1

Screening of Bacterial Recombinants: Strategies and Preventing False Positives

Sriram Padmanabhan, Sampali Banerjee and Naganath Mandi
Lupin Limited, Biotechnology, R & D, Ghotawade Village, Mulshi Taluka, India

1. Introduction

Complete decoding of complex eukaryotic genomes is a prerequisite for understanding varied gene functions. Gene silencing (point mutations, gene deletions, etc), sub cellular localization of proteins, gene expression pattern analysis by promoter activity assay, structure-function analysis, and *in vitro* or *in vivo* biochemical assays (Hartley et al., 2000; Curtis & Grossniklaus, 2003; Earley et al., 2006) are some of the approaches followed for understanding gene functions.

Typically, all the above approaches require the cloning of target genes with or without selective mutations, or cloning their promoter fragments into specialized vectors for further characterization. While the traditional approach for constructing expression cassettes that is based on the restriction enzyme/ligase cloning method is laborious and time consuming, the process is often hampered by length of the gene of interest, GC content of the gene, toxicity of the gene product to the expressing host and lack of relevant restriction sites for cloning purposes. All these factors render the production of expression constructs a significant technical obstacle for large-scale functional gene analysis.

After generating successful cloning/expression constructs, several steps followed are screening high number of colonies, avoiding false positive recombinants and requirement of dephosphorylation of vectors in case of single site cloning to ensure the generation of recombinants with rightly oriented gene of interest and to minimize vector background (non-recombinants).

Screening for recombinants is one of the most crucial and time-consuming steps in molecular cloning and several approaches available for this purpose include colony PCR screening, blue white screening, screening of recombinants, which have the gene of interest in the MCS region of the cloning vehicle, in such a way that the toxic gene reading frame is interrupted making the toxic gene inactivated upon insertion of any foreign gene; GFP fluorescence vectors wherein upon cloning, the GFP fluorescence disappears, etc. The method for screening of bacterial transformants that carry recombinant plasmid with the gene of interest, has become more rapid and simple by the use of vectors with visually detectable reporter genes.

2. Molecular cloning

A recombinant DNA comprises of two entities namely a vector and the gene of interest (GOI). The process of joining vector and any GOI is by making a phosphodiester bond by a

process called ligation. The ligation reaction is facilitated with the help of T4 DNA ligase in the presence of ATP. If a vector and any target DNA fragments are generated by the action of the same restriction endonuclease, they will join by base-pairing due to the compatibility of their respective ends. Such a construct is then transformed into a prokaryotic cell, where unlimited copies of the construct, an essentially the target DNA sequence is made inside the cell.

2.1 Steps in molecular cloning

The conventional restriction and ligation cloning protocol involves four major steps namely fragmentation of DNA with restriction endonucleases, ligation of DNA fragments to a plasmid vector, introduction into bacterial cells by transformation and screening and selection of recombinants.

2.1.1 Selection and preparation of vector and insert

A cloning vehicle, also termed as a vector, can be classified as a carrier carrying a gene to be transferred from one organism to another. Other cloning vectors include plasmids, cosmids, bacteriophage, phagemids and artificial chromosomes. In the early days of producing proteins in *E. coli*, limitations to transcription initiation were believed to lead to lower protein expression levels (Gralla, 1990). This event resulted in efforts put into construction of expression vectors, which carried strong promoters to enhance mRNA yield and a stable mRNA eventually. The promoters used included phage promoters like T7 and T5, the synthetic promoters tac and trc, and the arabinose inducible araBAD (Trepe, 2006). Additional vectors that were made available included Lambda promoters, PR and PL, (Elvin et al., 1990), rhamnose promoter (Cardona & Valvano, 2005), Trp-lac promoter (Chernajovskyi et al., 1983) etc. Certain promoter variants as seen in the expression vector pAES25 yield the maximum level of soluble active target protein (Broedel & Papciak , 2007).

Downstream of each specific promoter, there is a multiple cloning site (MCS) for cloning the gene to be expressed. While the inducible promoters are used to drive the foreign gene expression, the constitutive promoters (Liang et al,., 1999) are used mainly to express the antibiotic expression marker genes for plasmid maintenance.

TA cloning vectors (Zhou & Gomez-Sanchez, 2000; Chen et al., 2009) that takes advantage of the well-known propensity of non-proofreading DNA polymerases (e.g., Taq, Tfl, Tth) to add a single 3´-A to PCR products are also employed for cloning large PCR fragments. The proof-reading polymerases lack 5'-3' proofreading activity and are capable of adding adenosine triphosphate residues to the 3' ends of the double stranded PCR product. Such a PCR amplified product can then be cloned in any linearized vector with complementary 3' T overhangs.

The GC cloning technology is based on the recent discovery that the above proof-reading enzymes similarly add a single 3´-G to DNA molecules, either during PCR or as a separate G-tailing reaction to any blunt DNA. GC cloning vectors pSMART® GC and pGC™ Blue (commercialized by Lucigen, USA) contain a single 3´-C overhang, which is compatible with the single 3´-G overhang on the inserts.

Mead and coworkers (Mead et al., 1991) report cloning of PCR products without any restriction digestion taking advantage of the single 3' deoxyadenylate extension that

Thermus aquaticus, Thermus flavus, and Thermococcus litoralis DNA polymerases add to the termini of amplified nucleic acid.

Gateway cloning system is a relatively new trend in the field of molecular cloning, where in the site specific recombination system of lambda phage is used (Katzen 2007). This system enables the researchers to efficiently transfer DNA fragments between different vector and expression systems, without changing the orientation of the gene or its reading frame. The specific sequences are called "Gateway att sites" and recombination is facilitated by two enzymes "LR clonase" and "BP clonase". This easy Ligase-free cloning system is very beneficial for cloning, combining and transferring of DNA segments between different expression platforms in a high-throughput manner, but making the gateway entry clone usually involves conventional restriction enzyme based cloning, and this is a major drawback of this system.

DNA vectors that are used in many molecular biology gene cloning experiments need not necessarily result in protein expression. Expression vectors are often specifically designed to contain regulatory sequences that act as enhancer and promoter regions, and lead to efficient transcription of the gene that is carried on the expression vector. The regularly used cloning cum expression vectors include pET vectors, pBAD vectors, pTrc vectors etc wherein the GOI is cloned with a suitable promoter of the vector using the start codon of the vector or using a gene of interest with its own start codon into an apopropriate restriction site in the MCS.

RNA polymerases are enzyme complexes that synthesize RNA molecules using DNA as a template. The transcription begins when RNA polymerase binds to the DNA double helix which is at a promoter site just upstream of the gene to be transcribed. While in prokaryotes, one DNA-dependent RNA polymerase transcribes all classes of DNA molecules and the core *Escherichia coli* enzyme called *E. coli* RNA polymerase consists of three types of subunit, α, β, and β′, and has the composition $\alpha_2\beta\beta'$; the holoenzyme contains an additional σ subunit or sigma factor (Aaron, 2001). The phage RNA polymerase like T7 RNA polymerase found in pET based expression vectors are much smaller and simpler than bacterial ones: the polymerases from phage T3 and T7 RNA, e.g., are single polypeptide chains of <100 kDa.

The DNA fragment to be cloned is first isolated by a number of ways like cDNA preparation, nuclease fragments of genomic DNA, synthetic DNA's, amplified DNA fragments by means of polymerase chain reaction. After appropriate restriction enzyme digestion and purification, the purified inserts are ligated to the vector of choice.

2.1.2 Ligation of vector and insert

The ligation step is carried out with bacteriophage T4 DNA ligase using ATP required for the reaction and a suitable buffer condition. This process involves the joining of two DNA molecule ends with a phosphodiester bond between the 3' hydroxyl of one nucleotide and the 5' phosphate of the other. The ligation event is of two types sticky or blunt based on the types of restriction enzyme used for digestion of the vector and insert.

2.1.3 Transformation

Following ligation, the ligation product (recombinant plasmid) is transformed into bacteria for propagation. The transformed bacteria are then plated on selective agar to select for bacteria that have the plasmid of interest. Individual colonies are picked up and tested for

the desired insert. The transformation is achieved by chemical method (Hanahan, 1983; Inoyue, 1997; Bergmans, 1981) or electroporation (Morrison, 2001).

2.1.3.1 Chemical transformation

For transformation of bacterial cells by chemical means, the cells are grown to mid-log phase, harvested and treated with divalent cations such as $CaCl_2$ to make them competent. After mixing DNA with such competent cells, on ice, followed by a brief heat shock at 42 ^{0}C, the cells are incubated with rich medium for 30-60 minutes prior to plating on suitable antibiotic containing LB agar plates. The biggest advantage of this method includes no special equipment for transformation with no requirement to remove salts from the DNA used for transformation.

2.1.3.2 Electroporation

For electroporation, cells are also grown to mid-log phase but are then washed extensively with water to eliminate all salts from the growth medium, and glycerol added to the water to a final concentration of 10% so that the cells can be stored frozen and saved for future experiments. To electroporate DNA into cells, washed *E. coli* are mixed with the DNA to be transformed and then pipetted into a plastic cuvette containing electrodes. A short electric pulse, about 2400 volts/cm, is applied to the cells causing small pores in the membrane through which the DNA enters. The cells are then incubated with broth as above before plating. For electroporation, the DNA must be free of all salts so the ligations are first precipitated with alcohol before they are used.

2.2 Types of E. coli host cells used for transformation

For most cloning applications, *E. coli* k12 hosts like DH5α which are OmpT protease expressing cells (Salunkhe etal., 2010) are used. These cells are compatible with *lacZ* blue/white selection procedures, are easily transformed, and good quality plasmid DNA can be recovered from transformants. DH5α is one of the most preferred strains for plasmid propagation, because it is an *EndA1* knockout which inactivates the endonucleases ,and a recA knock out which prevents rapid homologous recombination, hence ensuring that the plasmids are stable inside the cells. One notable exception is when transforming with plasmid constructs containing recombinant genes under control of the T7 polymerase, these constructs are typically transformed into DH5α cells during the cloning stage and later introduced into a bacterial strain expressing T7 RNA polymerase for expression of the recombinant protein. The derivatives available for this purpose include BL21(DE3), BL21A1 which are all lon and OmpT protease negative strains (Banerjee et al., 2009; Mandi et al, 2008) and ER2566 (Yu et al., 2004) strains.

3. Selection of recombinants

The need to identify the cells that contain the desired insert at the appropriate and right orientation and isolate these from those not successfully transformed is of utmost importance to researchers. Modern cloning vectors include selectable markers (most frequently antibiotic resistance markers) that allow only cells in which the vector, but not necessarily the insert, has been transformed to grow. Additionally, the cloning vectors may contain color selection markers which provide blue/white screening (via α-factor complementation) on X-gal medium. Nevertheless, these selection steps do not

absolutely guarantee that the DNA insert is present in the cells. Further investigation of the resulting colonies is required to confirm that cloning was successful. This may be accomplished by means of PCR, restriction fragment analysis and/or DNA sequencing.

3.1 Colony immunoassay

Telford et al., (1977) reported identification of bacterial plasmids carrying DNA upon lysis and mixing with molten agar with ethidium bromide stain while in as early as 1978, a colony hybridization was developed for screening recombinants by Cami and Kourilsky, (1978) where upon blotting onto nitrocellulose filters and hybridization with a highly radioactive probe, screening of many thousands of colonies per plate for the presence of a DNA sequence carried by a plasmid and complementary to the probe was achieved.

An immunological approach to screen recombinant clones is possible if the gene of interest encodes a polypeptide for which specific antibodies can be prepared. In one approach, DNA complementary to mRNA is inserted in frame with the coding regions of genes present in *E. coli* plasmids. These results in "fused polypeptides" consisting of the *N*-terminal region of an *E. coli* polypeptide covalently linked to a sequence encoded by the cloned cDNA segment. The identification of cloned genes by colony immunoassays has not been common and one limitation of all previous colony immunoassays is that each fused polypeptide molecule must simultaneously bind to two different antibody molecules. Typically, the first antibody, immobilized on a solid support such as chemically activated paper, is used to entrap the fused polypeptide at the site of the lysed colony, and a second labelled antibody is then bound to the fused polypeptide and detected by autoradiography. A potential disadvantage of all immunological methods is that only one in six sequences inserted at random into the vector would have the orientation and frame consistent with translation into a recognizable fused polypeptide. Kemp & Cowman (1981) have described a method by which fused polypeptides can be detected by a colony immunoassay that demands binding of only one antibody molecule. *E. coli* colonies containing recombinant plasmids are lysed in situ, and proteins in the lysate are immobilized by binding directly to CNBr-activated paper. Antigens attached to the paper are then allowed to react with antiserum, and the antibodies that bind to them are in turn detected by reaction with ^{125}I-labeled protein A (^{125}I-protein A) from *Staphylococcus aureus*, followed by autoradiography. The limitations of this method are mRNA instability, inefficient translation, or rapid proteolytic degradation of the fused polypeptides that restrict their accumulation within the cells.

A simple immunoassay has been developed by Reggie and Comeron (1986) for isolation of particular gene(s) from a clone bank of recombinant plasmids. A clone bank of the DNA is constructed with a plasmid vector in *Escherichia coli* and filtered onto a hydrophobic grid membrane and grown up into individual colonies, and a replica was made onto nitrocellulose paper. The bacterial cells upon lysis are immobilized onto the nitrocellulose paper which is reacted with a rabbit antibody preparation made against the particular antigenic product to detect the recombinant clone which carries the corresponding gene. The bound antibodies can be detected easily by a colorimetric assay using goat anti-rabbit antibodies conjugated to horseradish peroxidase.

3.2 Visual screening

3.2.1 The blue-white screening

The blue white screening is one of the most common molecular techniques that allow detecting the successful ligation of gene of interest in vector (Langley et al., 1975; Zamenhof & Villarejo 1972; Ausubel et al., 1988). The α-Complementation plasmids are among the most commonly used vectors for cloning and sequencing the DNA fragments, as they generally have a good multiple cloning site and an efficient blue-white screening system for identification of recombinants in presence of a histochemical dye, 5-bromo-4-chloro-3-indolyl-β-d-galactoside (X-gal), and binding sites for commercially available primers for direct sequencing of cloned fragments (Manjula 2004).

The molecular mechanism for blue/white screening is based on a genetic engineering of the *lac* operon in the *E. coli* as a host cell combined with a subunit complementation achieved with the cloning vector. The *lacZ* product, a polypeptide of 1029 amino acids, gives rise to the functional enzyme after tetramerization (Jacobson et al., 1994) and is easily detected by chromogenic substrates either in cell lysates or directly on fixed cells *in situ* (Ko et al., 1990). The tetramerization is dependent on the presence of the *N*-terminal region spanning the first 50 residues. Deletions in the *N*-terminal sequence generate a so-called omega peptide that is unable to tetramerize and does not display enzymatic activity. The activity of the omega peptide can be fully restored either in bacteria or *in vitro,* if a small fragment (called alpha peptide) corresponding to the intact N-terminal portion is added *in trans* (Gallagher et al., 1994). The phenomenon is called α-complementation and the small *N*-terminal peptide is called alpha peptide. This effect has been widely exploited for studies in prokaryotes, where special strains that constitutively express omega peptide exist and allow the detection of expression of the small alpha peptide.

The vector (e.g. pBluescript) encodes the α subunit of LacZ protein with an internal multiple cloning site (MCS), while the chromosome of the host strain encodes the remaining omega subunit to form a functional β-galactosidase enzyme upon complementation. The foreign DNA can be inserted within the MCS of *lacZ*α gene, thus disrupting the formation of functional β-galactosidase. The chemical required for this screen is X-gal, a colorless modified galactose sugar that is metabolized by β-galactosidase to form 5-bromo-4-chloro-indoxyl which is spontaneously oxidized to the bright blue insoluble pigment 5,5'-dibromo-4,4'-dichloro-indigo and thus functions as an indicator. Isopropyl β-D-1-thiogalactopyranoside (IPTG) which functions as the inducer of the Lac operon, can be used to enhance the phenotype. The hydrolysis of colorless X-gal by the β-galactosidase causes the characteristic blue colour in the colonies indicating that the colonies contain vector without insert. White colonies indicate insertion of foreign DNA and loss of the cells ability to hydrolyze the marker. Bacterial colonies in general, however, are white, and so a bacterial colony with no vector will also appear white. These are usually suppressed by the presence of an antibiotic in the growth medium. Blue white screening is thus a quick and easy technique that allows for the screening of successful cloning reactions through the color of the bacterial colony. However, the correct type of vector and competent cells are important considerations when planning a blue white screen.

Although the *lacZ* and many other systems have been extensively used for gram negative bacteria like *E. coli*, there are limited options available for screening recombinants transformed in Gram positive bacteria. Chaffin & Rubens (1998) have developed a gram positive cloning vector pJS3, that utilizes the interruption of an alkaline phosphatase gene, *phoZ*, to identify

recombinant plasmids. A multiple cloning site (MCS) was inserted distal to the region coding for the putative signal peptide of phoZ where the alkaline phosphatase protein expressed from the *phoZ* gene (*phoZMCS*) retained activity similar to that of the native protein and cells displayed a blue colonial phenotype on agar containing 5-bromo-4-chloro-3-indolyl phosphate (X-p). Introduction of any foreign DNA into the MCS of phoZ produced a white colonial phenotype on agar containing X-p and allowed discrimination between transformants containing recombinant plasmids versus those maintaining self-annealed or uncut vector. This cloning vector has improved the efficiency of recombinant DNA experiments in gram-positive bacteria.

Cloning inserts into the multiple cloning region of the pGEMÂ®-Z Vectors disrupts the alpha-peptide coding sequences, and thus inactivates the beta-galactosidase enzyme resulting in white colonies. Recombinant plasmids are transformed into the appropriate strain of bacteria (i.e. JM109, DH5α), and subsequently plated on indicator plates containing 0.5 mM IPTG and 40 µg/ml X-gal.

A new version of TA cloning vector with directional enrichment and blue-white color screening has been reported by Horn (2005).

3.2.2 Limitations of blue-white screening

The "blue screen" technique described above suffers from the disadvantage of using a screening procedure (discrimination) rather than a procedure for selecting the clones. Discrimination is based on visually identifying the recombinant within the population of clones on the basis of a color. The *LacZ* gene, in the vector used for generating recombinants, may be non-functional and may not produce β-galactosidase. As a result, these cells cannot convert X-gal to the blue substance so the white colonies seen on the plate may not be recombinants but just the background vector.

A few white colonies might not contain the desired recombinant but a small piece of DNA to be ligated into the vector's MCS might change the reading frame for *LacZa*, and thus prevent its expression giving rise to false positive clones. Furthermore, a few linearized vectors may get transformed into the bacteria, the ends "repaired" and ligated together such that no LacZα is produced as a result, these cells cannot convert X-gal to the blue substance. On the other hand, in some cases, blue colonies may contain the insert, when the insert is "in frame" with the *LacZa* gene and is devoid of stop codon. This could sometimes lead to the expression of a fusion protein that is still functional as LacZα. Small inserts which happen to be in frame with the alpha-peptide coding region may produce light blue colonies, as beta-galactosidase activity is only partially inactivated.

Last but not the least, this complex procedure requires the use of the substrate X-gal which is very expensive, unstable and is cumbersome to use.

3.3 Reporter gene based screening

Another method for screening and identification of recombinant clones is by using the green fluorescent protein (GFP) obtained from jellyfish *Aequorea victoria*. It is a reporter molecule for monitoring gene expression, protein localization, protein-protein interaction etc. GFP has been expressed in bacteria, yeast, slime mold, plants, drosophila, zebrafish and in mammalian cells. Inouye et al., (1997) have described a bacterial cloning vector with mutated *Aequorea* GFP protein as an indicator for screening recombinant plasmids. The pGREENscript A when expressed in *E. coli* produced colonies showing yellow color in day

light and strong green fluorescence under long-UV. Inserted foreign genes are selected on the basis of loss of the fluorescence caused by inactivation of the GFP production. The vector used in the study is a derivative of pBluescript SK(+) (Short et al., 1998) and encodes for the same MCS flanked by T3 and T7 promoters, but lacks the *lacZ* gene and the f1 origin region (for single strand DNA production). Instead, the GFP-S65A cDNA is substituted in its place and is under the control of the *lac* promoter/operator. The insertion of foreign DNAs into MCS of pGreenscript A interferes with the production of GFPS65A and causes a loss in the green fluorescence and yellow color of *E. coli* colonies. While GFP solubility appears to be one of the limiting factors in whole cell fluorescence, Davis & Vierstra (1998) have reported about soluble derivatives of GFP for use in *Arabidopsis thaliana.*

A system for direct screening of recombinant clones in *Lactococcus lactis*, based on secretion of the staphylococcal nuclease (SNase) in the organism, was developed by Loir and co-workers (Loir et al, 1994). *L. lactis* strains containing the nuc' plasmids secrete SNase and are readily detectable by a simple plate test. An MCS was introduced just after the cleavage site between leader peptide and the mature SNase, without affecting the nuclease activity. Cloning foreign DNA fragments into any site of the MCS interrupts nuc gene and thus results in nuc mutant clones which are easily distinguished from nuc' clones on plates. The biggest advantage of this vector is the possibility of assessing activity of the fusion protein since the nuclease activity is not diminished by its N-terminal tail and is also reported to be unaffected by denaturing agents such as sodium dodecyl sulfate (SDS) or trichloroacetic acid.

3.3.1 Limitations of the reporter gene based screening

All the above described plasmids could also result in false positive clones, which is a major concern for researchers. Loss of GFP fluorescence due to medium composition is also known to lead to false positive results. Although the SNase based screening would give absolute 100% recombinants, the active nuc fusion protein expression might render the cell fragile and enhance its susceptible to the lethal action of the fusion protein upon hyper expression. Also, for all the above cases, there is a requirement of transfer the genes of interest from the cloning vector to the expression vectors which calls for fresh cloning followed by screening for recombinants.

Hence, it is evident that the commonly used method for screening and identification of recombinant clones are associated with problems of false positive results forcing researchers to look for alternate methods of screening bacterial recombinants and availability of vectors that would act both as cloning vectors and expression vector are user friendly and advantageous.

3.4 Reporter gene based screening- new concepts

The recent approach of screening recombinants is the use of vector for one-step screening and expression of foreign genes (Banerjee et al., 2010) (Fig. 1). The strategy uses the cloning of GFP gene into any expression vector with a stop codon other than the amber stop codon upstream of the ORF of GFP. Upon induction, the GFP would not express and hence would not fluoresce due to presence of the stop codon. Upon in frame cloning of any gene of interest upstream of GFP in such a vector, would then excise the initial stop codon and the resultant fusion protein would fluoresce. The gene of interest contains an amber stop codon and the recombinant screening is carried out in an amber suppressor *E. coli* strain. For expression studies, the same clone can be used for expression in a non amber suppressor *E. coli* host like LE392 cells.

This report describes a vector where screening the transformants *in situ* for the presence of recombinants is possible without any false positive results. All other commercially available vectors show loss of color or loss of fluorescence that may not be unfailing while the major advantage of the vector described in this report, takes the property of color or fluorescence obtained after cloning. This unique vector would also be applicable with any other reporter genes like beta galactosidase gene, luciferase gene, DsRed protein instead of the described gene for GFP in the same vector constructed similarly. It also provides researchers to skip setting up the control ligation mix (without insert) and the dephosphorylation step (CIP or SAP step) since the religated vector would never glow and all the fluorescing colonies are indicative of only the recombinants and also indicative of correct reading frame of the inserted target gene. Since GFP fluorescence is brightest when it is expressed in soluble form (Davis & Vierstra 1998), the intensity of the fluorescence after cloning any foreign gene would also indicate the extent of solubility of the fusion protein. The major advantage of the vector described in this report, takes the property of color or fluorescence obtained after cloning, which is hitherto unreported. Such a vector used for identification, selection and expression of recombinants has been patented by the authors (Deshpande etal., 2010) and is published under PCT (Patent no. WO2010/0226601A2)

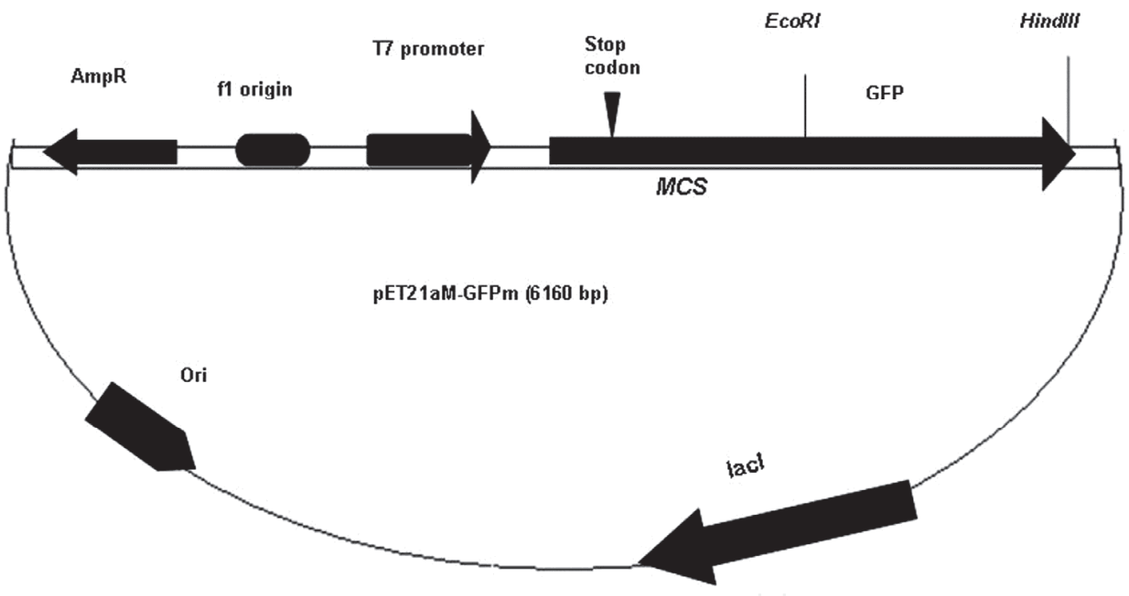


Fig. 1. Plasmid map of the vector pET21aM-GFPm that carries the reporter gene GFP at the 3′ end of the MCS. Amp^R refers to ampicillin resistant marker gene.

3.5 Screening clones by positive selection

A variety of positive selection cloning vectors has been developed that allow growth of only those bacterial colonies that carry recombinant plasmids. Typically, these plasmids express a gene product that is lethal for certain bacterial hosts and insertion of any DNA fragment that insertionally inactivates the expression of the toxic gene product resulting in growth of colonies.

Positive selection has been a powerful method of screening insert containing transformants. Here the toxic property of the molecule to the host cells is utilized for recombinant selection. The DNA sequence coding for the toxic product is directly cloned under the promoter elements recognized by the host cells. Positive selection in these vectors is achieved by either inactivation or replacement of toxic gene by the target gene. In general former is much more convenient than the latter (For a detailed review on positive selection vectors ; see reference Choi et al., 2002). The advantage of these systems is that no background colonies (non recombinants) appear on vector alone plates since the religated vectors carrying toxic intact genes are lethal to the host cells.

The genes of toxin-antitoxin system (Finbarr, 2003) of *E. coli* are being utilized for the development of positive selection vectors. A cloning vector carrying cloned *ccdB* gene which encodes poisonous topoisomerase II (DNA gyrase) causing unrecoverable DNA damage has been developed and is widely used as zero background cloning kit from Invitrogen (Bernard 1995). The other toxins from this system which are used successfully are *parD* and *parE* toxins (Gabant et al., 2000; Kim et al., 2004). Apart from this system, the other toxic gene used for the development of positive selection vectors are Colicin E3 and mutated form (E181Q) of catabolite activator protein (Ohashi-Kunihoro et al., 2006; Schlieper et al., 1998). These powerful selection strategies, however, are often only suitable for cloning and require a special host strain for propagation which carry a gene encoding antidote to the product of toxic gene.

The authors of this chapter have developed a positive selection vector which would be used for cloning and expression of heterologous genes simultaneously (Mandi et al., 2009). Here the toxic gene is derived from the antisense strand of *ccdB* gene and cloned under tightly regulated araBAD promoter downstream to the multiple cloning sites (Fig. 2). Multiple cloning sites facilitate cloning of foreign genes and doesn't affect the lethality of toxic gene. A simple method is used to screen the recombinants in that the transformed cells are plated on LB agar medium containing 13 mM arabinose. Moreover, this vector is used also for the expression of genes with authentic *N*-terminus and does not require special host strain for its propagation. Recently, Haag & Ostermeier (2009) have also reported the development of a novel positive selection vector, RHP-Amp^s, that is suitable for cloning and high level protein expression in *E. coli*. Although some limitations exist, positive selection vectors are useful in recombinant DNA experiments thereby reducing the time, effort and cost spent on identifying the correct clones.

Some of other well documented lethal genes include bacteriophage λ repressor, EcoRI methylase, EcoRI endonuclease, galactokinase, colicin E3, transcription factor GATA-1, lysis protein of φX174, barnase (Sambrook et al, 1989), SacB protein of *Bacillus subtilis* (Gay et al., 1985), RpsL protein of *E. coli* (Hashimoto-Gotoh et al., 1993), and also the ParD system of the R1 plasmid (Gabant et al., 2000).

A recently described strategy by Manjula, (2004) involves the use of galactose sensitivity exhibited by galactose epimerase (galE) mutants of *Escherichia coli*. Here, the *E. coli* cells that are lacZ+ galE, but not lacZ− galE are killed upon addition of lactose due to the accumulation of a toxic intermediate, UDP-galactose, by hydrolysis of lactose. Such a method has been suggested to be useful during primary cloning experiments such as construction of genomic or cDNA libraries and also in instances involving selection for rare recombinants.

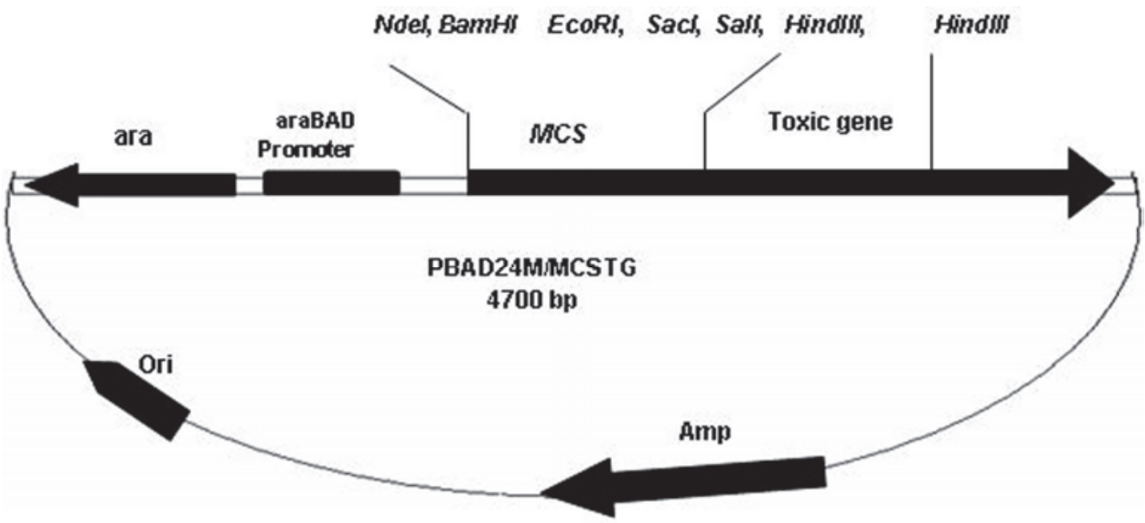


Fig. 2. Plasmid map of pBAD24MMCSTG which has a toxic gene under arabinose promoter as a HindIII/HindIII fragment and having a MCS for cloning any foreign gene to interrupt the expression of the TG gene to select recombinants based on cell survival in presence of the added inducer.

3.5.1 Potential limitations of screening recombinants based on positive selection

Requirements of tightly controlled promoter for expression of the anti-dote of the lethal gene. Requirements of special host cells for the lethal gene inserted or integrated in the bacterial genome is also one of the potential drawbacks and renders this system with limited applications. Moreover, since different genes respond to different promoters, requiring different kinds of host RNA polymerases, the modification of the host with the required lethal gene becomes a prerequisite with various genes which involves cost and is time-consuming. However, for efficient positive selection, the lethality of the marker gene must be strong enough to completely kill the clones harboring vector without insert.

4. Prediction of solubility of recombinant clones during screening

The structural and functional genomics require large supply of soluble, pure and functional proteins for high throughput analysis and as far as screening of soluble or insoluble recombinants is concerned, soluble protein production in *E. coli* is still a major bottleneck for investigators and a couple of efforts have been reported to improve the solubility or folding of recombinant protein produced in *E. coli* (Smith 2007). These include strategies like co-expression of chaperone proteins such as GroES, GroEL, DnaK and DnaJ lowering incubation temperature, use of weak promoters, addition of sucrose and betaine in a growth media, use of a richer media with phosphate buffer such as terrific broth (TB), use of signal sequence to export the protein to the periplasmic fraction, fermentation at extreme pH's and use of fusion tags to aid in expression and protein purification (De Marco et al., 2004).

A colony filtration (CoFi) blot method for rapid identification of soluble protein expression in *E. coli*, based on a separation of soluble protein from inclusion bodies by a filtration step at the colony level is described to screen a deletion mutagenesis library by Cornvik et al.,

(2005) while Coleman et al (2004) report a fluorescent based screening of soluble protein expression where specifically labeled proteins in cellular lysates are detected in one of three formats: a microplate using a fluorescence plate reader, a dot-blot using a fluorescence scanner or a microarray using a laser scanner.

Fusion tags have become indispensable tools for structural and functional proteomics initiatives (Banerjee et al., 2010). Fusion tags that are available for the ease of expression and purification of recombinant proteins include His tag (6-10 aa) (Smith et al., 1988), thioredoxin (109 aa) (Lavallie et al., 1993), Glutathione S-transferase (236 aa) (Smith & Johnson 1988), maltose binding protein (363 aa) (di Guan et al., 1998), NusA (435 aa) (Davis et al., 1999) etc. Although these tags are mainly used to promote the solubility of the target proteins and there by prevent the formation of inclusion bodies in *E. coli*, its use in screening recombinants for solubility of protein of interest is also demonstrated. Maxwell et al., (1999) report cloning of a gene of interest as chloramphenicol acyl transferase (CAT) fusion. Based on the resistance to higher concentration of chloramphenicol, it would be easier to predict the solubility of target protein. The authors have developed an *in vivo* system for assessing protein or protein domain of interest with CAT, the enzyme responsible for conferring bacterial resistance to chloramphenicol. CAT is highly soluble homodimeric protein that has shown to maintain activity when fused to various other proteins. It has been observed that CAT fusions to insoluble proteins confer lower chloramphenicol resistance than that of a fusion with highly soluble partner.

Similarly, Banerjee et al., (2010) have shown that the solubility of the target protein could be predicted *in situ* at the time of recombinant screening based on the intensity of the GFP fusion proteins. This work demonstrates that higher the solubility of the target protein , the intensity of the GFP fluorescence on the agar plate is higher rendering the screening of the recombinants a dual objective of identification and also predicting the solubility of the gene of interest attached to the reporter gene. The article as described by Banerjee et al., (2010) demonstrates this clearly. While GCSF, a human granulocyte colony stimulating factor gene which is known to get expressed as insoluble aggregates in *E. coli* shows lesser fluorescence as GFP fusion (Figure 3, panel B) the *E. coli* methionine amino peptidase, the well soluble *E. coli* protein exhibits higher intensity of GFP fusion (Figure 3, panel A).

A
MAP-GFP fusion clones

B
GCSF-GFP fusion clones

Fig. 3. Photograph showing two examples of GFP fusions in the pBAD24mGFP vector to demarcate the solubility difference by means of fluorescence. Panel A shows the *E. coli* methionine amino peptidase GFP fusion *E. coli* clone under long UV while Panel B shows GCSF-GFP fusion clone fluorescence under long UV. Higher glow indicates higher solubility. Clones #3 refers to a non-recombinant in case of MAP-GFP fusion while clone #2 and 5 are non-recombinants of GCSF-GFP fusions.

A similar strategy has been reported by Colman and Homes (2005) for integration confirmation in Pseudomonas strains. Here a pUC-based reporter plasmid (pUS23) was developed containing a recombination site [*aadB* 59 base element (59-be)] upstream of promoterless *aadB* [gentamicin (Gm) resistance] and *gfp* (green fluorescence) genes, and this construct was used to investigate the recombination and expression activities of the CI in *Pseudomonas stutzeri* strain Q capture of pUS23 at *attI* by an integron results in Pc-mediated expression of the aadB and gfp genes, which are silent in the initial construct. The final end result is gentamicin resistant and green fluorescent recombinants for positive integrants.

5. Conclusions

In this article, we have summarized some updated information about (1) new vectors that are commercially available to make the screening system of bacterial recombinants antibiotic free (2) new concepts about easy screening of recombinants utilizing the solubility property of the protein of interest and (3) specialized host strains and using the same clone for expression studies.

Developments in recombinant DNA technology have allowed rapid progress in the analysis of gene structure and function, and the production of potentially useful polypeptides in *Escherichia coli*. Often the experimentally limiting step has been the lack of a suitable screening methodology for expressed cDNAs. The extensive variety of screening bacterial colonies for recombinant plasmids arises from the fact that there is no single method for achieving fool-proof recombinant clone. Conventionally the screening methods employed routinely in academia and industry, for bacterial recombinants include colony hybridization, PCR and plasmid preparations. While all the methods involve cloning the gene of interest in a cloning vehicle and then reintroduction of the recombinant clone into another host cell for expression of the interest making the entire process time-consuming, laborious and expensive. While colony hybridizations require several days to a week and may involve the use of radioactivity, PCR based methods are expensive and have lengthy set-up and reaction times, plasmid preps require considerable hours for cell growth and preparation of mini-prep plasmids. An additional step of screening such recombinants for the solubility of the protein of interest makes the entire process labor-intensive and challenging.

6. Acknowledgements

The authors thank Dr. Kamal Sharma, Managing Director, Lupin Limited for being a constant source of encouragement. Thanks are also due to KrishnaMohan Padmanabha Das for critical reading of the manuscript, formatting the main text and the references.

7. References

Aaron K A Marvellous Machine for Making Messages.(2001) *Science* ,Vol. 292, No. 5523, pp. 1844-1846.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, & K. Struhl. 1988. Current protocols in molecular biology. Wiley Interscience, John Wiley and Sons, NY. Cambridge, MA.

Banerjee S, Deshpande AA, Mandi N & Padmanabhan S (2009) A novel cytokine derived fusion tag for over-expression of heterologous proteins in *E. coli*. *International Journal of Bioogical Life Sciencesl Vol. 5,* pp.129-133.

Banerjee S, Kumar J, Deshpande AA, & Padmanabhan S (2010) A novel prokaryotic vector for identification and selection of recombinants: Direct use of the vector for expression studies in *E. coli. Microbial Cell Factories*, Vol. 9, No. 30

Bergmans HE, Van Die IM, & Hoekstra WP (1981) Transformation in *Escherichia coli*: stages in the process. *Journal of Bacteriology*. Vol.146, No.2, pp. 564–570.

Bernard, P (1995). New ccdB positive-selection cloning vectors with kanamycin or chloramphenicol selectable markers. *Gene*, 162, pp. 159-160.

Broedel, Jr SE & Papciak,SM (2007) ACES™ Promoter Selection Vector Technical Brief

Cami B & Kourilsky P (1978) Screening of cloned recombinant DNA in bacteria by in situ colony hybridization. Nucleic Acids Research, Vol. 5, No. 7, pp. 2381–2390.

Cardona ST & Valvano MA (2005) An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in *Burkholderia cenocepacia. Plasmid*, Vol. 54, No. 3, pp. 219-228.

Chaffin DO & Rubens CE (1998) Blue/white screening of recombinant plasmids in Gram-positive bacteria by interruption of alkaline phosphatase gene (*phoZ*) expression. *Gene*, Vol. 28, pp.91-99.

Chen S, Songkumarn P, Liu J & Wang Guo-Liang (2009) A Versatile Zero Background T-Vector System for Gene Cloning and Functional Genomics. *Plant Physiologyl*. Vol. 150, pp. 1111-1121.

Chernajovsky Y, Mory Y, Vaks B, Feinstein SI, Segev D & Revel M (1983) Production of human interferon in *E. coli* under lac and tryplac promoter control. *Annals in New York Academic Sciences*., Vol. 413, pp. 88-96.

Choi YJ, Wang TT & Lee BH (2002) Positive selection vectors. *Critical Reviews in Biotechnology*, Vol. 22, pp. 225-244.

Coleman MA, Lao VH, Segelke BW, & Beernink PT (2004) High-Throughput, Fluorescence-Based Screening for Soluble Protein Expression. *Journal of Proteome Research*, Vol 3, No. 5, pp. 1024–1032.

Coleman, NV & Holmes AJ (2005) The native Pseudomonas stutzeri strain Q chromosomal integron can capture and express cassette-associated genes. *Microbiology.* Vol 151, pp. 1853-1864.

Cornvik T, Dahlroth Sue-Li, Magnusdottir A, Herman MD, Knaust R, Ekberg M & Nordlund P (2005) Colony filtration blot: a new screening method for soluble protein expression in Escherichia coli. *Nature Methods*, Vol 2, 507 – 509.

Curtis MD, Grossniklaus U (2003) A Gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiology* ,Vol 133, pp. 462–469.

Davis GD, Elisee C, Newham DM & Harrison RG (1999) New fusion protein systems designed to give soluble expression in *Escherichia coli*. *Biotechnology and Bioengineering*, Vol 65, No. 4, pp 382-388.

Davis SJ & Vierstra RD (1998) Soluble, highly fluorescent variants of green fluorescent protein (GFP) for use in higher plants. *Plant Molecular Biology*, Vol. 36, pp. 581–588

De Marco, V., G. Stier, S. Blandin, & A. de Marco (2004) The solubility and stability of recombinant proteins are increased by their fusion to NusA. *Biochemical and Biophysical Research Communications,* Vol. 322, pp. 766-771.

Deshpande AA, Sampali Banerjee, Jitendra Kumar & Padmanabhan S (2010) Vector for identification, selection and expression of recombinants. International Publication number WO2010/026601 A2.

di Guan C, Li P, Riggs PD & Inouye H. (1998) Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene*, Vol. 67, No. 1, pp. 21-30.

Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K & Pikaard CS (2006) Gateway-compatible vectors for plant functional genomics and proteomics. *The Plant Journal, Vol.* 45, pp. 616–629.

Elvin CM, Thompson PR, Argall ME, Hendry P, Stamford NP, Lilley PE, Dixon NE (1990) Modified bacteriophage lambda promoter vectors for overproduction of proteins in *Escherichia coli*. *Gene*, Vol.87, No.1, pp. 123-126.

Finbarr H (2003) Toxins-antitoxins: plasmid maintenance, programmed cell death and cell cycle arrest. *Science,* Vol. 310, pp. 1496–1499.

Gabant PT, Van Reeth PL, Dreze M F, Szpirer C & Szpirer J (2000). New positive selection system based on the parD(kis/kid) system of the R1 plasmid. *BioTechniques,* Vol. 28 , pp. 784-788.

Gallagher C N, Roth N J & Huber R E (1994) A rapid method for the purification of large amounts of an alpha-complementing peptide derived from beta-galactosidase (*E. coli*). *Preparative Biochemistry*, Vol. 24, No. 34, pp 297-304.

Gay P, Le Coq D, Steinmetz M, Berkelman T & Kado CI (1985) Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. *Journal of Bacteriology*, Vol 164, pp. 918-921.

Gralla, J.D. (1990) Promoter Recognition and mRNA Initiation by *Escherichia coli* Es70. *Meth Enzymol*., Vol. 185, pp.37-54.

Haag AF & Ostermeier C (2009) Positive selection vector for direct protein expression. *Biotechniques* Vol. 46, pp 453-457.

Hartley JL, Temple GF & Brasch MA (2000) DNA cloning using in vitro site-specific recombination. *Genome Reearch*., Vol.10, 1788–1795.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*. Vol.166, 557-580.

Hashimoto-Gotoh, T, Tsujimura A, Kuriyama K & Matsuda S (1993) Construction and characterization of new host-vector systems for the enforcement-cloning method. *Gene* , Vol.137, pp. 211-216.

Horn D (2005) Directional enrichment of directly cloned PCR products. BioTechniques, Vol. 39, No. 1, pp. 40–46.

Inouye S, Ogawa H, Yasuda K, Umesono K & Tsuji, F (1997) A bacterial cloning vector using mutated *Aequorea* green fluorescent protein as an indicator. *Gene* Vol. 189, pp 159-162.

Jacobson RH, Zhang XJ, Dubose RF & Matthews BW (1994) Three-dimensional structure of β-galactosidase from *E. coli*. *Nature*, Vol. 369, No. 6483, pp 761-766.

Katzen F (2007) Gateway recombination cloning: a biological operating system. *Expert Opinion in Drug Discovery*, Vol . 2, No. 4, pp 571-581

Kemp DJ & Cowman AF (1981) Direct immunoassay for detecting *E. coli* colonies that contain polypeptides encoded by cloned DNA segments. *Proceedings of National Academy of Sciences USA,* Vol. 78, No. 7, pp. 4520-4524.

Kim HG, Hwang HJ, Kim MS, Lee DY, Chung SK, Lee JM, Park JH & Chung DK (2004) pTOC-KR: a positive selection cloning vector based on the ParE toxin. *BioTechniques,* Vol. 36, pp. 60-64.

Ko MS, Nakauchi H & Takahashi N (1990) The Dose dependence of glucocorticoid-inducible gene expression results from changes in the number of transcriptionally active templates. EMBO Journal, Vol 9, No. 9, pp. 2835-2842

Langley KE, Villarejo MR, Fowler AV, Zamenhof PJ & Zabin I (1975) Molecular basis of beta-galacosidase α-complentation. *Proceedings of Naional Academy of Sciences USA*, Vol. 72, No. 4, pp. 1254-1257.

LaVallie ER, DiBlasio EA, Kovacic S, Grant KL, Schendel PF & McCoy JM (1993) A Thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Nature Biotechnology* Vol. 11, pp. 187 – 193.

Liang S, Bipatnath M, Xu Y, Chen S, Dennis P, Ehrenberg M & Bremer H. (1999) Activities of constitutive promoters in *Escherichia coli*. *Journal of Molecular Biology*, Vol. 292, No.1, pp.19-37.

Loir YLe, Gruss A, Ehrilch SD & Langella P (1994) Direct Screening of Recombinants in Gram-Positive Bacteria Using the Secreted Staphylococcal Nuclease as a Reporter gene. *Journal of Bacteriology*, Vol.176, No. 16, pp. 5135-5139.

Mandi N, Kotwal P & Padmanabhan S (2009) Construction of a novel zero background prokaryotic expression vector: potential advantages. *Biotechnology Letters*, Vol. 31, No. 12, pp.1905-1910.

Mandi N, Soorapaneni S, Rewanwar S, Kotwal P, Prasad B, Mandal G & Padmanabhan S (2008) High yielding recombinant staphylokinase in bacterial expression system-cloning, expression, purification and activity studies. *Protein Express Purification*, Vol. 64, pp. 69-75.

Manjula R (2004) Positive selection system for identification of recombinants using α-complementation plasmids *BioTechniques*,Vol. 37, No. 6, pp. 948-952.

Maxwell KL, Mittermaier AK, Forman-Kay JD & Davidson AR (1999) A simple in vivo assay for increased protein solubility. *Protein Science,* Vol. 8, pp. 1908–1911.

Mead DA, Pey NK, Herrnstadt C, Marcil RA & Smith LM (1991) A universal method for the direct cloning of PCR amplified nucleic acid. *Biotechnology (*N Y), Vol. 9, No.7, pp. 657-663.

Morrison SL. Transformation of *E. coli* by electroporation. *Current Protocols in Immunology,* 2001 May; Appendix 3: Appendix 3N.

Ohashi-Kunihiro S, Hagiwara H, Yohda M, Masaki H & Machida M. (2006) Construction of a positive selection marker by a lethal gene with the amber stop codon(s) regulator. *Bioscience Biotechnology Biochemistry*, Vol. 70, No. 1, pp. 119-125.

Reggie YCLo & Cameron LA (1986) A simple immunological detection method for the direct screening of genes from clone banks. *Biochemistry and Cell Biology*, Vol 64, pp 73-76.

Salunkhe S, Raiker VA,Rewanwar S, Kotwal P, Kumar A & Padmanabhan S (2010) Enhanced fluorescent properties of an OmpT site deleted mutant of green fluorescent protein. *Microbial Cell Factories,Vol.* 9, pp. 26

Sambrook J, Fritsch EF, & Maniatis. T (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. CSH Laboratory Press, Cold Spring Harbor, NY.

Schlieper D, Schmidt M, Sobek H & Brigitte von Wilcken-Bergmann (1998) Blunt end cloning of PCR fragments using a new positive-selection vector. *Biochemica,* No. 2, pp. 34-35.

Short JM, Fernandez JM, Sorage JA & Huse WD (1998) λ ZAP: a bacteriophage λ expression vector with in vivo excision properties. *Nucleic Acid Research,* Vol. 16, No. 15, pp. 7583-7600.

Smith DB & Johnson KS (1988) Modified glutathione S-transferase fusion proteins for simplified analysis of protein - protein interactions. *Gene*, Vol. 67, pp. 31-40.

Smith HE (2007) The transcriptional response of *Escherichia coli* to recombinant protein insolubility. *Journal of Structural and Functional Genomics,* Vol. 8, No. 1, pp. 27-35.

Smith MC, Furman TC, Ingoliaq TD, & Idgeony CP (1988) Chelating peptide-immobilized metal Ion affinity chromatography: A new concept in affinity chromatography for recombinant proteins. *The Journal of Biological Chemistry*, Vol. 263, No. 15, pp. 7211-7215.

Telford J, Boseley P, Schaffner W & Birnstiel M (1977) Novel screening procedure for recombinant plasmids. Science, Vol. 195 no. 4276, pp. 391-393.

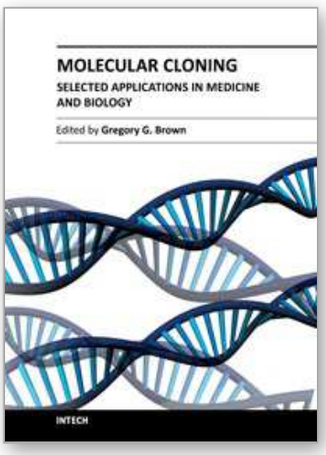
Trepe, K, (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and bio- chemical fundamental to commercial systems. *Applied Microbiology and Biotechnology*. ,Vol.72, No.2, pp. 211-222.

Yu RJ, Hong A, Dai Y & Gao Y (2004) Intein-mediated rapid purification of recombinant human pituitary adenylate cyclase activating polypeptide. B*iochimica et Biophysica Acta*, Vol. 36, No.11, pp. 759-66.

Zamenhof P & Villarejo M (1972) Construction and properties of *E. coli* strains exhibiting α-complementation of beta-galactosidase fragments in-vivo. *Journal of Bacteriology,* Vol. 110, No. 1, pp. 171-178

Zhou MY & Gomez-Sanchez (2000) Universal TA Cloning. *Current Issues in Molecular Biology*, Vol. 2, No.1, pp. 1-7.

IntechOpen



Molecular Cloning - Selected Applications in Medicine and Biology

Edited by Prof. Gregory Brown

ISBN 978-953-307-398-9
Hard cover, 324 pages
Publisher InTech
Published online 12, October, 2011
Published in print edition October, 2011

The development of molecular cloning technology in the early 1970s created a revolution in the biological and biomedical sciences that extends to this day. The contributions in this book provide the reader with a perspective on how pervasive the applications of molecular cloning have become. The contributions are organized in sections based on application, and range from cancer biology and immunology to plant and evolutionary biology. The chapters also cover a wide range of technical approaches, such as positional cloning and cutting edge tools for recombinant protein expression. This book should appeal to many researchers, who should find its information useful for advancing their fields.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen