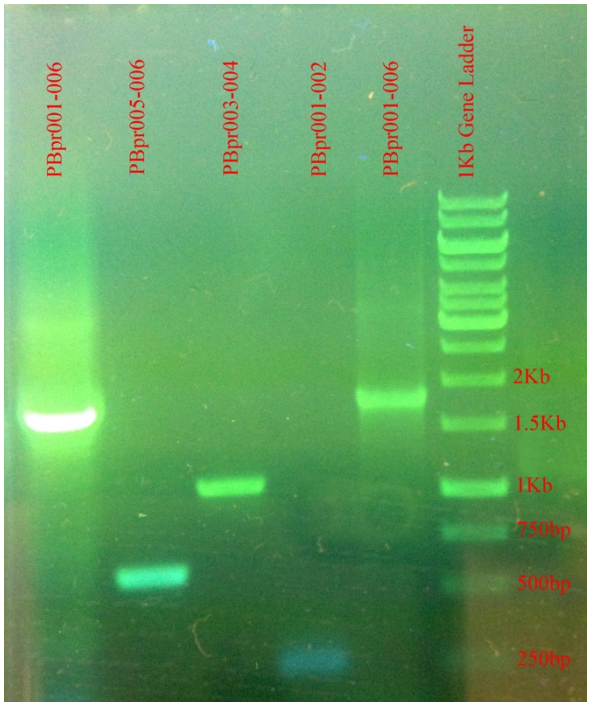

Figure 19 – SOE-PCR of PBpr001-006

Successful assembly of Hyg^R (PBpr001-006) sub-cassette in lanes 1 and 5.
Component pieces are shown in lanes 2-4.

Fragments for construction of the mAG sub-cassette (007-008, 009-010, 011-012) and the $Puro^r$ cassette (013-014, 015-016) were isolated via PCR from their respective sources. The $Puro^r$ CDS (PBpr013-014) was amplified such that it excluded the initial ATG-codon. All but PBpr011-012 were amplified successfully **(Figure 20)**. PBpr011-012 was successfully resolved by using PBpr015-016 as a template in a subsequent amplification. To avoid the issue encountered with the PBpr005-006 fragment, the 5' overlap region of PBpr011-012 was preemptively extended from 20bp to 37bp creating PBpr028-012. A 3-way SOE-PCR was then performed, using fragments PBpr007-008, PBpr009-010, and PBpr028-012, to finish

construction of the mAzamiGreen sub-cassette. The 2-way SOE-PCR between PBpr013-014 and PBpr015-016 was unsuccessful however.

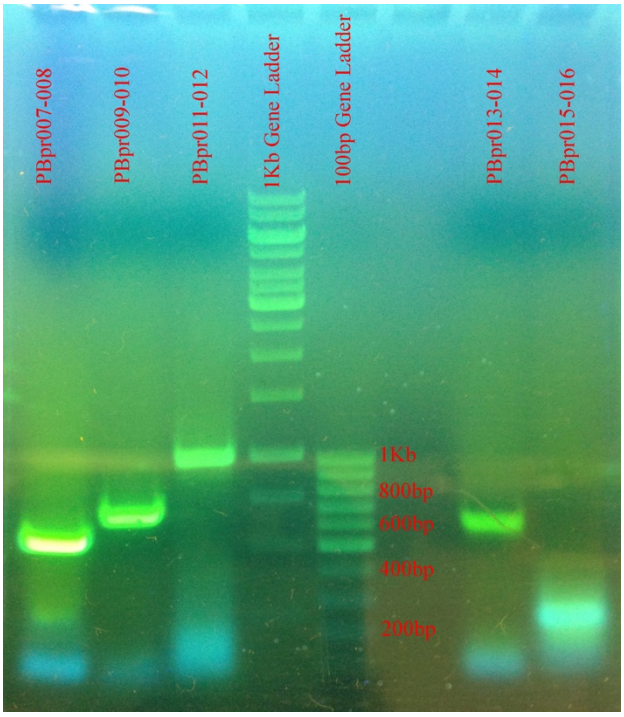

Figure 20 – Components of mAG and Puro^R Sub-Cassettes

Lanes 1-3 show PCR-isolated components of the mAG sub-cassette. Lanes 7-8 show PCR-isolated components of Puro^R sub-cassette. PCR, Polymerase Chain Reaction.

The 3' overlap region of the Puro^r CDS was extended from 20bp to 49bp, creating fragment PBpr013-031 **(Figure 21)**. A new SOE-PCR was attempted between 013-031 and 015-016, using elongation times of both 1 and 2 minutes, but these were unsuccessful **(Figure B 3, Figure B 4)**. A re-amplification of PBpr015-016 using itself as a template showed no bands, indicating that this fragment may have degraded **(Figure 22)**. This fragment was subsequently

re-isolated from the original plasmid. In addition to this, two new sets of external primers were designed that exhibited improved T_a compatibility. A successful SOE-PCR was performed between PBpr013-031 and PBpr015-016 using external primers PBpr029 & PBpr030, and PBpr013 & PBpr032 **(Figure 23)**. The resulting PBpr029-030 codes for a transcriptionally deficient Puro^r gene. Once all three sub-cassettes were constructed, they were sequenced to verify their integrity.

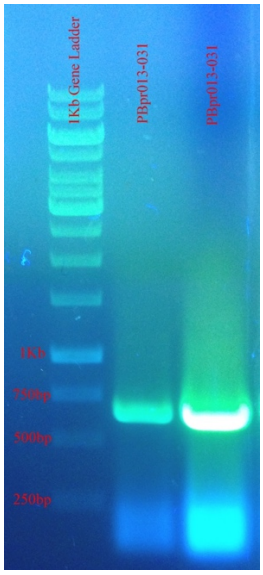

Figure 21 – Puro CDS Extension

Lanes 2-3 show 3'-extended Puro^R CDS. CDS, Coding Sequence.

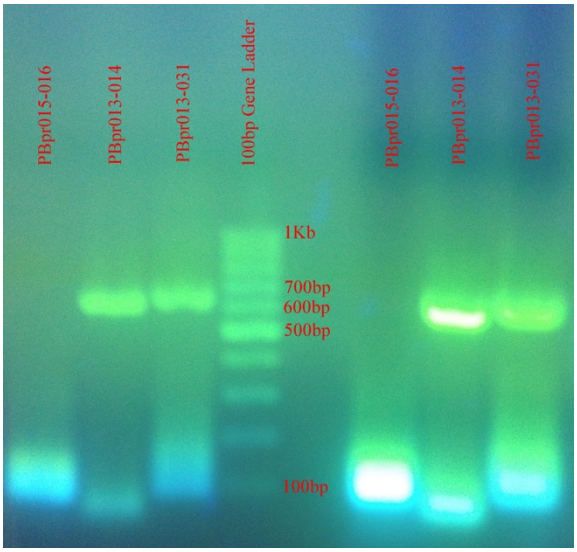

Figure 22 – Diagnostic of Puro^R Fragments

Gel showing unexpected absence of PBpr015-016 Fragment necessary for construction of the Puro^R sub-cassette.

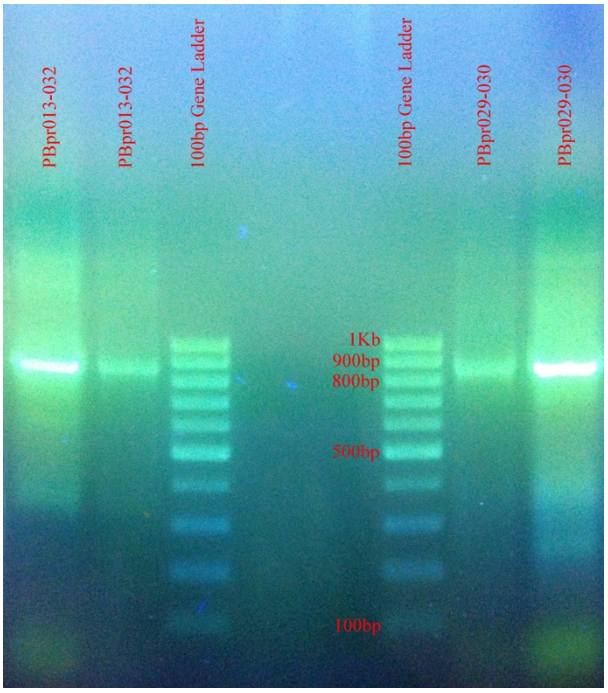

Figure 23 – Assembly of Puro^R sub-cassette

Successful SOE-PCR protocol using redesigned outer primers. SOE-PCR, Splicing by Overlap Extension Polymerase Chain Reaction.

6.1.2 Assembling the Sub-Cassettes

The high GC content of the ϕC31 attachment sites makes them difficult to manipulate through traditional cloning. To avoid issues in this area, a "recombinatorial backbone" was synthesized via Thermo GeneArt which contained the FRT regions, attP regions, attB regions, as well as restriction sites between them.

The contiguous region containing the FRT sites and the attP sites was cloned into the plasmid pMTL85141 via Gibson Assembly. pMTL85141 is a basic *E. coli* cloning vector

exhibiting chloramphenicol resistance, and was acquired from the Aucoin Lab (University of Waterloo). Primers used were PBga001, PBga002 (on pMTL85141) and PBga003, PBga004 (on the synthesized DNA PBsynbackbone001). The assembly (pMTL-synbackbone) was transformed into *E. coli* DH5α, grown overnight, and miniprepped.

The Hyg^r cassette was cloned into pMTL-synbackbone via Gibson Assembly, such that it is placed in between the two FRT sites. Primers used were PBga005, PBga006 (on pMTL-synbackbone) and PBga007, PBga008 (on Hyg^r cassette). The assembly (pMTL-synbackbone-Hyg) was transformed into *E. coli* DH5α, grown overnight, and miniprepped.

Next, preparations to clone the mAzamiGreen cassette into pMTL-synbackbone-Hyg were made. However, the initial PCR with Gibson primers (PBga009, 010, 011, 012) resulted in smeary diffuse products for PBga011-012 **(Figure 24)**. Bands were extracted and a Gibson reaction was attempted, but was unsuccessful. An unsuccessful attempt was made to form a Gibson assembly between PBga009-010 and the initial mAG cassette (PBpr007-012). An attempt was made to re-prep PBpr007-012 for Gibson assembly using PBga011 and PBga012 at a range of temperatures (60°C, 66°C, 72°C), with the thought that lower annealing temperatures might allow for a successful amplication. This still resulted in some high-molecular-weight smearing, but with a distinct band edge **(Figure 25)**. The band resulting from the 72°C reaction was extracted and used in a Gibson assembly with PBga009-010 and transformed into *E. coli*. This reaction appeared to be successful, although at a relatively low transformation efficiency (12 colonies grew successfully). Three colonies were backed up and miniprepped.

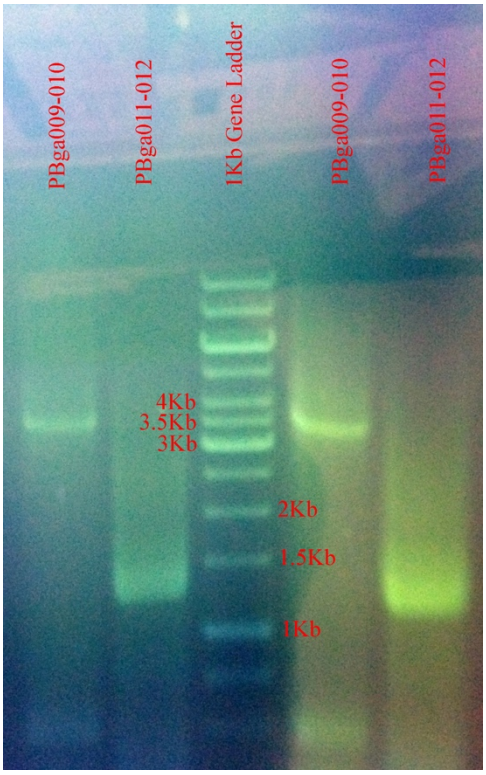

Figure 24 – Gibson Prep of mAG Cassette

Amplification of the mAG sub-cassette with Gibson primers (PBga011-012) results in diffuse bands, lanes 2 & 5. Gibson prep of pMTL-synbackbone-Hyg^{R} shown in lanes 1 & 4.

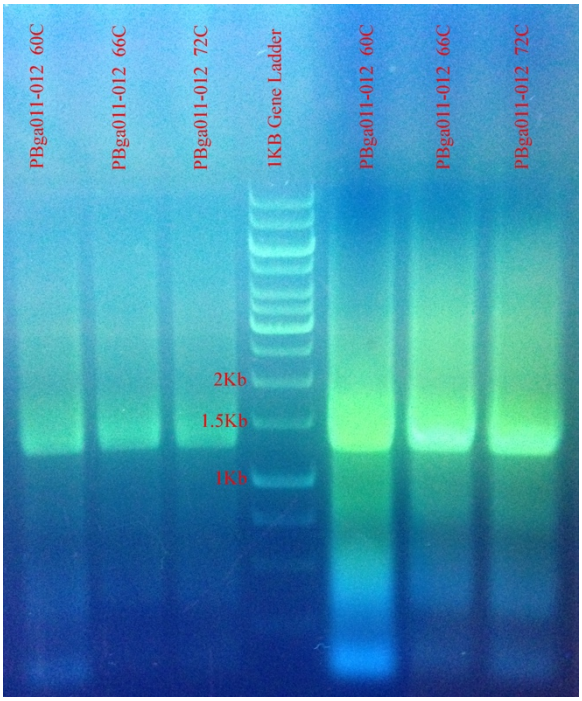

Figure 25 – Temperature Course of Gibson Prep, PBga011-012

Gibson prep of PBga011-012 performed at 60°C, 66°C, and 72°C. All result in diffuse bands with high-molecular weight smears.

pMTL-synbackbone-Hyg-mAG was prepped for Gibson assembly using Gibson primers PBga013 and PBga014. Unexpectedly, this PCR showed no results when run on a gel. Gibson prep of PBpr029-030 with primers PBga015-016 also showed nothing. These PCR reactions were retried at new temperatures (67°C, 72°C) but still showed no bands **(Figure 26)**. A control re-amplification of PBpr029-030 was included and successful.

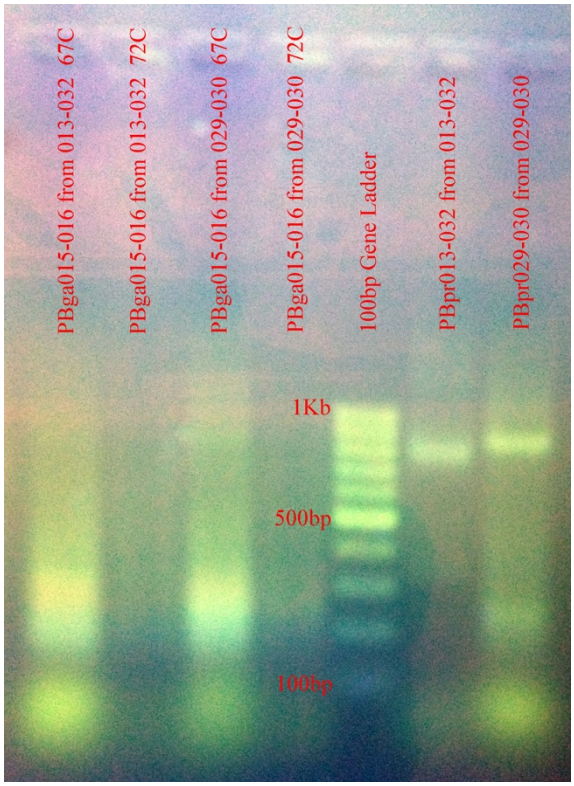

Figure 26 – Temperature Course of Gibson Prep on PBga015-016 from Two Templates

Fragments PBpr013-032 and PBpr029-030 were both used as templates for an attempted Gibson prep of PBga015-016. Reactions were carried out at 67°C and 72°C. Control re-amplifications of PBpr013-032 and PBpr029-030 are shown in Lanes 6 and 7 respectively.

Gibson assemblies were then sequenced to confirm the proper construction of pMTL-synbackbone-Hyg-mAG. However, sequencing results showed an incorrect assembly; the mAG cassette appeared to be absent from the plasmid.

PBga009-010 and PBpr007-012 were both reamplified, and another Gibson assembly was attempted, but no colony growth was detected. Extension primers (PBga009X, PBga010X) were designed in an effort to increase PBga009-010's overlap regions with PBpr007-012.

However, the PCR amplification using these primers showed no bands. Alternate annealing temperatures were explored (60°C, 66°C, 72°C) but still showed no bands. It was then decided to attempt extending the PBpr007-012 fragment instead. Extension primers PBpr007X and PBpr012X were used in a PCR on PBpr007-012 at 64°C, 66°C, and 68°C. All showed proper sized bands, but with some full-lane smearing. The 64°C sample appeared to be the clearest, and was extracted for later use. However, it was noted in the nanodrop that the 260/230 ratio was demonstrably low, at 0.04. A Gibson assembly was attempted between PBga009-010 and PBpr007X-012X, but was unsuccessful.

It was decided to eschew Gibson assembly in favour of restriction cloning. Double digests of pMTL-synbackbone-Hyg and PBpr007-012 were performed using BamHI and SdaI, and the fragments were then ligated and transformed into *E. coli*. However, no colony growth was seen. Instead of re-attempting with PBpr007-012, a restriction cloning protocol was used on the Puro^r cassette. pMTL-synbackbone-Hyg and PBpr029-030 were digested with HindIII and NotI enzymes, ligated, and transformed into *E. coli.* Minipreps of three different colonies were prepared, but nanodrop readings again showed very low 260/230 readings (0.31-0.62) and low DNA concentration (2.24-17.12 ng/µl).

6.1.3 PiggyBac Transposase

Plasmids pBSII-IE1-orf, containing the piggyBac transposase CDS, and pXLBacII, containing the piggyBac terminal repeat domains were obtained from Dr. Donald Jarvis (University of Wyoming). Although these plasmids were listed as containing Amp^r regions,

but no colonies resulted when plated on ampicillin. Transformants were plated on kanamycin, tetracycline, and chloramphenicol plates. pXLBacII-transformants produced colony growth on kanamycin, and these colonies were miniprepped to recover the plasmid. However, pBSII-IE1-orf did not produce colonies on any of these plates.

To confirm that the plasmid labelled pXLBacII did indeed contain piggyBac TRDs, it was sequenced using TRD-binding primers. Sequencing results confirmed that this plasmid did contain the piggyBac TRD regions.

To determine whether the sample labelled pBSII-IE1-orf was the proper plasmid, a restriction map using BamHI was produced. This should have produced fragments of 2010bp and 3513bp, but this was not reflected in the gel **(Figure 27)**. Additionally, a PCR was performed using primers designed to amplify a 755bp inner portion of the piggyBac CDS (PBpr023, PBpr024) which resulted in no amplified fragment. These two procedures confirmed that the piggyBac Transposase CDS was not present in the samples received.

To remedy the lack of a piggyBac CDS, a new DNA fragment, PBtransIE1, was synthesized as a gBlock via IDT Inc (Integrated DNA Technologies Inc., Skokie IL). This fragment contained an Sf9-optimized piggyBac Transposase CDS with an upstream IE1 promoter. This synthesized fragment required the addition of a 3' UTR, and needed to be placed in a plasmid backbone.

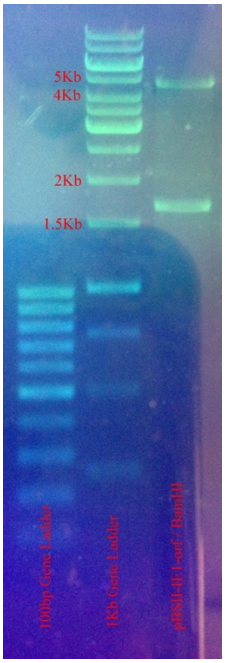

Figure 27 – Restriction Map of Sample Labelled pBSII-IE1-orf

Digestion of sample with BamHI resulted in a banding pattern not consistence with the expected restriction map of pBSII-IE1-orf.

Primers PBpr033 and PBpr034 were used in a PCR amplification of PBtransIE1 to prepare it for an SOE-PCR reaction. However, the resulting gel showed smearing, as well as a double band pattern **(Figure 28)**. A second attempt at this reaction, using redesigned primers PBpr037 and PBpr038, showed reduced smearing but still presented a double-banding pattern **(Figure 29)**.

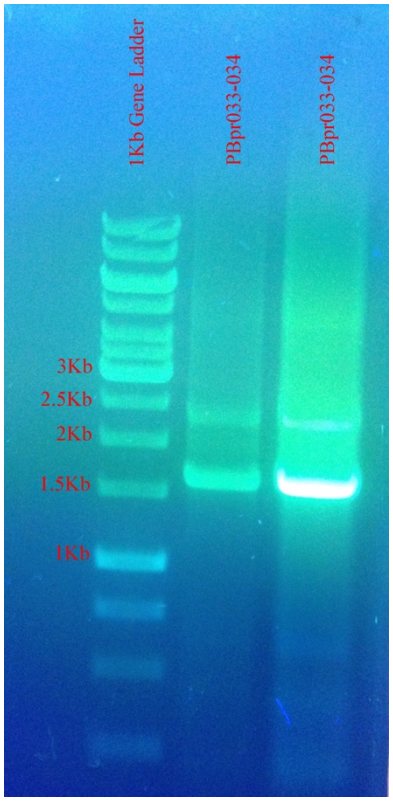

Figure 28 – Amplification of PBtransIE1 (PBpr033-034)

Amplification of PBtransIE1 with primers PBpr033 and PBpr034 results in double-band pattern.

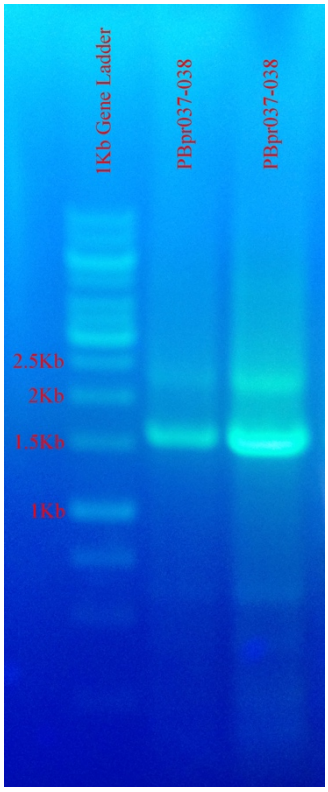

Figure 29 – Amplification of PBtransIE1 (PBpr037-038)

Amplification of PBtransIE1 with redesigned primers PBpr037 and PBor038 still results in double-band pattern.

It was subsequently decided to move PBtransIE1 into a plasmid first, and append the 3' UTR afterwards. Gibson primers PBga021 and PBga022 were used to amplify the pMTL backbone, but this produced no amplified fragment. These primers were subsequently used on pMTL-synbackbone with a similar negative result. A second attempt at a PCR amplification of pMTL using PBga021 and PBga022, but with the annealing temperature raised from 59°C to 72°C, resulted in a properly sized band (**Figure 30)**.

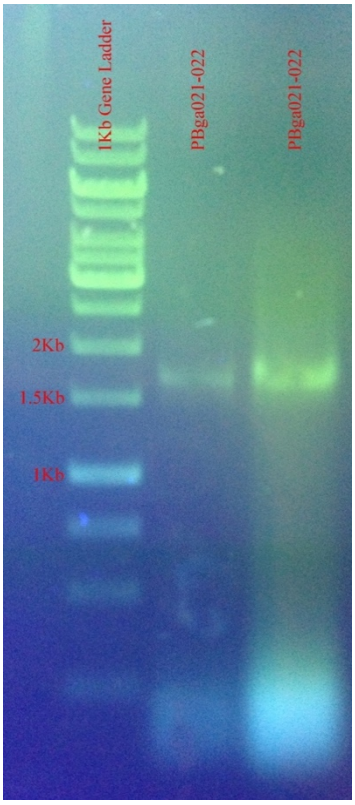

Figure 30 – Gibson Prep of pMTL as PBga021-022, 72°C

Amplifying plasmid pMTL with Gibson primers PBga021 and PBga022 results in defined band after raising temperature from 59°C (not shown) to 72°C (shown)

A Gibson assembly was attempted between PBga021-022 and PBtransIE1, but no colony growth was detected. An amplification of PBtransIE1 with primers PBga023 and PBga24 resulted in a similar double band pattern as seen before.

A series of forward primers were designed (PBpr040, PBpr042, PBpr043) that bind to intermittent regions along PBtransIE1. These were paired with reverse primer PBpr041, and

amplified via PCR. The result was a series a series of double bands, shifted in size. The space between the bands, however, remained consistent **(Figure 31)**.

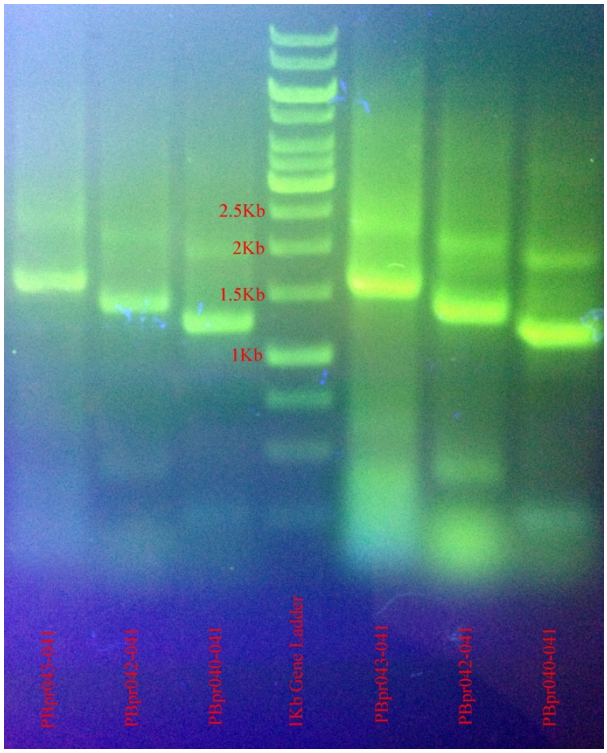

Figure 31 – Forward Primer Series on PBtransIE1

A series of PCR amplications on PBtransIE1 using various forward primers (PBpr040, PBpr042, PBpr043), always with reverse primer PBpr041, shows double-band pattern that shits in size relative to the binding position of the forward primer. PCR, Polymerase Chain Reaction.

6.2 Discussion

The cassette construction process was heavily marred with unexpected complications. Thus, a great deal of the methodology and results was dedicated to attempts at troubleshooting these complications. While ultimately unsuccessful in completing the final product, a number of useful pieces of information were collected through this process.

SOE-PCR protocols often suggest that overlap regions of between 20-30bp are sufficient for a successful reaction (Warrens, Jones, & Lechler, 1997). While this was sufficient for several SOE-PCR reactions, ones that became troublesome were invariably remedied by extending the overlap regions to >=35bp. When appending a 3' UTR onto the Hygro^r CDS a 35bp overlap was used, while the same reaction in the Puro^r cassette required a 49bp overlap. In retrospect, failed SOE-PCRs often surfaced when instituting a 3' UTR. Extending the overlap regions here appeared to successfully counteract the low binding affinity of these sequences.

The Gibson Assembly phase of cassette construction was initially successful, with the recombinatorial backbone slotting into pMTL and the Hygro^r cassette being cleanly assembled afterwards. However, things became problematic when attempting to assemble the mAG cassette (PBpr007-012). The presence of high molecular weight smears, such as those seen in Figure 24 and Figure 25, are often caused by plasmid template contamination. However, in these instances the template (PBpr007-012) was 24bp shorter than the predicted product, so template contamination could be safely ruled out.

There was some clear evidence of primer dimers forming in reactions run on these gels, so some element of mispriming was a possibility. The temperature series shown in **Figure 25** did

show some qualitative improvement with respect to primer dimers and smearing. This would indicate that increasing the stringency of the PCR reaction was positively improving the result. Unfortunately, a T_a higher than 72°C was untenable.

A similar type of smearing was found when attempting to amplify the Puro^r cassette (PBpr029-030) with PBga015 and PBga016 **(Figure 26)**. After the fact, it was noted that the reverse primers in the PBga011-012 and PBga015-016 PCR preps exhibited very high ΔG score for homodimers, -12.64kcal and -9.26kcal respectively (calculated via Benchling). It is a reasonable hypothesis that 3' dimerization was contributing heavily to the unsuccessful results shown in **Figure 25** and **Figure 26**.

The highly AT rich 3' regions in conjunction with the presence of palindromic restriction sites made designing useful primers very difficult for these cassettes, as these characteristics lend themselves to dimerization problems. However, the restriction sites offered an alternative approach to assembly. A pseudo-successful attempt at integrating the Puro^r cassette into pMTL-synbackbone-Hygro was made via a double digest/ligation. However, upon nanodropping the miniprepped plasmids, the 260/230 readings were prohibitively low.

Low 260/230 readings typically indicate a contaminant of some kind, such as carbohydrates, phenols, or glycogen, has entered the sample. In instances where a commercial miniprep column is used, and in particular an older kit such as the one used here, residual guanidine contamination was likely the issue (Scientific, 2007).

The lack of a piggyBac transposase CDS, which was originally to be sourced from the plasmid pBSII-ie1-orf, instigated a series of problems. The alternative to sourcing this CDS

from a plasmid was to have the sequence synthesized. An advantage of this method was it allowed for the piggyBac CDS to be codon optimized for expression in Sf9 cells. The sequence designed for synthesis, PBtransIE1, included an upstream IE1 promoter. However, as a cost-cutting measure, the 3' UTR necessary for proper gene transcription was omitted – a decision that had unexpected ramifications.

Amplification of PBtransIE1, be it preparation for SOE-PCR (using PBpr033/PBpr034 and PBpr037/PBpr038) or Gibson Assembly (PBga023/PBpr024), invariably resulted in a double-band pattern **(Figure 28 and Figure 29)**. Initially, it was attempted to perform a Gibson Assembly without doing a PCR preparation on PBtransIE1, which would have avoided the need to resolve the double banding issue. However this was unsuccessful.

The brightest band in the double band was showing up at ~1600bp, whereas the predicted product should be ~2300bp. It is likely that one or both of the primers was mispriming relative to the template, resulting in a shorter amplicon. To gain insight into this, a series of forward primers (PBpr040, PBpr042, PBpr043) was used to amplify PBtransIE1, such that each reaction had a differently sized amplicon **(Figure 31)**. The double banding pattern remained present, but each lane was shifted in size. The size shift indicates that the forward primers were binding properly. Therefore, the double band pattern must have been the result of mispriming by the reverse primer.

Chapter 7

Limiting Dilution of Sf9 Cell Line

During the construction of the recombinant cassettes, and the trouble-shooting therein, a series of Limiting Dilution test runs were enacted. This would help determine optimal parameters for clonal isolation of the recombinant Sf9 cell line. Parameters of interest were: fresh vs. conditioned media, supplementation with Fetal Bovine Serum (FBS), and plate material.

Limiting dilution functions by diluting a healthy culture of cells, and aliquoting them such that only 1 new cell is present in each new culture vessel. In this instance, a culture vessel consisted of 1 well in a 96-well plate. Cells are then grown to confluency and a new maintenance culture is established.

Two types of VWR tissue-culture-treated 96 well plates were examined: flat-bottom and U-bottom. Media treatments consisted of 100% conditioned Sf900III media, and a 50/50 mixture of conditioned and fresh Sf900III media. Trials of FBS supplementation were performed at levels of 0%, 2%, and 10% FBS by volume.

7.1 Materials and Methods

7.1.1 Preparing Conditioned Media

Allow an Sf9 suspension culture to grow to ~$3x10^6$ cell/ml in SF900III nutrient media. Transfer culture into 50ml conical Falcon centrifuge tubes. Pellet cultures at 150xG for 5 minutes. In a biosafety hood, vacuum filter the media through a 0.2μm pore filter. Securely cap conditioned media and store at 4°C.

7.1.2 Limiting Dilution

In a biosafety hood, using sterile serological pipettes, initial media treatments were prepared in 20ml volumes. A cell dilution was prepared beginning with a suspension culture of Sf9 cells at ~$5x10^5$ cells/ml. Using 27°C fresh Sf900III media, serial 1/10 dilutions were performed to produce a suspension of $5x10^2$ cells/ml. 100μl of the Sf9 cell dilution was transferred into each media preparation to produce a cell density of 2.5 cells/ml. Tubes were gently inverted to mix cells into media. Cell/media preparations were aliquoted into 96-well tissue-culture treated plates, at a volume of 200μl/well. Plates were covered and incubated in an airtight container with a clean-bath treated sponge at 27°C.

Wells were inspected within 12 hours of plating to indicate wells containing single cells. Inspections were carried out using a Zeiss Axiovert 40 C phase contrast microscope. Cells were inspected weekly to monitor growth, and replace media.

7.2 Results

Trial 1 for performing limiting dilution on an Sf9 culture examined two different media preparations, as well as two different tissue-culture treated plate types **(Table 8)**. While the flat-bottom plates were fully transparent, the round-bottom plates exhibited a slight translucence. While it was initially suspected that cells would still be discernable through them, this was incorrect. Therefore, all entries for the round bottom plates are listed as Not Determinable (ND) **(Table 9)**.

Table 8 - Trial 1, Limiting Dilution Treatments

Plate	Conditioned Media (%)	Fresh Media (%)	Plate Type
1	100	0	Flat-Bottom
2	100	0	Round-Bottom
3	50	50	Flat-Bottom
4	50	50	Round-Bottom

Table 9 – Trial 1, Limiting Dilution Cell Counts

Week-over-week data is presented as a percentage of wells containing isolated, viable cells.
ND, Not Determinable

Plate	1	2	3	4
Total Cells Plated	44	ND	43	ND
Wells Containing Isolated, Viable Cells	25	ND	15	ND
Week 1 – Viable Cells Remaining (%)	0	ND	33.33	ND
Week 1 – Cells Showing Division (%)	0	ND	13.33	ND
Week 2 – Viable Cells Remaining (%)	0	ND	0	ND
Week 2 – Cells Showing Division (%)	0	ND	0	ND

Trial 2 for performing a limiting dilution was performed entirely on flat-bottom polystyrene tissue-culture-treated 96-well plates. Six different media formulations were used in this trial **(Table 10)**. At two weeks, plate C was found to be contaminated by a fungus, and was subsequently discarded. All entries for past this point on the plate were labelled as C (contaminated) **(Table 11)**.

Table 10 - Trial 2, Limiting Diluiton Treatments

FBS, Fetal Bovine Serum

Plate	Conditioned Media (%)	Fresh Media (%)	FBS (%)
A	100	0	0
B	98	0	2
C	90	0	10
D	50	50	0
E	49	49	2
F	45	45	10

For both trials, week-over-week data was only collected for wells containing isolated cells. Information was compartmentalized based on whether a well contained cells that appeared alive (viability), and whether there appeared to be new cells present during inspection relative to the previous inspection (division). During this inspection process, viability was assessed visually via phase-contrast microscopy; cells that appeared qualitatively healthy were deemed to be viable.

Table 11 – Trial 2, Limiting Dilution Cell Counts

Week-over-week data is presented as a percentage of wells containing isolated, viable cells.

C, Contaminated

Plate	A	B	C	D	E	F
Total Cells Plated	49	34	47	56	41	32
Wells Containing Isolated, Viable Cells	19	15	21	33	18	20
Week 1 – Viable Cells Remaining (%)	21.05	46.66	42.85	36.36	72.22	90
Week 1 – Cells Showing Division (%)	0	0	4.76	6.06	11.11	20
Week 2 – Viable Cells Remaining (%)	0	20	C	6.06	11.11	25
Week 2 – Cells Showing Division (%)	0	0	C	0	0	5
Week 3 – Viable Cells Remaining (%)	0	0	C	0	5.55	5
Week 3 – Cells Showing Division (%)	0	0	C	0	0	0
Week 4 – Viable Cells Remaining (%)	0	0	C	0	0	0
Week 4 – Cells Showing Division (%)	0	0	C	0	0	0

7.3 Discussion

Limiting Dilution, while time consuming, is generally considered to be the optimal method for cell-line development when producing a clonal-isolate is a priority. By mechanically isolating individual cells, great confidence can be had as to the monoclonal nature of the resultant population. However, such an extreme dilution can cause a great deal of stress on a cell. It has been noted that not all established insect cell lines are compatible with the limiting dilution procedure, and that even in successful trials upwards of 50% of isolates will die (Harrison & Jarvis, 2010). Nevertheless, several lepidopteran lines, including the Sf9 and

High-Five lines, have successfully undergone this procedure in the past (Fernandes et al., 2012; Steele et al., 2017).

In the literature, there are some variations in the approach to limiting dilution. One example using Sf9 cells utilized fully conditioned media supplemented with 10% FBS (Fernandes et al., 2012). A trial with the High-Five cell line was successfully implemented using a 50/50 split of fresh and conditioned media, with no FBS supplementation (Steele et al., 2017).

Additionally, there are some variations on the exact definition of conditioned media; the former listed the media as the supernatant of exponentially growing cells, while the latter listed CM as media from 48-hour growth cultures (Fernandes et al., 2012; Steele et al., 2017). Other groups have harvested CM from cultures between 2-5x10^6 cells/ml, or from 3 day growth cultures (Calles et al., 2006; Doverskog et al., 2000).

Unfortunately, no healthy cell lines were successfully produced in these trials. The reasons for this are unclear. However, by examining the number of cells that survived to the 1 week mark in the second trial, a number of suggestions can be made for further work. Detailed statistics for these comparisons can be found in Appendix C.

In the treatment set with a 50/50 media split (plates D-F), the addition of FBS provided a statistically significant boost in cell health. Viability improved from 36.4% in plate D to 72.2% in plate E and 90% in plate F (p values of 6.6x10^{-3} and 5.0x10^{-6} respectively) **(Table C 3,**

Table C 4**)**. The evidence for improved cell viability due to FBS is weaker in the plates with fully conditioned media. Viability improved from 21.1% in plate A to 46.7% in plate B and 42.9% in plate C (p values of $6.5x10^{-2}$ and $7.3x10^{-2}$ respectively) **(Table C 1, Table C 2)**.

Interestingly, increasing the amount of FBS from 2% to 10% did not have a pronounced effect on cell viability. The difference between plate B & C, and E & F, were not statistically significant. Further data would be useful to be certain that this conclusion holds. Assuming it does, however, this would suggest that limiting dilutions should be carried out using 2% FBS rather than 10%. The added difficulty of weaning a cell-line off of 10% FBS once established does not appear to be accompanied by any advantage during the development process.

A comparison between plate D and A shows an increase in viability from 21.1% to 36.4% **(Table C 7)**. However, this comparison falls well short of p=0.05. When data from plates 1 and 3 from trial 1 are included, a statistically significant increase of 9.1% to 35.4% was seen when going from fully conditioned media to split media (p=0.001) **(**

Table C 10**)**. Additionally, comparing plate F and C wherein media conditions change but FBS remains constant showed an increase from 42.9% to 90% in favour of split media (p=0.0005) **(Table C 9)**. These results would suggest that a 50/50 split of conditioned and fresh media is advantageous for Sf9 viability in limiting dilution conditions.

Chapter 8

Conclusions & Recommendations

While the final goal of a ϕC31-competent Sf9 cell line was not realized, a number of useful takeaways can be made from this project. Firstly, the design choice of utilizing mAzamiGreen and mKOk as a paired fluorescent reporter system appears to be a successful one. Each of the three possible expression permutations (singular mAG, singular mKOk, and co-expressed mAG+mKOk) were clearly differentiable. No treatment saw more than 0.7% of expressing cells register in a competing treatment's gate. This result also demonstrates that mKOk exhibits strong enough flouresence to be used with mAG in any future co-expression studies.

Limiting dilution of an Sf9 cell culture remains a difficult procedure. Results from this project may provide suggestions on how to maximize the chance of successfully creating a clonal isolate. There is some evidence to suggest that an application of mixed fresh/conditioned media better promotes cell viability when compared to a treatment of purely conditioned media. This experiment only tested two permutations - 100% conditioned and 50/50 fresh/conditioned – and so exploration of other ratios may prove to be fruitful in the future. Additionally, trials examining to what effect the degree of media conditioning has on Sf9 cells in the context of limiting dilution would be useful. Some investigation has been performed in the context of standard growth conditions, but not in single-cell growth conditions (Doverskog et al., 2000).

Unsurprisingly, evidence showed that the addition of FBS improved Sf9 viability week-over-week during a limiting dilution protocol. What is interesting, however, is that a 10% FBS

treatment was not significantly better at improving cell viability week-over-week than a 2% FBS treatment. When taken with the fact that any successfully isolated cell line will need to be subsequently weaned off of FBS, a low-concentration FBS supplement appears to be advisable. However, given that the end goal of the limiting dilution protocol was unsuccessful, further work is required on this front.

Ultimately, construction of the ϕC31-competent gene cassette was unsuccessful. In retrospect, the construction failures often revolved around manipulating the AT-rich 3' UTR fragments. The extremely low GC-content of these regions made them difficult to handle in isolation, and the construction plan did not adequately account for this. This was exacerbated by including restriction sites, which are inherently palindromic, at the 3' ends of these regions. In the end, the design decision to include "backwards compatibility" for restriction cloning protocols proved to be a poor one.

Given that restriction cloning is an aging technology and restriction sites may interfere with constructing genes via Gibson assembly by promoting 3' fragment dimerization, it is recommended to eschew, or minimize, the inclusion of restriction sites in future synthetic biology endeavours. Furthermore, it is recommended that handling AT-rich transcriptional terminator regions in isolation should be avoided whenever possible. Their low T_m and predilection towards mispriming make them difficult assemble. It is also recommended that further work be done to optimize and standardize a media formulation for Sf9 limiting dilution protocols. This includes exploration of Conditioned Media/Fresh Media ratios, as well as the degree of media conditioning.

At the time the experimental section of this project was concluded, the genetic components of the proposed ϕC31-competent Sf9 cell line remain in a state of semi-completion. Based on the design considerations and the limited results attained, the end goal of a functional cell line still appears to be achievable. In conclusion, it is recommended that pursuing the completion of this work is still a worthwhile endeavour.

Appendix A

Oligonucleotide Primers

Oligonucleotide primers are used in PCR amplification protocols. Labels include the code "pr" if the intended use was for PCR or SOE-PCR reactions. Labels including the code "ga" were intended for use in preparing a Gibson Assembly.

Table 12 - Oligonucleotide Primers

Dir. – Direction: **F**orward, **R**everse

Name	Dir.	Sequence
PBpr001	F	atcgacaattgggtacccatgatgataaacaatgtatggtgctaatgttgc
PBpr002	R	agacgtcgcggtcagctcaggctttcccataacagatgctgttcaactgtgtttacc
PBpr003	F	atgggaaagcctgagctgacc
PBpr004	R	ctaagataatgacagttcctttgccctcggacgag
PBpr005	F	aggaactgtcattatcttagtttgtattgtcatgttttaatacaatatg
PBpr006	R	atcgcgctagccgtcagcaacgaccctgc
PBpr007	F	atcgcggatcccatgatgataaacaatgtatggtgctaatgttgc
PBpr008	R	cttcatctcgggcttgatcacgctcaccataacagatgctgttcaactgtgtttacc
PBpr009	F	atggtgagcgtgatcaagcccgag
PBpr010	R	ctaagataatgacagttacttggcctggctgggcagc
PBpr011	F	agtaactgtcattatcttagtttgtattgtcatgttttaatacaatatg
PBpr012	R	actatcctgcaggcgtcagcaacgaccctgc
PBpr013	F	atcgtaagcttaccgagtacaagcccacgg
PBpr014	R	ctaagataatgacagtcaggcaccgggcttgc
PBpr015	F	cctgactgtcattatcttagtttgtattgtcatgttttaatacaatatg
PBpr016	R	attgcgcggccgccgtcagcaacgaccctgc
PBpr017	F	cgttggaagacacagcatgacc
PBpr018	R	ttccgagggctccgataaatgg
PBpr019	F	tagccatttatcggagccctcg
PBpr020	R	ttcgttagccagtccacagagg
PBpr021	F	gccatttatcggagccctcgacttaccgcattgacaagca
PBpr022	R	tcatgctgtgtcttccaacgccgggctgcaggaattcgat
PBpr023	F	caaaatgtgccacacttgcagc
PBpr024	R	ttcagtgcagagactcggctac
PBpr025	F	ctcgtccgagggcaaaggaactgtcattatcttagtttgtattgtcatg

PBpr026	R	atcgcgctagccgtcagc
PBpr027	F	atcgacaattgggtacccatgatgataaac
PBpr028	F	gctgcccagccaggccaagtaactgtcattatcttag
PBpr029	F	atcgtaagcttaccgagtacaagcccacggtgc
PBpr030	R	attgcgcggccgccgtcag
PBpr031	R	catattgtattaaaacatgacaatacaaactaagataatgacagtcaggcaccgggc
PBpr032	R	attgcgcggccgccg
PBpr033	F	cgatgtctttgtgatgcg
PBpr034	R	caatacaaactaagataatgacagttagaagcaggactgg
PBpr035	F	ccagtcctgcttctaactgtcattatcttagtttgtattgtcatgttttaatacaatatg
PBpr036	R	cgtcagcaacgaccctgc
PBpr037	F	cgatgtctttgtgatgcgcc
PBpr038	R	ttagaagcaggactggcacatg
PBpr039	F	cgcgagcacaacatcgacatgtgccagtcctgcttctaactgtcattatc
PBpr040	F	atgggctgctccctggac
PBpr041	R	ttagaagcaggactggcacatgtc
PBpr042	F	ccgactgttttcgtatccgctc
PBpr043	F	caaaatgtgccacacttgcagc
PBpr044	R	tgtcgatgttgtgctcg
PBpr045	R	cagatcactttcttgcac
PBpr007X	F	ggatcccatgatgataaacaatg
PBpr012X	R	cctgcaggcgtcagcaa
PBga001	F	tatcggagccctcggaagaggcaaatgaaatagattgac
PBga002	R	ctggccgtcgttttacatcttccgcttcctcgct
PBga003	F	agcgaggaagcggaagatgtaaaacgacggccag
PBga004	R	tctatttcatttgcctcttccgagggctccgata
PBga005	F	cgttgctgacggctagcgggtggatggtttaagaa
PBga006	R	tcatgggtacccaattgagacagcgtatctcacg
PBga007	F	cgtgagatacgctgtctcaattgggtacccatgatga
PBga008	R	tcttaaaccatccacccgctagccgtcagcaacg
PBga009	F	ttgctgacgcctgcaggaacctgttaagcgcacc
PBga010	R	ttatcatcatgggatcctggttcctctatatcccc
PBga011	F	gggatatagaggaaccaggatcccatgatgataaacaatg
PBga012	R	ggtgcgcttaacaggttcctgcaggcgtcagcaa
PBga013	F	ttgctgacggcggccgctatgtgaagcacagcatagcc
PBga014	R	ttgtactcggtaagcttcgcccccaactgagaga
PBga015	F	tctctcagttgggggcgaagcttaccgagtacaagcccacgg

PBga016	R	atgctgtgcttcacatagcggccgccgtcagcaa
PBga017	F	tatcggagccctcggaaacttaccgcattgacaagc
PBga018	R	ctggccgtcgttttacacccgggctgcaggaatt
PBga019	F	aattcctgcagcccgggtgtaaaacgacggccag
PBga020	R	ttgtcaatgcggtaagtttccgagggctccgata
PBga021	F	tgtgccagtcctgcttctaagaggcaaatgaaatagattgac
PBga022	R	gtcggcgcatcacaaagacatcgtcttccgcttcctcgct
PBga023	F	agcgaggaagcggaagacgatgtctttgtgatgc
PBga024	R	tctatttcatttgcctcttagaagcaggactggc
PBga025	F	cagggtcgttgctgacggaggcaaatgaaatagattg
PBga026	R	acaatacaaactaagatttagaagcaggactggc
PBga027	F	gccagtcctgcttctaaatcttagtttgtattgtcatgtttt
PBga028	R	tctatttcatttgcctccgtcagcaacgaccctg
PBga007X	F	ggatcccatgatgataaacaatg
PBga009X	F	acccgaaaagcagggtcgttgctgacgcctgcagg
PBga010X	R	cattagcaccatacattgtttatcatcatgggatcctggttc
PBga012X	R	cctgcaggcgtcagcaa
PBga021X	F	gcgagcacaacatcgacatgtgccagtcctgcttctaag
PBga022X	R	caataacctacaaaaatgtcggcgcatcacaaagac

Appendix B

Electrophoresis Gels

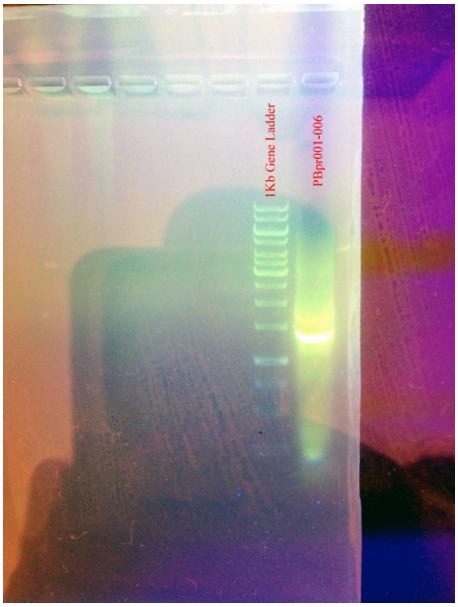

Figure B 1 – Unsuccessful SOE-PCR of PBpr-001-006

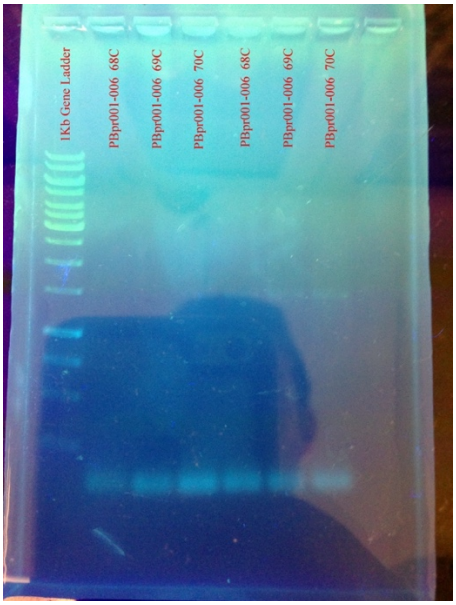

Figure B 2 – Temperature Series of SOE-PCR, PBpr001-006

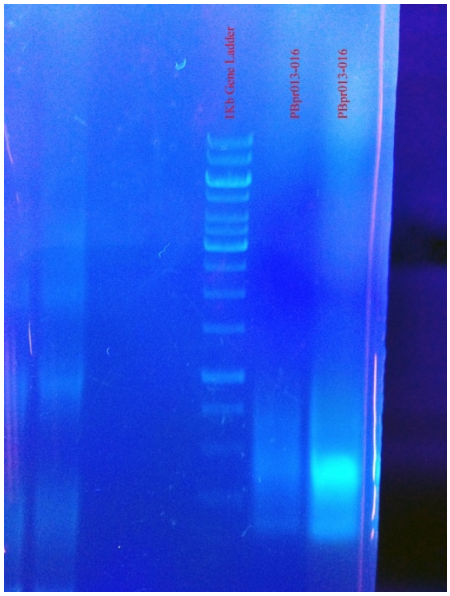

Figure B 3 – Unsuccessful SOE-PCR, PBpr013-016, 1min T_E

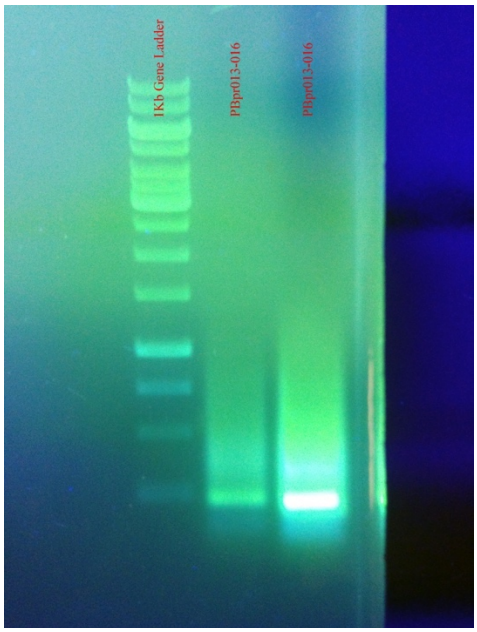

Figure B 4 – Unsuccessful SOE-PCR, PBpr013-016, 2min T_E

Appendix C

Statistical Comparisons of Limiting Dilution Treatments

All tests of statistical significance referenced in Chapter 7 are T-tests comparing mean values, while assuming unequal variances. Values were calculated in Microsoft Excel. All P-values referenced in Chapter 7 are one-tailed (pertinent cells are highlighted).

Table C 1 - Means Comparison of 100% CM (A) vs. 98% CM, 2% FBS (B)

Means comparison of cell viability in isolated wells after 1 week, between plates A & B in Trial 2, Chapter 7.

	B	*A*
Mean	0.4667	0.2105
Variance	0.2667	0.1754
Observations	15	19
Hypothesized Mean Dif.	0	
df	27	
t Stat	1.5585	
P(T<=t) one-tail	0.0654	
t Critical one-tail	1.7033	
P(T<=t) two-tail	0.1308	
t Critical two-tail	2.0518	

Table C 2 - Means Comparison of 100% CM (A) vs. 90% CM, 10% FBS (C)

Means comparison of cell viability in isolated wells after 1 week, between plates A & C in Trial 2, Chapter 7.

	C	*A*
Mean	0.4286	0.2105
Variance	0.2571	0.1754
Observations	21	19
Hypothesized Mean Dif.	0	
df	38	
t Stat	1.4878	
P(T<=t) one-tail	0.0725	
t Critical one-tail	1.6860	
P(T<=t) two-tail	0.1451	
t Critical two-tail	2.0244	

Table C 3 - Means Comparison of 50% CM, 50% FM (D) vs. 49% CM, 49% FM, 2% FBS (E)

Means comparison of cell viability in isolated wells after 1 week, between plates D & E in Trial 2, Chapter 7.

	E	*D*
Mean	0.7222	0.3636
Variance	0.2124	0.2386
Observations	18	33
Hypothesized Mean Dif.	0	
df	37	
t Stat	2.5992	
P(T<=t) one-tail	0.0067	
t Critical one-tail	1.6871	
P(T<=t) two-tail	0.0133	
t Critical two-tail	2.0262	

Table C 4 - Means Comparison of 50% CM, 50% FM (D) vs. 45% CM, 45% FM, 10% FBS (F)

Means comparison of cell viability in isolated wells after 1 week, between plates D & F in Trial 2, Chapter 7.

	F	*D*
Mean	0.9000	0.3636
Variance	0.0947	0.2386
Observations	20	33
Hypothesized Mean Dif.	0	
df	51	
t Stat	4.9028	
P(T<=t) one-tail	0.000005	
t Critical one-tail	1.6753	
P(T<=t) two-tail	0.0000	
t Critical two-tail	2.0076	

Table C 5 - Means Comparison of 98% CM, 2% FBS (B) vs. 90% CM, 10% FBS (C)

Means comparison of cell viability in isolated wells after 1 week, between plates B & C in Trial 2, Chapter 7.

	C	*B*
Mean	0.4286	0.4667
Variance	0.2571	0.2667
Observations	21	15
Hypothesized Mean Dif.	0	
df	30	
t Stat	-0.2199	
P(T<=t) one-tail	0.4137	
t Critical one-tail	1.6973	
P(T<=t) two-tail	0.8275	
t Critical two-tail	2.0423	

Table C 6 - Means Comparison of 49% CM, 49% FM, 2% FBS (E) vs. 45% CM, 45% FM, 10% FBS (F)

Means comparison of cell viability in isolated wells after 1 week, between plates E & F in Trial 2.

	F	*E*
Mean	0.9000	0.7222
Variance	0.0947	0.2124
Observations	20	18
Hypothesized Mean Dif.	0	
df	29	
t Stat	1.3824	
P(T<=t) one-tail	0.0887	
t Critical one-tail	1.6991	
P(T<=t) two-tail	0.1774	
t Critical two-tail	2.0452	

Table C 7 - Means Comparison of 100% CM (A) vs. 50% CM, 50% FM (D)

Means comparison of cell viability in isolated wells after 1 week, between plates A & D in Trial 2, Chapter 7.

	D	*A*
Mean	0.3636	0.2105
Variance	0.2386	0.1754
Observations	33	19
Hypothesized Mean Dif.	0	
df	43	
t Stat	1.1932	
P(T<=t) one-tail	0.1197	
t Critical one-tail	1.6811	
P(T<=t) two-tail	0.2393	
t Critical two-tail	2.0167	

Table C 8 - Means Comparison of 98% CM, 2% FBS (B) vs. 49% CM, 49% FM, 2% FBS (E)

Means comparison of cell viability in isolated wells after 1 week, between plates B & E in Trial 2, Chapter 7.

	E	*B*
Mean	0.7222	0.4667
Variance	0.2124	0.2667
Observations	18	15
Hypothesized Mean Dif.	0	
df	28	
t Stat	1.4859	
P(T<=t) one-tail	0.0742	
t Critical one-tail	1.7011	
P(T<=t) two-tail	0.1485	
t Critical two-tail	2.0484	

Table C 9 - Means Comparison of 90% CM, 10% FBS (C) vs. 45% CM, 45% FM, 10% FBS (F)

Means comparison of cell viability in isolated wells after 1 week, between plates C & F in Trial 2, Chapter 7.

	F	*C*
Mean	0.9000	0.4286
Variance	0.0947	0.2571
Observations	20	21
Hypothesized Mean Dif.	0	
df	33	
t Stat	3.6176	
P(T<=t) one-tail	0.0005	
t Critical one-tail	1.6924	
P(T<=t) two-tail	0.0010	
t Critical two-tail	2.0345	

Table C 10 - Means Comparison of 100% CM (A+1) vs. 50% CM, 50% FM (D+3)

Means comparison of cell viability in isolated wells after 1 week. Comparison is between the aggregate data of plate A in Trial 2 & plate 1 in Trial 1 vs plate D in Trial 2 & plate 3 in Trial 1, Chapter 7.

	D+3	*A+1*
Mean	0.3542	0.0909
Variance	0.2336	0.0846
Observations	48	44
Hypothesized Mean Dif.	0	
df	78	
t Stat	3.1951	
P(T<=t) one-tail	0.0010	
t Critical one-tail	1.6646	
P(T<=t) two-tail	0.0020	
t Critical two-tail	1.9908	

Appendix D

Example Reactions

Example reactions for protocols listed in Chapter 4 are listed here.

Table D 1 - Example PCR Set-Up

Component	Value
5x HF Buffer	10 µl
10mM dNTPs	1 µl
Phusion Polymerase	0.5 µl
Primer 1	1.25 µl
Primer 2	1.25 µl
Template (1/20)	2 µl
dH_2O	34 µl
Length	1500 bp
Elongation Time	1.5 min
T_a	65°C

Table D 2 - Example SOE-PCR Set-Up

Component	Mix A	Mix B
5x HF Buffer	5 µl	5 µl
10mM dNTPs	0.5 µl	0.5 µl
Phusion Polymerase	0.25 µl	0.25 µl
Fragment 1	1 µl	0 µl
Fragment 2	2 µl	0 µl
Primer 1	0 µl	1.25 µl
Primer 2	0 µl	1.25 µl
dH_2O	16.25 µl	16.25 µl
Combined Length	1600bp	1600bp
Elongation Time	90 sec	90 sec
T_a	67°C	63°C

Table D 3 - Example Gibson Assembly Set-Up

Component	Volume
2x Gibson Assembly Master Mix	5 µl
Insert Fragment	2 µl
Backbone Fragment	0.5 µl
dH_2O	2.5 µl

Table D 4 - Example Restriction Digest Set-Up

Component	Volume
DNA	2 µl
FastDigest Green Buffer	2 µl
SdaI	1 µl
BamHI	2 µl
dH_2O	14 µl

Table D 5 - Example Ligation Set-Up

Component	Volume
10x T4 Buffer	2 µl
T4 Ligase	0.2 µl
Insert DNA	6 µl
Backbone DNA	2 µl
dH_2O	9.8 µl

Table D 6 - Example Agarose Gel Set-Up

Component	Volume
TAE Buffer	50 ml
Agarose Powder	0.8%-1.4% w/v
20,000x RedSafe Dye	2.5 μl

Appendix E

Pseudo-attP Search in *S. frugiperda*

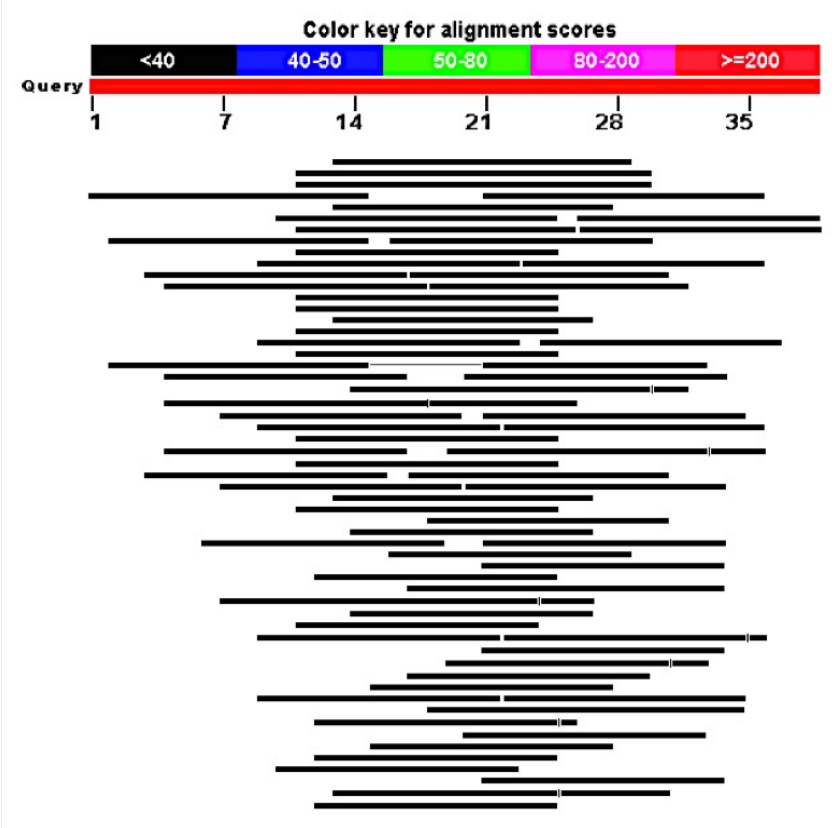

Figure E 1 - Search for Pseudo-attP Sites in S. Frugiperda

An overview of results from a basic local alignment search of ϕC31-attP against the *S. frugiperda* genome. No result showed higher that 41% identity. Genome data from (Kakumani, Malhotra, Mukherjee, & Bhatnagar, 2014)

Bibliography

Abbani, M. A., Papagiannis, C. V., Sam, M. D., Cascio, D., Johnson, R. C., & Clubb, R. T. (2007). Structure of the cooperative Xis-DNA complex reveals a micronucleoprotein filament that regulates phage lambda intasome assembly. *Proceedings of the National Academy of Sciences*, *104*(7), 2109–2114. http://doi.org/10.1073/pnas.0607820104

Abremski, K., & Hoess, R. (1984). Bacteriophage P1 Site-specific Recombination. *The Journal of Biological Chemistry*, *60*, 1509–1514.

Aihara, H., Kwon, H. J., Nunes-Düby, S. E., Landy, A., & Ellenberger, T. (2003). A conformational switch controls the DNA cleavage activity of lambda integrase. *Molecular Cell*, *12*(1), 187–98. http://doi.org/10.1016/S1097-2765(03)00268-5

Akopian, A., He, J., Boocock, M. R., & Stark, W. M. (2003). Chimeric recombinases with designed DNA sequence recognition. *Proceedings of the National Academy of Sciences*, *100*(15), 8688–8691.

Allen, M. D., Yamasaki, K., Ohme-Takagi, M., Tateno, M., & Suzuki, M. (1998). A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *The EMBO Journal*, *17*(18), 5484–96. http://doi.org/10.1093/emboj/17.18.5484

Andrews, B. J., Proteau, G. A., Beatty, L. G., & Sadowski, P. D. (1985). The FLP recombinase of the 2μ circle DNA of yeast: Interaction with its target sequences. *Cell*, *40*(4), 795–803. http://doi.org/10.1016/0092-8674(85)90339-3

Atkinson, A. E., Henderson, J., Hawes, C. R., & King, L. A. (1996). Efficient membrane targeting of the chick nicotinic acetylcholine receptor alpha-subunit in stable insect cell lines. *Cytotechnology*, *19*(1), 37–42. http://doi.org/10.1007/BF00749753

Aumiller, J. J., Mabashi-Asazuma, H., Hillar, A., Shi, X., & Jarvis, D. L. (2012). A new glycoengineered insect cell line with an inducibly mammalianized protein N-glycosylation pathway. *Glycobiology*, *22*(3), 417–428. http://doi.org/10.1093/glycob/cwr160

Baldwin, E. P., Martin, S. S., Abel, J., Gelato, K., Kim, H., Schultz, P. G., & Santoro, S. W.

(2003). A Specificity Switch in Selected Cre Recombinase Variants Is Mediated by Macromolecular Plasticity and Water. *Chemistry and Biology*, *10*, 1085–1094. http://doi.org/10.1016/j

Bateman, J. R., Lee, A. M., & Wu, C. T. (2006). Site-specific transformation of Drosophila via phi-C31 integrase-mediated cassette exchange. *Genetics*, *173*(2), 769–777. http://doi.org/10.1534/genetics.106.056945

Bernard, P., & Couturier, M. (1992). Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *Journal of Molecular Biology*, *226*(3), 735–745. http://doi.org/10.1016/0022-2836(92)90629-X

Bibb, L. A., & Hatfull, G. F. (2002). Integration and excision of the Mycobacterium tuberculosis prophage-like element , phiRv1. *Molecular Microbiology*, *45*(6), 1515–1526.

Biosciences, B. (2010). BD FACSCalibur Flow Cytometry System Fluidics. BD Biosciences.

Biswas, T., Aihara, H., Radman-livaja, M., Filman, D., Landy, A., & Ellenberger, T. (2005). A structural basis for allosteric control of DNA recombination by λ integrase, *435*(7045), 1059–1066.

Bleckmann, M., Fritz, M. H.-Y., Bhuju, S., Jarek, M., Schürig, M., Geffers, R., … van den Heuvel, J. (2015). Genomic Analysis and Isolation of RNA Polymerase II Dependent Promoters from Spodoptera frugiperda. *Plos One*, *10*(8), 1–16. http://doi.org/10.1371/journal.pone.0132898

Brown, W. R. A., Lee, N. C. O., Xu, Z., & Smith, M. C. M. (2011). Serine recombinases as tools for genome engineering. *Methods*, *53*(4), 372–379. http://doi.org/10.1016/j.ymeth.2010.12.031

Bryan, A., Roesch, P., Davis, L., Moritz, R., Pellett, S., & Welch, R. A. (2006). Regulation of Type 1 Fimbriae by Unlinked FimB- and FimE-Like Recombinases in Uropathogenic Escherichia coli Strain CFT073, *74*(2), 1072–1083. http://doi.org/10.1128/IAI.74.2.1072

Buchholz, F., Angrand, P. O., & Stewart, A. F. (1998). Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nature Biotechnology*, *16*(7), 657–662. http://doi.org/10.1038/nbt0798-657

Buchholz, F., Ringrose, L., Angrand, P. O., Rossi, F., & Stewart, A. F. (1996). Different thermostabilities of FLP and Cre recombinases: Implications for applied site-specific recombination. *Nucleic Acids Research*, *24*(21), 4256–4262. http://doi.org/10.1093/nar/24.21.4256

Buchholz, F., & Stewart, A. F. (2001). Alteration of Cre recombinase site specificity by substrate-linked protein evolution. *Nature Biotechnology*, *19*(11), 1047–1052. http://doi.org/10.1038/nbt1101-1047

Butler, M. (2005). Animal cell cultures: Recent achievements and perspectives in the production of biopharmaceuticals. *Applied Microbiology and Biotechnology*, *68*(3), 283–291. http://doi.org/10.1007/s00253-005-1980-8

Calles, K., Svensson, I., Lindskog, E., & Häggström, L. (2006). Effects of conditioned medium factors and passage number on Sf9 cell physiology and productivity. *Biotechnology Progress*, *22*(2), 394–400. http://doi.org/10.1021/bp050297a

Camerini-Otero, R. D., & Hsieh, P. (1995). Homologous recombination proteins in prokaryotes and eukaryotes. *Annual Review of Genetics*, *29*, 509–552. http://doi.org/10.1146/annurev.genet.29.1.509

Campbell, A. M. (1963). Episomes. *Advances in Genetics*, *11*, 101–145. http://doi.org/10.1016/S0065-2660(08)60286-2

Chan, K. M., Liu, Y. T., Ma, C. H., Jayaram, M., & Sau, S. (2013). The 2 micron plasmid of Saccharomyces cerevisiae: A miniaturized selfish genome with optimized functional competence. *Plasmid*, *70*(1), 2–13. http://doi.org/10.1016/j.plasmid.2013.03.001

Chen, X. L., Reindle, A., & Johnson, E. S. (2005). Misregulation of 2 microm circle copy number in a SUMO pathway mutant. *Molecular and Cellular Biology*, *25*(10), 4311–20. http://doi.org/10.1128/MCB.25.10.4311-4320.2005

Chen, Y., Narendra, U., Iype, L. E., Cox, M. M., Rice, P. A., & Gorman, O. (2000). Crystal Structure of a Flp Recombinase – Holliday Junction Complex : Assembly of an Active Oligomer by Helix Swapping, *6*, 885–897.

Chen, Y., & Rice, P. A. (2003a). New insight into the site-specific recombination from FLP recombinase-DNA structures. *Annu Rev Biophys. Biomol. Struct.*, *32*, 135–59.

http://doi.org/10.1146/annurev.biophys.32.110601.141732

Chen, Y., & Rice, P. A. (2003b). The Role of the Conserved Trp 330 in Flp-mediated Recombination. *The Journal of Biological Chemistry*, *278*(27), 24800–24807. http://doi.org/10.1074/jbc.M300853200

Cheo, D. L., Titus, S. A., Byrd, D. R. N., Hartley, J. L., Temple, G. F., & Brasch, M. A. (2004). Concerted assembly and cloning of multiple DNA fragments using in vitro site-specific recombination: Functional analysis of multi-site expression clones. *Genome Research*, *14*, 2111–2120. http://doi.org/10.1101/gr.2512204.)

Choi, S., Begum, D., Koshinsky, H., Ow, D., & Wing, R. (2000). A new approach for the identification and cloning of genes: the pBACwich system using Cre/lox site-specific recombination. *Nucleic Acids Research*, *28*(7), 1–7. http://doi.org/10.1093/nar/28.7.e19

Chompoosri, J., Fraser, T., Rongsriyam, Y., Komalamisra, N., Siriyasatien, P., Thavara, U., … Fraser, M. J. (2009). Intramolecular integration assay validates integrase PHI C31 and R4 potential in a variety of insect cells. *Southeast Asian Journal of Tropical Medicine and Public Health*, *40*(6), 1235–1253.

Colloms, S. D., Merrick, C. A., Olorunniji, F. J., Stark, W. M., Smith, M. C. M., Osbourn, A., … Rosser, S. J. (2014). Rapid metabolic pathway assembly and modification using serine integrase site-specific recombination. *Nucleic Acids Research*, *42*(4), e23–e23. http://doi.org/10.1093/nar/gkt1101

Dafhnis-Calas, F., Xu, Z., Haines, S., Malla, S. K., Smith, M. C. M., & Brown, W. R. A. (2005). Iterative in vivo assembly of large and complex transgenes by combining the activities of φC31 integrase and Cre recombinase. *Nucleic Acids Research*, *33*(22). http://doi.org/10.1093/nar/gni192

Dale, E. C., & Ow, D. W. (1991). Gene transfer with subsequent removal of the selection gene from the host genome. *Proceedings of the National Academy of Sciences*, *88*(23), 10558–10562. http://doi.org/10.1073/pnas.88.23.10558

Das, B., Martinez, E., Midonet, C., & Barre, F.-X. (2013). Integrative mobile elements exploiting Xer recombination. *Trends in Microbiology*, *21*(1), 23–30. http://doi.org/10.1016/j.tim.2012.10.003

Douris, V., Swevers, L., Labropoulou, V., Andronopoulou, E., Georgoussi, Z., & Iatrou, K. (2006). Stably Transformed Insect Cell Lines: Tools for Expression of Secreted and Membrane-anchored Proteins and High-throughput Screening Platforms for Drug and Insecticide Discovery. *Advances in Virus Research*, (68), 113–156. http://doi.org/10.1016/S0065-3527(06)68004-4

Doverskog, M., Bertram, E., Ljunggren, J., Ohman, L., Sennerstam, R., & Haggstrom, L. (2000). Cell cycle progression in serum-free cultures of Sf9 insect cells: Modulation by conditioned medium factors and implications for proliferation and productivity. *Biotechnology Progress*, *16*(5), 837–846. http://doi.org/10.1021/bp000108i

Elick, T. A., Bauser, C. A., & Fraser, M. J. (1996). Excision of the piggyBac transposable element in vitro is a precise event that is enhanced by the expresion of its encoded transposase. *Genetica*, *98*, 33–41.

Enders, J. F., Weller, T. H., & Robbins, F. . (1949). Cultivation of the Lansing Strain of Poliomyelitis Virus in Cultures of Various Human Embryonic Tissues. *Science*, *109*(2822), 85–87.

Ennifar, E., Meyer, J. E. W., Buchholz, F., Stewart, A. F., & Suck, D. (2003). Crystal structure of a wild-type Cre recombinase-IoxP synapse reveals a novel spacer conformation suggesting an alternative mechanism for DNA cleavage activation. *Nucleic Acids Research*, *31*(18), 5449–5460. http://doi.org/10.1093/nar/gkg732

Ennis, T. J., & Sohi, S. S. (1976). Chromosomal characterisation of five lepidopteran cell lines of Malacosoma disstria (Lasiocampidae) and Christoneura fumiferana (Tortricidae). *Canadian Journal of Genetics and Cytology*, *18*(3), 471–477. http://doi.org/10.1139/g76-057

Esposito, D., & Scocca, J. J. (1997). The integrase family of tyrosine recombinases : evolution of a conserved active site domain. *Nucleic Acids Research*, *25*(18), 3605–3614.

Fadeev, E. A., Sam, M. D., & Clubb, R. T. (2009). NMR Structure of the Amino-Terminal Domain of the Lambda Integrase Protein in Complex with DNA: Immobilization of a Flexible Tail Facilitates Beta-Sheet Recognition of the Major Groove. *Journal of*

Molecular Biology, *388*(4), 682–690. http://doi.org/10.1016/j.jmb.2009.03.041

Fath-Goodin, A., Kroemer, J. A., & Webb, B. A. (2009). The campoletis sonorensis ichnovirus vankyrin protein p-vank-1 inhibits apoptosis in insect sf9 cells. *Insect Molecular Biology*, *18*(4), 497–506. http://doi.org/10.1111/j.1365-2583.2009.00892.x

Fernandes, F., Dias, M. M., Vidigal, J., Sousa, M. F. Q., Patrone, M., Teixeira, A. P., & Alves, P. M. (2014). Production of rotavirus core-like particles in Sf9 cells using recombinase-mediated cassette exchange. *Journal of Biotechnology*, *171*, 34–8. http://doi.org/10.1016/j.jbiotec.2013.11.020

Fernandes, F., Vidigal, J., Dias, M. M., Prather, K. L. J., Coroadinha, A. S., Teixeira, A. P., & Alves, P. M. (2012). Flipase-mediated cassette exchange in Sf9 insect cells for stable gene expression. *Biotechnology and Bioengineering*, *109*(11), 2836–2844. http://doi.org/10.1002/bit.24542

Fogg, P. C. M., Colloms, S., Rosser, S., Stark, M., & Smith, M. C. M. (2014). New Applications for Phage Integrases. *Journal of Molecular Biology*, *426*(15), 2703–2716. http://doi.org/10.1016/j.jmb.2014.05.014

Fogg, P. C. M., Younger, E., Fernando, B. D., Khaleel, T., Stark, W. M., & Smith, M. C. M. (2018). Recombination directionality factor gp3 binds ΦC31 integrase via the zinc domain, potentially affecting the trajectory of the coiled-coil motif. *Nucleic Acids Research*, *46*(3), 1308–1320. http://doi.org/10.1093/nar/gkx1233

Freuler, F., Stettler, T., Meyerhofer, M., Leder, L., & Mayr, L. M. (2008). Development of a novel Gateway-based vector system for efficient, multiparallel protein expression in Escherichia coli. *Protein Expression and Purification*, *59*(2), 232–241. http://doi.org/10.1016/j.pep.2008.02.003

Futcher, A. B. (1986). Copy Number Amplification of the 2 um Circle Plasmid of Saccharomyces cerevisiae. *Journal of Theoretical Biology*, (119), 197–204.

Gardiner, G. R., & Stockdale, H. (1975). Two tissue culture media lepidopteran cells and nuclear for production polyhedrosis viruses. *Journal of Invertebrate Pathology*, *25*, 363–370.

Ghosh, K., Chi, K. L., Guo, F., Segall, A. M., & Van Duyne, G. D. (2005). Peptide trapping

of the holliday junction intermediate in Cre-loxP site-specific recombination. *Journal of Biological Chemistry*, *280*(9), 8290–8299. http://doi.org/10.1074/jbc.M411668200

Ghosh, K., Guo, F., & Van Duyne, G. D. (2007). Synapsis of loxP sites by cre recombinase. *Journal of Biological Chemistry*, *282*(33), 24004–24016. http://doi.org/10.1074/jbc.M703283200

Ghosh, K., Lau, C., Gupta, K., & Duyne, G. D. Van. (2005). Preferential synapsis of loxP sites drives ordered strand exchange in Cre- loxP site-specific recombination. *Nature Chemical Biology*, *1*(5), 275–283. http://doi.org/10.1038/nchembio733

Ghosh, P., Kim, A. I., & Hatfull, G. F. (2003). The Orientation of Mycobacteriophage Bxb1 Integration Is Solely Dependent on the Central Dinucleotide of attP and attB. *Molecular Cell*, *12*, 1101–1111.

Ghosh, P., Pannunzio, N. R., & Hatfull, G. F. (2005). Synapsis in Phage Bxb1 Integration : Selection Mechanism for the Correct Pair of Recombination Sites. *Journal of Molecular Biology*, *349*, 331–348. http://doi.org/10.1016/j.jmb.2005.03.043

Gibb, B., Gupta, K., Ghosh, K., Sharp, R., Chen, J., & Duyne, G. D. Van. (2010). Requirements for catalysis in the Cre recombinase active site. *Nucleic Acids Research*, *38*(17), 5817–5832. http://doi.org/10.1093/nar/gkq384

Giuraniuc, C. V., MacPherson, M., & Saka, Y. (2013). Gateway Vectors for Efficient Artificial Gene Assembly In Vitro and Expression in Yeast Saccharomyces cerevisiae. *PLoS ONE*, *8*(5). http://doi.org/10.1371/journal.pone.0064419

Greene, A. E., Charney, J., Nichols, W. W., & Coriell, L. L. (1972). Species Identity of Insect Cell Lines. *In Vitro*, *7*(5), 313–322.

Gregory, M. A., Till, R., & Smith, M. C. M. (2003). Integration Site for Streptomyces Phage phiBT1 and Development of Site-Specific Integrating Vectors. *Journal of Bacteriology*, *185*(17), 5320–5323. http://doi.org/10.1128/JB.185.17.5320

Grindley, N. D. F., Whiteson, K. L., & Rice, P. A. (2006). Mechanisms of Site-Specific Recombination. *Annual Review of Biochemistry*, *75*, 567–605. http://doi.org/10.1146/annurev.biochem.73.011303.073908

Gross, A., Schoendube, J., Zimmermann, S., Steeb, M., Zengerle, R., & Koltay, P. (2015).

Technologies for single-cell isolation. *International Journal of Molecular Sciences*, *16*(8), 16897–16919. http://doi.org/10.3390/ijms160816897

Groth, A. C., Fish, M., Nusse, R., & Calos, M. P. (2004). Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. *Genetics*, *166*(4), 1775–82. http://doi.org/10.1534/genetics.166.4.1775

Groth, A. C., Olivares, E. C., Thyagarajan, B., & Calos, M. P. (2000). A phage integrase directs efficient site-specific integration in human cells. *Proceedings of the National Academy of Sciences*, *97*(11), 5995–6000. http://doi.org/10.1073/pnas.090527097

Guo, F., Gopaul, D. N., & Duyne, G. D. Van. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature*, *389*(9), 40–46.

Gupta, M., Till, R., & Smith, M. C. M. (2007). Sequence in attB that affect the ability of ΦC31 integrase to synapse and to activate DNA cleavage. *Nucleic Acids Research*, *35*(10), 3407–3419. http://doi.org/10.1093/nar/gkm206

Hambly, E., & Suttle, C. A. (2005). The viriosphere, diversity, and genetic exchange within phage communities. *Current Opinion in Microbiology*, *8*(4), 444–450. http://doi.org/10.1016/j.mib.2005.06.005

Harrison, R. G., Greenman, M. J., Mall, F. P., & Jackson, C. M. (1907). Observations of the living developing nerve fiber. *The Anatomical Record*, *1*(5), 116–128. http://doi.org/10.1002/ar.1090010503

Harrison, R. L., & Jarvis, D. L. (2010). Transforming Lepidopteran Insect Cells for Continuous Recombinant Protein Expression. *Baculovirus and Insect Cell Expression Protocols*, *19*, 470. http://doi.org/10.1007/978-1-4939-3043-2

Hartley, J. L., Temple, G. F., & Brasch, M. A. (2000). DNA Cloning Using In Vitro Site-Specific Recombination. *Genome Research*, *10*(11), 1788–1795. http://doi.org/10.1101/gr.143000.that

Hatfull, G. F., & Grindley, N. D. F. (1986). Analysis of gamma-delta resolvase mutants in vitro : Evidence for an interaction between serine-10 of resolvase and site I of res. *Proceedings of the National Academy of Sciences*, *83*, 5429–5433.

Heffron, F., & Mccarthy, B. J. (1979). DNA Sequence Analysis of the Transposon Three

Genes and Three Sites Involved in Transposition of Tn3. *Cell*, *18*(12), 1153–1163.

Heichman, K. A., Moskowitz, I. P. G., & Johnson, R. C. (1991). Configuration of DNA strands and mechanism of strand exchange in the Hin invertasome as revealed by analysis of recombinant knots. *Genes and Development*, *5*, 1622–1634.

Heil, J. R., Cheng, J., & Charles, T. C. (2012). Site-specific Bacterial Chromosome Engineering: ΦC31 Integrase Mediated Cassette Exchange (IMCE). *Journal of Visualized Experiments*, (61), 1–6. http://doi.org/10.3791/3698

Hink, W. F. (1970). Established insect cell line from the cabbage looper, Trichoplusia ni. *Nature*, *226*(5244), 466–467. http://doi.org/10.1038/226466b0

Hoess, R., Abremski, K., & Sternberg, N. (1984). Interaction of the bacteriophage P1 recombinase Cre with the recombining site loxP. *Proceedings of the National Academy of Sciences*, *81*(February), 1026–1029. http://doi.org/10.1101/SQB.1984.049.01.086

Hollister, J. R., Shaper, J. H., & Jarvis, D. L. (1998). Stable expression of mammalian β 1 , 4-galactosyltransferase extends the N-glycosylation pathway in insect cells. *Glycobiology*, *8*(5), 473–480.

Houdt, R. Van, Leplae, R., Lima-mendez, G., Mergeay, M., & Toussaint, A. (2012). Towards a more accurate annotation of tyrosine- based site-specific recombinases in bacterial genomes. *Mobile DNA*, *3*(6), 1–11.

Invitrogen. (2002). *Growth and Maintenance of Insect Cell lines*. *Life Technologies Corporation*.

Jarvis, D. L., Fleming, J. A., Kovacs, G. R., Summers, M. D., & Guarino, L. A. (1990). Use of Early Baculovirus Promoters for Continuous Expression and Efficient Processing of Foreign Gene Products in Stably Transformed Lepidopteran Cells. *Bio/Technology*, *8*, 950–955.

Jayaram, M. (1997). The cis-trans paradox of integrase. *Science*, *276*, 49–51.

Jehle, J. A., Blissard, G. W., Bonning, B. C., Cory, J. S., Herniou, E. A., Rohrmann, G. F., … Vlak, J. M. (2006). On the classification and nomenclature of baculoviruses: A proposal for revision. *Archives of Virology*, *151*(7), 1257–1266. http://doi.org/10.1007/s00705-006-0763-6

Kakumani, P. K., Malhotra, P., Mukherjee, S. K., & Bhatnagar, R. K. (2014). A draft genome assembly of the army worm, Spodoptera frugiperda. *Genomics*, *104*(2), 134–143. http://doi.org/10.1016/j.ygeno.2014.06.005

Kallmeyer, J., Pockalny, R., Adhikari, R. R., Smith, D. C., & D'Hondt, S. (2012). Global distribution of microbial abundance and biomass in subseafloor sediment. *Proceedings of the National Academy of Sciences*, *109*(40), 16213–16216. http://doi.org/10.1073/pnas.1203849109

Karasawa, S., Araki, T., Yamamoto-Hino, M., & Miyawaki, A. (2003). A Green-emitting Fluorescent Protein from Galaxeidae Coral and Its Monomeric Version for Use in Fluorescent Labeling. *Journal of Biological Chemistry*, *278*(36), 34167–34171. http://doi.org/10.1074/jbc.M304063200

Karpinski, J., Hauber, I., Chemnitz, J., Schäfer, C., Paszkowski-Rogacz, M., Chakraborty, D., … Buchholz, F. (2016). Directed evolution of a recombinase that excises the provirus of most HIV-1 primary isolates with high specificity. *Nature Biotechnology*. http://doi.org/10.1038/nbt.3467

Kempf, J., Snook, L. a, Vonesch, J. L., Dahms, T. E. S., Pattus, F., & Massotte, D. (2002). Expression of the human mu opioid receptor in a stable Sf9 cell line. *Journal of Biotechnology*, *95*(2), 181–7.

Kersulyte, D., Mukhopadhyay, A. K., Shirai, M., Nakazawa, T., & Berg, D. E. (2000). Functional Organization and Insertion Specificity of IS 607 , a Chimeric Element of Helicobacter pylori. *Journal of Bacteriology*, *182*(19), 5300–5308.

Khaleel, T., Younger, E., Mcewan, A. R., Varghese, A. S., & Smith, M. C. M. (2011). A phage protein that binds φC31 integrase to switch its directionality. *Molecular Microbiology*, *80*(6), 1450–1463. http://doi.org/10.1111/j.1365-2958.2011.07696.x

Kim, A. I., Ghosh, P., Aaron, M. A., Bibb, L. A., Jain, S., & Hatfull, G. F. (2003). Mycobacteriophage Bxb1 integrates into the Mycobacterium smegmatis groEL1 gene. *Molecular Microbiology*, *50*(2), 463–473. http://doi.org/10.1046/j.1365-2958.2003.03723.x

Kitts, P. A., Ayres, M. D., & Possee, R. D. (1990). Linearization of baculovirus DNA

enhances the recovery of recombinant virus expression vectors. *Nucleic Acids Research*, *18*(19), 5667–5672. http://doi.org/10.1093/nar/18.19.5667

Kleymann, G., Boege, F., Hahn, M., Hampe, W., Vasudevan, S., & Reilander, H. (1993). Human β2-adrenergic receptor produced in stably transformed insect cells is functionally coupled via endogenous GTP-binding protein to adenylyl cyclase. *European Journal of Biochemistry*, *213*(2), 797–804. http://doi.org/10.1111/j.1432-1033.1993.tb17822.x

Konieczka, J. H., Paek, A., Jayaram, M., & Voziyanov, Y. (2004). Recombination of hybrid target sites by binary combinations of Flp variants: Mutations that foster interprotomer collaboration and enlarge substrate tolerance. *Journal of Molecular Biology*, *339*(2), 365–378. http://doi.org/10.1016/j.jmb.2004.03.060

Kost, T. A., Condreay, J. P., & Jarvis, D. L. (2005). Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nature Biotechnology*, *23*(5), 567–575. http://doi.org/10.1038/nbt1095.Baculovirus

Kostrewa, D., Granzin, J., Koch, C., Choe, H.-W., Raghunathan, S., Wolf, W., … Saenger, W. (1991). Three-dimensional structure of the E. coli DNA-binding protein FIS. *Nature*, *349*(1), 178–180.

Kostriken, R., Morita, C., & Heffron, F. (1981). Transposon Tn3 encodes a site-specific recombination system: Identification of essential sequences, genes, and actual site of recombination. *Proceedings of the National Academy of Sciences*, *78*(7), 4041–4045.

Kranz, A., Fu, J., Duerschke, K., Weidlich, S., Naumann, R., Francis Stewart, A., & Anastassiadis, K. (2010). An improved flp deleter mouse in C57Bl/6 based on flpo recombinase. *Genesis*, *48*(8), 512–520. http://doi.org/10.1002/dvg.20641

Kretzmer, G. (2002). Industrial processes with animal cells. *Applied Microbiology and Biotechnology*, *59*, 135–142. http://doi.org/10.1007/s00253-002-0991-y

Krogh, B. O., & Shuman, S. (2000). Catalytic Mechanism of DNA Topoisomerase IB. *Molecular Cell*, *5*, 1035–1041.

Kuhstoss, S., & Rao, N. (1991). Analysis of the Integration Function of the Streptomycete Bacteriophage phiC31. *Journal of Molecular Biology*, *222*, 897–908.

Kunkel, B., Losick, R., & Stragier, P. (1990). The Bacillus subtilis gene for the developmental transcription factor is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. *Genes and Development*, *4*, 525–535.

Kwon, H. J., Tirumalai, R., Landy, A., & Ellenberger, T. (1997). Flexibility in DNA Recombination : Structure of the Lambda Integrase Catalytic Core. *Science*, *276*(5309), 126–131.

Landy, A. (1989). Dynamic, Sructural, and Regulatory Aspects of Lambda Site-Specific Recombination. *Annual Review of Biochemistry*, *58*, 913–949.

Lauth, M. (2002). Stable and efficient cassette exchange under non-selectable conditions by combined use of two site-specific recombinases. *Nucleic Acids Research*, *30*(21), 115e–115. http://doi.org/10.1093/nar/gnf114

Lee, J., & Jayaram, M. (1995). Role of partner homology in DNA recombination. Complementary base pairing orients the 5'-hydroxyl for strand joining during Flp site-specific recombination. *Journal of Biological Chemistry*, *270*, 4042–4052.

Lee, J., Tribble, G., & Jayaram, M. (2000). Resolution of tethered antiparallel and parallel holliday junctions by the Flp site-specific recombinase. *Journal of Molecular Biology*, *296*(2), 403–419. http://doi.org/10.1006/jmbi.1999.3472

Li, H., Wang, J., Wilhelmsson, H., Hansson, A., Thoren, P., Duffy, J., … Larsson, N. G. (2000). Genetic modification of survival in tissue-specific knockout mice with mitochondrial cardiomyopathy. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(7), 3467–72. http://doi.org/10.1073/pnas.97.7.3467

Li, S., Flisikowska, T., Kurome, M., Zakhartchenko, V., Kessler, B., Saur, D., … Schnieke, A. (2014). Dual fluorescent reporter pig for cre recombination: Transgene placement at the ROSA26 locus. *PLoS ONE*, *9*(7), 3–10. http://doi.org/10.1371/journal.pone.0102455

Li, W., Kamtekar, S., Xiong, Y., & Sarkis, G. J. (2005). Structure of a Synaptic gd Resolvase Tetramer Covalently Linked to Two Cleaved DNAs. *Science*, *309*(8), 1210–1216.

Li, X., Harrell, R. A., Handler, A. M., Beam, T., Hennessy, K., & Fraser, M. J. (2005). piggyBac internal sequences are necessary for efficient transformation of target

genomes. *Insect Molecular Biology*, *14*(1), 17–30. http://doi.org/10.1111/j.1365-2583.2004.00525.x

Lin, C. H., & Jarvis, D. L. (2013). Utility of temporally distinct baculovirus promoters for constitutive and baculovirus-inducible transgene expression in transformed insect cells. *Journal of Biotechnology*, *165*(1), 11–17. http://doi.org/10.1016/j.jbiotec.2013.02.007

Lin, C. Y., Wang, Y. H., Li, K. C., Sung, L. Y., Yeh, C. L., Lin, K. J., … Hu, Y. C. (2015). Healing of massive segmental femoral bone defects in minipigs by allogenic ASCs engineered with FLPo/Frt-based baculovirus vectors. *Biomaterials*, *50*(1), 98–106. http://doi.org/10.1016/j.biomaterials.2015.01.052

Loessner, M. J., Inman, R. B., & Lauer, P. (2000). Complete nucleotide sequence , molecular analysis and genome structure of bacteriophage A118 of Listeria monocytogenes : implications for phage evolution. *Molecular Mic*, *35*(2), 324–340.

Lomovskaya, N. D., Chater, K. F., & Mkrtumian, N. M. (1980). Genetics and molecular biology of Streptomyces bacteriophages. *Microbiological Reviews*, *44*(2), 206–29.

Long, D., Lu, W., & Hao, Z. (2016). Highly efficient and inducible DNA excision in transgenic silkworms using the FLP / FRT site-specific recombination system. *Transgenic Research*, *25*(6), 795–811. http://doi.org/10.1007/s11248-016-9970-4

Long, D., Zhao, A., Xu, L., Lu, W., Guo, Q., Zhang, Y., & Xiang, Z. (2013). In vivo site-specific integration of transgene in silkworm via PhiC31 integrase-mediated cassette exchange. *Insect Biochemistry and Molecular Biology*, *43*(11), 997–1008. http://doi.org/10.1016/j.ibmb.2013.08.001

Loomis, K. H., Yaeger, K. W., Batenjany, M. M., Mehler, M. M., Grabski, A. C., Wong, S. C., & Novy, R. E. (2005). InsectDirect™ System: Rapid, high-level protein expression and purification from insect cells. *Journal of Structural and Functional Genomics*, *6*(2–3), 189–194. http://doi.org/10.1007/s10969-005-5241-y

López, M. G., Alfonso, V., Carrillo, E., & Taboga, O. (2010). Trans-complementation of polyhedrin by a stably transformed Sf9 insect cell line allows occ- baculovirus occlusion and larval per os infectivity. *Journal of Biotechnology*, *145*(2), 199–205. http://doi.org/10.1016/j.jbiotec.2009.10.015

Louwerse, J. D., van Lier, M. C. M., van der Steen, D. M., de Vlaam, C. M. T., Hooykaas, P. J. J., & Vergunst, A. C. (2007). Stable recombinase-mediated cassette exchange in Arabidopsis using Agrobacterium tumefaciens. *Plant Physiology*, *145*(4), 1282–1293. http://doi.org/10.1104/pp.107.108092

Lyznik, L. A., Rao, K. V., & Hodges, T. K. (1996). FLP-mediated recombination of FRT sites in the maize genome. *Nucleic Acids Research*, *24*(19), 3784–3789. http://doi.org/10.1093/nar/24.19.3784

Ma, C., Kwiatek, A., Bolusani, S., Voziyanov, Y., & Jayaram, M. (2007). Unveiling Hidden Catalytic Contributions of the Conserved His/Trp-III in Tyrosine Recombinases: Assembly of a Novel Active Site in Flp Recombinase Harboring Alanine at this Position. *Journal of Molecular Biology*, *368*, 183–196. http://doi.org/10.1016/j.jmb.2007.02.022

Malanowska, K., Salyers, A. A., & Gardner, J. F. (2006). Characterization of a conjugative transposon integrase IntDOT. *Molecular Microbiology*, *60*(5), 1228–1240. http://doi.org/10.1111/j.1365-2958.2006.05164.x

Martin, S. S., Pulido, E., Chu, V. C., Lechner, T. S., & Baldwin, E. P. (2002). The order of strand exchanges in Cre-LoxP recombination and its basis suggested by the crystal structure of a Cre-LoxP holliday junction complex. *Journal of Molecular Biology*, *319*(1), 107–127. http://doi.org/10.1016/S0022-2836(02)00246-2

McCarroll, L., & King, L. a. (1997). Stable insect cell cultures for recombinant protein production. *Curr Opin Biotechnol*, *8*(5), 590–594. http://doi.org/S0958-1669(97)80034-1 [pii]

McEwan, A. R., Raab, A., Kelly, S. M., Feldmann, J., & Smith, M. C. M. (2011). Zinc is essential for high-affinity DNA binding and recombinase activity of ρc31 integrase. *Nucleic Acids Research*, *39*(14), 6137–6147. http://doi.org/10.1093/nar/gkr220

Mcewan, A. R., Rowley, P. A., & Smith, M. C. M. (2009). DNA binding and synapsis by the large C-terminal domain of φC31 integrase. *Nucleic Acids Research*, *37*(14), 4764–4773. http://doi.org/10.1093/nar/gkp485

Misiura, A., Pigli, Y. Z., Boyle-vavra, S., Daum, R. S., Boocock, M. R., & Rice, P. A.

(2013). Roles of two large serine recombinases in mobilizing the methicillin-resistance cassette SCCmec. *Molecular Microbiology*, *88*(May), 1218–1229. http://doi.org/10.1111/mmi.12253

Mullany, P., Roberts, A. P., & Wang, H. (2002). Mechanism of integration and excision in conjugative transposons. *Cell and Molecular Life Sciences*, *59*, 2017–2022.

Mumm, J. P., Landy, A., & Gelles, J. (2006). Viewing single lambda site-specific recombination events from start to finish. *The EMBO Journal*, *25*(19), 4586–4595. http://doi.org/7601325 [pii] 10.1038/sj.emboj.7601325

Nanto, K., & Ebinuma, H. (2008). Marker-free site-specific integration plants. *Transgenic Research*, *17*(3), 337–344. http://doi.org/10.1007/s11248-007-9106-y

Nash, H. A. (1974). Purification of bacteriophage λ Int protein. *Nature*, *247*, 543–545.

Nash, H. a. (1996). Site-Specific Recombination: Integration, Excision, Resolution, and Inversion of Defined DNA Segments. *Escherichia Coli and Salmonella: Cellular and Molecular Biology*, *2*, 2363–2376.

Nimmo, D. D., Alphey, L., Meredith, J. M., & Eggleston, P. (2006). High efficiency site-specific genetic engineering of the mosquito genome. *Insect Molecular Biology*, *15*(2), 129–136. http://doi.org/10.1111/j.1365-2583.2006.00615.x

Nollmann, M., He, J., Byron, O., & Stark, W. M. (2004). Solution Structure of the Tn3 Resolvase-Crossover Site Synaptic Complex. *Molecular Cell*, *16*, 127–137.

Nunes-Düby, S. E., Azaro, M. A., & Landy, A. (1995). Swapping DNA strands and sensing homology without branch migration in lambda site-specific recombination. *Current Biology*, *5*(2).

Nunes-Düby, S. E., Kwon, H. J., Tirumalai, R. S., Ellenberger, T., & Landy, A. (1998). Similarities and differences among 105 members of the Int family of site-specific recombinases. *Nucleic Acids Research*, *26*(2), 391–406. http://doi.org/10.1093/nar/26.2.391

Nunes-Düby, S. E., Smith-Mungo, L. I., & Landy, A. (1995). Single base-pair precision and structural rigidity in a small IHF-induced DNA loop. *Journal of Molecular Biology*, *253*(2), 228–242. http://doi.org/10.1006/jmbi.1995.0548

Osuna, R., Lienau, D., & Hughes, K. T. (1995). Sequence, Regulation, and Functions of fis in Salmonella typhimurium. *Journal of Bacteriology*, *177*(8), 2021–2032.

Papagiannis, C. V., Sam, M. D., Abbani, M. A., Yoo, D., Cascio, D., Clubb, R. T., & Johnson, R. C. (2007). Fis Targets Assembly of the Xis Nucleoprotein Filament to Promote Excisive Recombination by Phage Lambda. *Journal of Molecular Biology*, *367*(2), 328–343. http://doi.org/10.1016/j.jmb.2006.12.071

Pfeifer, T. a., Hegedus, D. D., Grigliatti, T. a., & Theilmann, D. a. (1997). Baculovirus immediate-early promoter-mediated expression of the Zeocin resistance gene for use as a dominant selectable marker in Dipteran and Lepidopteran insect cell lines. *Gene*, *188*(2), 183–190. http://doi.org/10.1016/S0378-1119(96)00756-1

Pijlman, G. P., van Schinjndel, J. E., & Vlak, J. M. (2003). Spontaneous excision of BAC vector sequences from bacmid-derived baculovirus expression vectors upon passage in insect cells. *Journal of General Virology*, *84*(10), 2669–2678. http://doi.org/10.1099/vir.0.19438-0

Plasterk, R. H. A., Brinkman, A. D., & Putte, P. V. A. N. D. E. (1983). DNA inversions in the chromosome of Escherichia coli and in bacteriophage Mu: Relationship to other site-specific recombination systems. *Proceedings of the National Academy of Sciences*, *80*(9), 5355–5358.

Poulter, R. T. M., & Goodwin, T. J. D. (2005). DIRS-1 and the other tyrosine recombinase retrotransposons. *Cytogenetic and Genome Research*, *110*(1–4), 575–588. http://doi.org/10.1159/000084991

Ptashne. (2004). *A genetic switch: phage lambda revisted* (3rd ed.). Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Ravin, N. V. (2011). Plasmid N15 : The linear phage – plasmid. *Plasmid*, *65*, 102–109. http://doi.org/10.1016/j.plasmid.2010.12.004

Raymond, C. S., & Soriano, P. (2007). High-efficiency FLP and ΦC31 site-specific recombination in mammalian cells. *PLoS ONE*, *2*(1), 1–4. http://doi.org/10.1371/journal.pone.0000162

Reed, R. R. (1981). Resolution of cointegrates between transposons yv and Tn3 defines the

recombination site. *Proceedings of the National Academy of Sciences*, *78*(6), 3428–3432.

Reed, R. R., & Grindley, N. D. F. (1981). Transposon-Mediated Recombination in Vitro : DNA Cleavage and Protein-DNA Linkage at the Recombination Site. *Cell*, *25*, 721–728.

Reed, R. R., & Moser, C. D. (1984). Resolvase-mediated recombination intermediates contain a serine residue covalently linked to DNA. *Cold Spring Harbor Symposia on Quantitative Biology*, *49*, 245–249.

Ren, L., Pang, D., Zhang, M., Li, L., Wang, T., Ouyang, H., & Li, A. (2011). Comparative analysis of the activity of two promoters in insect cells. *African Journal of Biotechnology*, *10*(44), 8930–8941. http://doi.org/10.5897/AJB11.1044

Rice, P. A., & Steitz, T. A. (1994). Refinement of gamma-delta resolvase reveals a strikingly flexible molecule. *Structure*, *2*(5), 371–384.

Rice, P. A., Yang, S. W., Mizuuchi, K., & Nash, H. A. (1996). Crystal structure of an IHF-DNA complex: A protein-induced DNA U-turn. *Cell*, *87*(7), 1295–1306. http://doi.org/10.1016/S0092-8674(00)81824-3

Rodrigues, M. E., Costa, A. R., Henriques, M., Azeredo, J., & Oliveira, R. (2014). *Animal Cell Biotechnology - Methods and Protocols*. *Animal Cell Biotechnology: Methods and Protocols* (Vol. 1104). http://doi.org/10.1007/978-1-62703-733-4_10

Rojek, A., Füchtbauer, E.-M., Kwon, T.-H., Frøkiaer, J., & Nielsen, S. (2006). Severe urinary concentrating defect in renal collecting duct-selective AQP2 conditional-knockout mice. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(15), 6037–42. http://doi.org/10.1073/pnas.0511324103

Rongrong, L., Lixia, W., & Zhongping, L. (2005). Effect of deletion mutation on the recombination activity of Cre recombinase. *Acta Biochimica Polonica*, *52*(2), 541–544.

Rutherford, K., & Duyne, G. D. Van. (2014). The ins and outs of serine integrase site-specific recombination. *Current Opinion in Structural Biology*, *24*, 125–131. http://doi.org/10.1016/j.sbi.2014.01.003

Rutherford, K., Yuan, P., Perry, K., Sharp, R., & Duyne, G. D. Van. (2013). Attachment site

recognition and regulation of directionality by the serine integrases. *Nucleic Acids Research*, *41*(17), 8341–8356. http://doi.org/10.1093/nar/gkt580

Ryan, V. T., Grimwade, J. E., Camara, J. E., Crooke, E., & Leonard, A. C. (2004). Escherichla coli prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA. *Molecular Microbiology*, *51*(5), 1347–1359. http://doi.org/10.1046/j.1365-2958.2003.03906.x

Sanderson, M. R., Freernon, P. S., Rice, P. A., Goldman, A., Hatfull, G. F., Grindley, N. D. F., & Steitz, T. A. (1990). The Crystal Structure of the Catalytic Domain of the Site-Specific tecombination Enzyme ~ 6 Resolvase at 2 . 7 A Resolution. *Cell*, *63*(12), 1323–1329.

Santoro, S. W., & Schultz, P. G. (2002). Directed evolution of the site specificity of Cre recombinase. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(9), 4185–4190. http://doi.org/10.1073/pnas.022039799

Sarkar, D., Azaro, M. A., Aihara, H., Papagiannis, C. V., Tirumalai, R., Nunes-Düby, S. E., … Landy, A. (2002). Differential affinity and cooperativity functions of the amino-terminal 70 residues of λ integrase. *Journal of Molecular Biology*, *324*(4), 775–789. http://doi.org/10.1016/S0022-2836(02)01199-3

Sarkar, D., Radman-Livaja, M., & Landy, A. (2001). The small DNA binding domain of Lambda integrase is a context-sensitive modulator of recombinase functions. *EMBO Journal*, *20*(5), 1203–1212. http://doi.org/10.1093/emboj/20.5.1203

Scientific, T. (2007). *260/280 and 260/230 Ratios*. *NanoDrop Products* (Vol. 17).

Seah, N. E., Warren, D., Tong, W., Laxmikanthan, G., Van Duyne, G. D., & Landy, A. (2014). Nucleoprotein architectures regulating the directionality of viral integration and excision. *Proceedings of the National Academy of Sciences*, *111*(34), 12372–12377. http://doi.org/10.1073/pnas.1413019111

Shepherd, M. D., Kharel, M. K., Bosserman, M. A., & Rohr, J. (2010). Laboratory maintenance of Streptomyces species. *Current Protocols in Microbiology*, 1–10. http://doi.org/10.1002/9780471729259.mc10e01s18.Laboratory

Shi, X., Harrison, R. L., Hollister, J. R., Mohammed, A., Fraser, M. J., & Jarvis, D. L.

(2007). Construction and characterization of new piggyBac vectors for constitutive or inducible expression of heterologous gene pairs and the identification of a previously unrecognized activator sequence in piggyBac. *BMC Biotechnology*, *7*, 5. http://doi.org/10.1186/1472-6750-7-5

Shirai, M., Nara, H., Sato, A., Aida, T., & Takahashi, H. (1991). Site-Specific Integration of the Actinophage R4 Genome into the Chromosome of Streptomyces parvulus upon Lysogenization. *Journal of Bacteriology*, *173*(13), 4237–4239.

Shultz, J. L., Voziyanova, E., Konieczka, J. H., & Voziyanov, Y. (2011). A genome-wide analysis of FRT-like sequences in the human genome. *PLoS ONE*, *6*(3). http://doi.org/10.1371/journal.pone.0018077

Silverman, M., & Simon, M. (1980). Phase Variation: Mutants Genetic Analysis of Switching. *Cell*, *19*(4), 845–854.

Sinclair, R. B., & Bibb, M. J. (1988). The repressor gene (c) of the Streptomyces temperate phage φ{symbol}c31: Nucleotide sequence, analysis and functional cloning. *MGG Molecular & General Genetics*, *213*(2–3), 269–277. http://doi.org/10.1007/BF00339591

Singh, S., Ghosh, P., & Hatfull, G. F. (2013). Attachment Site Selection and Identity in Bxb1 Serine Integrase-Mediated Site-Specific Recombination. *PLoS Genetics*, *9*(5). http://doi.org/10.1371/journal.pgen.1003490

Smith, G. E., Summers, M. D., & Fraser, M. J. (1983). Production of human beta interferon in insect cells infected with a baculovirus expression vector . Production of Human Beta Interferon in Insect Cells Infected with a Baculovirus Expression Vector. *Molecular and Cellular Biology*, *3*(12), 2156–2165. http://doi.org/10.1128/MCB.3.12.2156.Updated

Smith, M. C. A., Till, R., & Smith, M. C. M. (2004). Switching the polarity of a bacteriophage integration system. *Molecular Microbiology*, *51*(6), 1719–1728. http://doi.org/10.1111/j.1365-2958.2003.03942.x

Smith, M. C. M., & Thorpe, H. M. (2002). Diversity in the serine recombinases. *Molecular Microbiology*, *44*, 299–307.

Srivastava, V., & Ow, D. W. (2001). Biolistic mediated site-specific integration in rice. *Molecular Breeding*, *8*(4), 345–350. http://doi.org/10.1023/A:1015229015022

Stanasila, L., Pattus, F., & Massotte, D. (1998). Heterologous expression of G-protein-coupled receptors: Human opioid receptors under scrutiny. *Biochimie*, *80*, 563–571.

Stark, W. M., Grindley, N. D. F., Hatfull, G. F., & Boocock, M. R. (1991). Resolvase-catalysed reactions between res sites differing in the central dinucleotide of subsite I. *The EMBO Journal*, *10*(11), 3541–3548.

Stark, W. M., Sherratt, D. J., Soocock, M. R., & Gil, G. (1989). Site-Specific Recombination by Tn3 Resolvase : Topological Changes in the Forward and Reverse Reactions, *58*, 779–790.

Staszewski, R. (1984). Cloning by limiting dilution: An improved estimate that an interesting culture is monoclonal. *Yale Journal of Biology and Medicine*, *57*(6), 865–868.

Steele, K. H., Stone, B. J., Franklin, K. M., Fath-Goodin, A., Zhang, X., Jiang, H., … Geisler, C. (2017). Improving the baculovirus expression vector system with vankyrin-enhanced technology. *Biotechnology Progress*, *33*(6), 1496–1507. http://doi.org/10.1002/btpr.2516

Summers, M. D., & Smith, G. E. (1987). *A Manual of Methods for Baculovirus Vectors and Insect Cell CUlture Procedures*. *Department of Entomology Texas A&M University*.

Suttle, C. A. (2007). Marine viruses — major players in the global ecosystem. *Nature Reviews Microbiology*, *5*(10), 801–812. http://doi.org/10.1038/nrmicro1750

Sze, C. C., Laurie, A. D., & Shingler, V. (2001). In Vivo and In Vitro Effects of Integration Host Factor at the DmpR-Regulated Sigma54 -Dependent Po Promoter. *Journal of Bacteriology*, *183*(9), 2842–2851. http://doi.org/10.1128/JB.183.9.2842

Tekle, M., Warren, D. J., Biswas, T., Ellenberger, T., Landy, A., & Nunes-Düby, S. E. (2002). Attenuating functions of the C terminus of λ integrase. *Journal of Molecular Biology*, *324*(4), 649–665. http://doi.org/10.1016/S0022-2836(02)01108-7

Theilmann, D. A., & Stewart, S. (1992). Molecular analysis of the trans-activating IE-2 gene of Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus. *Virology*, *187*(1), 84–96. http://doi.org/10.1016/0042-6822(92)90297-3

Thompson, J., Vargas, L. M. De, Koch, C., Kahmann, R., & Landy, A. (1987). Cellular factors couple recombination with growth phase: Characterization of a new component

in the λ site-specific recombination pathway. *Cell*, *50*, 901–908.

Thorpe, H. M., & Smith, M. C. M. (1998). In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proceedings of the National Academy of Sciences*, *95*(10), 5505–5510. http://doi.org/10.1073/pnas.95.10.5505

Thyagarajan, B., Olivares, E. C., Hollis, R. P., Ginsburg, D. S., & Calos, M. P. (2001). Site-Specific Genomic Integration in Mammalian Cells Mediated by Phage C31 Integrase. *Molecular and Cellular Biology*, *21*(12), 3926–3934. http://doi.org/10.1128/MCB.21.12.3926-3934.2001

Tirumalai, R. S., Healey, E., & Landy, A. (1997). The catalytic domain of lambda site-specific recombinase. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(12), 6104–9. http://doi.org/10.1073/pnas.94.12.6104

van Oers, M. M., Malarme, D., Jore, J. M. P., & Vlak, J. M. (1992). Expression of the Autographa californica nuclear polyhedrosis virus p10 gene: effect of polyhedrin gene expression. *Archives of Virology*, *123*(1–2), 1–11. http://doi.org/10.1007/BF01317134

van Oers, M. M., Pijlman, G. P., & Vlak, J. M. (2015). Thirty years of baculovirus-insect cell protein expression: from dark horse to mainstream technology. *Journal of General Virology*, *96*(Pt_1), 6–23. http://doi.org/10.1099/vir.0.067108-0

Vaughn, J. L., Goodwin, R. H., Tompkins, G. J., & Mccawley, P. (1977). The establishment of two cell lines from the insect Spodoptera frugiperda (Lepidoptera; Noctuidae). *In Vitro*, *13*(4), 213–217.

Vidigal, J., Fernandes, F., Coroadinha, A. S., Teixeira, A. P., & Alves, P. M. (2014). Insect Cell Line Development Using Flp-Mediated Cassette Exchange Technology. *Animal Cell Biotechnology*, *1104*(2), 137–147. http://doi.org/10.1007/978-1-62703-733-4

Volkert, F. C., & Broach, J. R. (1986). Site-specific recombination promotes plasmid amplification in yeast. *Cell*, *46*(4), 541–550. http://doi.org/10.1016/0092-8674(86)90879-2

Voziyanov, Y., Konieczka, J. H., Stewart, A. F., & Jayaram, M. (2003). Stepwise manipulation of DNA specificity in Flp recombinase: Progressively adapting Flp to

individual and combinatorial mutations in its target site. *Journal of Molecular Biology*, *326*(1), 65–76. http://doi.org/10.1016/S0022-2836(02)01364-5

Voziyanov, Y., Stewart, A. F., & Jayaram, M. (2002). A dual reporter screening system identifies the amino acid at position 82 in Flp site-specific recombinase as a determinant for target specificity. *Nucleic Acids Research*, *30*(7), 1656–63.

Wang, B., Kitney, R. I., Joly, N., & Buck, M. (2011). Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nature Communications*, *2*, 508. http://doi.org/10.1038/ncomms1516

Wang, J., Ho, B., & Ding, J. L. (2001). Functional expression of full length Limulus Factor C in stably transformed Sf9 cells. *Biotechnology Letters*, *23*(1), 71–76. http://doi.org/10.1023/A:1026764514737

Warren, D., Laxmikanthan, G., & Landy, A. (2008). A chimeric Cre recombinase with regulated directionality. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(47), 18278–83. http://doi.org/10.1073/pnas.0809949105

Warrens, A. N., Jones, M. D., & Lechler, R. I. (1997). Splicing by over-lap extension by PCR using asymmetric amplification: An improved technique for the generation of hybrid proteins of immunological interest. *Gene*, *186*(1), 29–35. http://doi.org/10.1016/S0378-1119(96)00674-9

Wasserman, S. A., Dungan, J. M., & Cozzarelli, N. R. (1985). Discovery of a Predicted DNA Knot Substantiates a Model for Site-Specific Recombination. *Science*, *229*(4709), 171–174.

Whiteson, K. L., Chen, Y., Chopra, N., Raymond, A. C., & Rice, P. A. (2007). Brief Communication Identification of a Potential General Acid/Base in the Reversible Phosphoryl Transfer Reactions Catalyzed by Tyrosine Recombinases: Flp H305. *Chemistry and Biology*, *14*(2), 121–129. http://doi.org/10.1016/j.chembiol.2007.01.011

Wojciak, J. M., Sarkar, D., Landy, A., & Clubb, R. T. (2002). Arm-site binding by lambda - integrase: solution structure and functional characterization of its amino-terminal domain. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(6), 3434–9. http://doi.org/10.1073/pnas.052017999

Yang, S. H., & Jayaram, M. (1994). Generality of the Shared Active Site among Yeast Family Site-specific Recombinases. *The Journal of Biological Chemistry*, *269*(17), 12789–12796.

Yang, W., & Steitz, T. A. (1995). Crystal Structure of the Site-Specific Recombinase Resolvase Complexed with a 34 bp Cleavage Site. *Cell*, *82*, 193–207.

Yonemura, N., Tamura, T., Uchino, K., Kobayashi, I., Tatematsu, K. I., Iizuka, T., … Kusakabe, T. (2013). phiC31-integrase-mediated, site-specific integration of transgenes in the silkworm, Bombyx mori (Lepidoptera: Bombycidae). *Applied Entomology and Zoology*, *48*(3), 265–273. http://doi.org/10.1007/s13355-013-0182-6