Validation of Editing Efficiency of Cytosine Deaminase Base Editor by Varying Guide Length and Specificity – A CRISPR way

Akshaya. S
(17PBC001)

A Thesis submitted to Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore-641 043

In partial fulfilment of the requirement for the degree of Master of Science in Biochemistry

April, 2019

Validation of Editing Efficiency of Cytosine Deaminase Base Editor by Varying Guide Length and Specificity – A CRISPR way

Akshaya.S

(17PBC001)

A Thesis submitted to Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore-641 043

In partial fulfilment of the requirement for the degree of Master of Science in Biochemistry

April, 2019

23/4/19

Signature of the Head of the Department

Signature of the Supervisor

ACKNOWLEDGEMENT

"Gratitude turns disappointment into lessons learned, discoveries made, alternatives explored and new plans set in motion". I am grateful to **The Almighty God** for establishing me to complete this dissertation. I sincerely thank **Ayya Avargal** and **Amma Avargal** for creating a portal to exhibit out abilities. I express my deepest appreciation to all the higher authorities of Avinashilingam Institute for Home Science and Higher Education for Women for their immense support.

I take this opportunity of expressing my sincere thanks to **Dr**. **Krishnakumar,** Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for permitting and providing the infrastructure to undertake this investigation. I immensely thanks to **Dr. Premavathy Vijayan,** Vice Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for providing the entire facilities essential to carry out and complete the study. I record my sincere thanks to **Dr. Kowsalya**, Registrar, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for timely help rendered to carry out the work.

I place on record, my sincere gratitude to **Dr. P.R.Padma**, Head of the Department of Biochemistry, Biotechnology and Bioinformatics and Dean of School of Biosciences, Avinahilingam Institute for Home Science and Higher Education for Women for support to complete the project successfully.

I offer my heartfelt thanks to **Dr. Alok Srivastava**, Director, Center for Stem Cell Research, for providing me with an opportunity to work in this esteemed institution.I

I consider it a privilege to express my deep gratitude to ***Dr. Mohankumar K. Murugesan***, the members of his lab and other members of the Center for Stem Cell Research (a Unit of InStem), Vellore for providing me with unconditional support and the resources to carry out this project. I also appreciate their patient guidance, enthusiastic encouragement and useful critiques that were exceedingly helpful in the planning and development of this project.

I am privileged to express my heart bound thanks to **Dr. NirmalaDevi R** Assistant Professor (SG), Department of Biochemistry, Biotechnology and Bioinformatics, who with her constant guidance and care made to achieve good things.

I profusely thank **Mr. Vignesh R** for his constant guidance and entire support to carry out this project successfully. And also I wish to extend my gratitude to Nithin Sam,

Nivedhitha.D, Nazar, Dr.Muthu (lab members), other lab members and core lab members of Center for Stem Cell Research who gave me constant care and support to complete the project.

I owe my heart full thanks to all the **Staff Members** of the Department for Biochemistry, Biotechnology and Bioinformatics, Avinashilingam Institute for Home Science and Higher Education for Women Coimbatore, who extended their kind support and constant encouragement in my endeavors.

I dedicate this project to my **family** and **friends** who inculcate me with the knowledge, values, ethics and given me moral support and lovable care, which has been source of strength in the successful completion of this dissertation.

I finally wish to thank all the other hands who have not been acknowledged and those who have extended their valuable support directly or indirectly during the course of my project work.

AKSHAYA. S

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

1.0 INTRODUCTION

Genome Engineering technology provides unparalleled potential for modifying genomes in plants, animals and as well as in humans. Genome engineering in humans involved in curing genetic diseases, whereas in other organisms it provides the method to alter their biosphere which is likely to have a benefit to environment and human societies (Joung *et al*., 2015). Genome editing is the process of modifying the nucleotide sequence of the genome. Initially, genome editing is commonly called as gene targeting, which may serve as a \powerful research tool for scientists. This technology is now available to recruit biomolecules that modulates the regulation of gene expression and modifies epigenetic marks at specific chromosomal loci. The impact of recent developments in genome editing on science and biotechnology is immense. However, this incredible power over our heritable information also comes with a great responsibility to use it ethically and safely. This responsibility relates to both how we modify our own genomes and the genomes of our progeny, and the genomes of the species that inhabit our planet (Doudna and Gersbach, 2015).

The genome editing tools may offer many advanced biotechnological methods which provides efficient modification in the genome of the organism. The transformation of genome editing technologies in the last few years has also taught us a tremendous amount about scientific progress. In the earliest days of gene therapy, it was recognized that genome editing might be the ideal. A critical breakthrough in gene therapy was the discovery that by creating a site-specific DNA doublestranded break (DSB) in the target gene it is possible to stimulate genome editing by homologous recombination by 2–5 orders of magnitude, providing overall frequencies of 5 % or more. Hence, the double stranded break may serve as a key principle in the development of gene editing. The principle of genome editing technology is being developed not only to cure monogenic diseases but also to cure more common diseases that have arisen from multifactorial origins. The use of genome editing to cure monogenic disease is conceptually simple (genome editing can be used to correct the underlying genomic typographical errors), but the power of genome editing is that it provides a mechanism that can do more than simply modify single nucleotides. It is a method that can make more sophisticated and nuanced genomic changes, which can be used to cure more common diseases or to modify their course (Porteus, 2015).

Genome editing techniques are broadly classified into four basic and two hybrid types that includes Engineered Meganucleases, Zinc Finger Nucleases (ZFNs), Transcription Activator Like Effector Nucleases (TALENs), Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas9 nucleases), these are commonly employed basic tools of genome editing. Mega-TAL nucleases and Cas9-FokI nucleases are hybrid genome editing tools. The gene editing tools have wide variety of applications and they are come under one platform called as Nucleases. There are some differences between each of these nuclease platforms. For example, the type of break that is created in the DNA double strand is different. Meganucleases and Mega-TALs generate breaks with 3′ overhangs whereas ZFNs create breaks with 5′ overhangs and TALENs create breaks that are variable in position that are usually (but not always) 5′ overhangs, as determined by the properties of the FokI nuclease (Fn) and CRISPR/Cas9 nucleases create blunt breaks. But all these nuclease platforms have a common property of creating double stranded breaks in DNA strand so that they can share a fundamental mechanism of action (Xi *et al*., 2015).

The genome editing technologies are not widely appreciated because they can make a break in double stranded DNA just they can act as nucleases. But their power comes from the fact that all the genome editing tools can be designed to make a break very specifically at essentially any target sequence that is chosen by the experimenter. So, it allows the modification of practically any locus in the genome of any organism. This modification mainly depends on the DNA repair capabilities of the cell where the DNA break occurs. In nature all the cells or organisms undergoes two types of DNA repair processes. One is Non-Homologous End Joining (NHEJ) and another one is Homology Dependent Repair (HDR). In Non-Homologous End Joining method, the ends at the break can simply be re-joined, either precisely or imprecisely by small insertions or deletions at the break site, the creating targeted mutations. Homology Dependent Repair (HDR) mechanism can proceed by copying sequences from template DNA that has extensive homology with sequences around the break. This method uses a matched sequence from other cellular chromatid as template; it can divert the use of DNA supplied by the experimenter that leads to the sequence changes which results in the targeted sequence replacement (Carroll and Charo, 2015).

With a variety of gene editing tools they are readily available and their mechanism of action through knowledge, they have many more applications in the field of Science, Biotechnology and Medicine. It can be used to study the gene function and also to find out the

societal challenges as well as to improve the quality of human life. In agricultural field it can be used to address the challenges of feeding. For example, studies have already demonstrated that the editing of plant genomes which confers resistance to viral infection and provides protection from drought. In the field of medicine, it can help to treat inherited diseases, cancers, regenerative diseases and damaged tissues by manipulating the human DNA sequences. Even though there are still some challenges with these strategies due to its efficiency, delivery, and safety, we focussing on other strategies to overcome these problems in nowadays (Kim *et al*., 2016).

Meganucleases are derived from homing endonucleases they are highly site-specific dsDNA endonucleases that can be reengineered to expand their target site repertoires using various strategies, such as computational structure-based design, domain swapping, combined with yeast surface display for efficient detection of meganucleases with desired sequence specificities (Guha *et* al., 2017). Zinc Finger Nucleases are artificial proteins in which a zinc-finger DNA-binding domain is fused to the nonspecific nuclease domain from FokI restriction enzyme. Every individual zinc finger that specifically recognizes and binds to a nucleotide triplet, and fingers are often assembled into groups to bind to specific targeted DNA sequences. The major drawback is the design of ZFNs remains a complicated and technically challenging process, and it has low efficacy compared to that of other gene editing tools. Similar to that of ZFNs, TALENs are also artificial proteins they are the fusions of transcriptional activator-like effector (TALE) repeats and the FokI restriction enzyme. Every individual of TALE repeat targets is a single nucleotide which allows more flexible target design and increasing the number of potential target sites. The TALENs technique has many advantages and also has some disadvantages. One of the important drawbacks is the construction of TALE repeats remains a challenge and the efficiency of gene targeting with TALENs is variable. This kind of engineered nuclease has more specificity than ZFNs (Zhang *et al*., 2018).

CRISPR are derived from a bacteria based adaptive immune system. The CRISPR/Cas9 nuclease system is different from other types of engineered nucleases because all other nuclease systems are derived specifically through protein-DNA interaction whereas CRISPR/Cas9 systems derived from RNA-DNA Watson-Crick base pairing. There are two important components namely, single stranded guide RNA (sgRNA) and non-specific endonuclease called Cas9 protein present in the CRISPR/Cas9 nuclease system. The Cas9 protein is a multifunctional protein which in complex with sgRNA, to unwind the double stranded DNA

and as well as to find out whether the guide RNA is sufficiently identical to target site and then it creates a blunt double stranded breaks in the target DNA. Thus, CRISPR/Cas9 nucleases can be engineered very easily and it seems to be active at their desired target site. There are two main classes of CRISPR-Cas systems, which encompass five major types and 16 different subtypes based on cas gene content, cas operon architecture, Cas protein sequences. The CRISPR/Cas9 nuclease system focuses on genome editing applications in eukaryotes, and it argubly afford the most applications in both native and engineered forms in bacteria. There are three primary ways to render CRISPR-Cas systems such as depending on the CRISPR immunity stage, Cas machinery and outcome being exploited. CRISPR-Cas immune systems can be used to vaccinate against invasive genetic elements either naturally or by engineering. CRISPR-Cas systems can be exploited to provide resistance against phages (Fagerlund *et al*., 2015).

The experimental analysis indicates that CRISPR/Cas9-based genome editing could have widespread off-target effects, resulting in a significant level of non-specific editing at other unintended genomic loci. It has more advantages over than other genomic tools. The CRISPR/Cas9 system in bacteria has two essential RNA components namely, mature CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA). These two RNAs have partial complementary sequences and together form a well-defined two-RNA structure that directs Cas9 to target invading viral or plasmid DNA. The gRNA domain of the sgRNA determines the efficacy and specificity of the genome editing activities by Cas9. A major application of using CRISPR/Cas9 nuclease system over other techniques is its simplicity and versatility (Wong *et al*., 2015).

Although CRISPR/Cas9 nuclease system has wide variety of applications, drawback of this method is low editing frequency of Homology-Directed Repair (HDR). All CRISPR applications use wild type Cas9, editing by HDR which has some cleavage problems even it has well designed sgRNA. To overcome this problem, David Liu's lab introduced a technique called single base editing with CRISPR by creating a new Cas9 fusion proteins called as "single base editors" rather than try to improve HDR. These fusions contain Cas9 nickase and rat cytidine deaminase, which converts cytosine to uracil without breaking the DNA strands. Uracil is subsequently converted into thymine by DNA replication or DNA repair mechanisms. Likewise, adenine deaminases convert adenine (A) to inosine (I), which is treated as guanine (G) by polymerases, creating A•T to G•C substitutions. Adenine-deaminase-mediated base

editing (ABE) is more complicated than Cytidine-deaminase-mediated base editing (CBE) because no known naturally occurring cytidine deaminases catalyze adenine deamination in DNA rather than RNA. Studies showed that hundreds of diseases are good targets for base editing therapies with this kind of fusions. The single base editing methods are very simple and three important requirements are necessary for their functions such as, Cas9 fused to a cytidine deaminase, gRNA targeting Cas9 to a specific locus and a target cytosine at position 4-8 in the non-targeted strand (Komar *et al*., 2016).

System fuses an inhibitor of this process to Cas9 which rising the editing frequency to maximum of ~20% and the double stranded break formation is very low that is <0.1%, since the DNA is not directly cleaved. To increase base editing efficiency more than 50% copy the edits into the opposite strand of DNA. Base editing efficiency will increase the improvement of CRISPR delivery. For base editing ribonucleoprotein delivery is the good option and it is a big achievement in the CRISPR field (Addgene, 2017).

To optimize staining in flow cytometry antibody titration is recommended. Whilst antibodies will bind to high affinity targets on a cell, if they are present in excess they will also bind to low affinity targets. This results in an increase in background fluorescence and consequently a reduction in your ability to resolve populations, especially if there are subtle differences. Further more if the antibody concentration is too high, it may result in a false negative prozone effect

AIM AND OBJECTIVE

This study is designed to validate the base editing efficiency of cytosine deaminase with guides of varying length and to determine the specificity of the guides to make specific edits in a series of possible editable bases. Also to optimize the antibody by titration process for staining and animal studies. Methodology adopted in the present study is described in the next chapter.

2.0 REVIEW OF LITERATURE

Currently we are facing many challenges in the field of biotechnology. One such challenge is to develop an efficient and reliable ways to make targeted changes within the genome of cells. Traditional approaches of mutagenesis which requires chemical agents or transposons that needs extensive screening in order to recover desired mutations in the genome. This kind of challenges can be effectively eradicated with the help of genome engineering technology. There is a tremendous advancement which has been seen in both number and feasibility of genome engineering techniques (Wolfs *et al*., 2016).

This review of literature encompasses the information about the,

2.1 Gene Editing

2.2 Gene Editing Tools

2.3 Meganucleases

2.4 Zinc Finger Nucleases

2.4 TALENs

2.5 CRISPR

2.6 Base Editing

2.7 Sequencing of Constructs

2.8 HEK 293 T Cell line

2.9 HUDEP Cell line

2.10 Antibody titration

2.1 GENE EDITING:

Genome-editing tools provide advanced biotechnological techniques that enable the precise and efficient targeted modification at genomic loci. Targeted genome editing which offers an approach to produce a specific mutation in the genome of the cell. Mostly the genome editing techniques may follows double-stranded (ds) DNA viral vectors in differentiated human cells and RNA interference (RNAi) mediated targeted gene knockdown approaches but it have some drawbacks such as the protein composition of the viral capsid can be potentially immunogenic. On the other hand, the use of exogenously introduced dsRNA in RNAi technology can disrupt the “homeostasis” of the cellular machinery involved in gene silencing. This genome editing technology continuously expanding its dimensions to higher precision, larger scale and diversified applications in the biotechnological area. The risks involved in altering genomes through the use of genome-editing technology are significantly lower because most edits alter only a few nucleotides, producing changes that

are not unlike those found throughout naturally occurring populations. Once the genomic-editing agents have segregated out, there is no way to find a difference between a naturally occurring mutation and a mutation caused by gene editing techniques (Burkard *et al*., 2018).

In the wake of the identification of these gene editing systems, it is critical to understand their fundamental mechanisms of action. This will enable the extension of these tools to more diverse applications and their optimization for user-defined specifications. The expanse of mechanistic information that is currently lacking is both wide and deep. Many knowledge gaps need to be filled for each system and comparisons of these properties across systems will facilitate the construction of general rules for gene editing techniques. We are only now learning how these molecular machines find and interact with their DNA target sites. Similarly, the various cellular DNA repair processes that control genome editing results that can be harnessed efficiently and precisely only if we fully understand the mechanisms by which DNA breaks are recognized, processed, and restored. The complete understanding of the properties of these systems will similarly advance our capacity to design optimal tools and achieve the field's long-term goal of developing accurate computational predictions of genome editing outcomes (Hilton and Gersbach, 2015).

Genome editing is the most powerful research tool for scientists and it is previously called as gene targeting. In particular, the ease of gene targeting in yeast was one factor that made yeast was an important model organism in studies of the pathophysiology of human disease. The importance of gene targeting as a research tool was further highlighted by honouring Dr. Oliver Smithies and Dr. Mario Capecchi for their development of gene targeting in mouse embryonic stem cells and for their subsequent precise genetic engineering of mice and their transformational advance in understanding human pathophysiology and they were received a Noble price in Physiology or Medicine in the year of 2007 (Xu *et al*., 2015).

Earlier gene therapy was identified as ideal genome editing approach for curing many genetic diseases. But the studies showed that there is a low absolute frequency of gene correction by homologous recombination in human somatic cells (10^{-6}) were observed after gene therapy. Hence, the DSB became a key principle in the development of genome editing techniques (Yao, 2017).

The basic process of genome editing is to create a specific DSB in the genome and then allow the cell's own endogenous repair machinery to repair the break. The cell can repair the break using one of two basic mechanisms namely, nonhomologous end-joining (NHEJ) and homologous recombination (HR). Editing a single break in the genome of the

cell can be created by nonhomologous end joining mechanism which involves insertions or deletions at the site of the break. When editing of a single break occurs by homologous recombination using a provided donor sequence, precise nucleotide changes in the genome range from a single base insertion to the introduction of large cassette of genes. When the editing of two breaks occurs by NHEJ, chromosomal deletions, inversions or translocations can be created. These gross chromosomal rearrangements can be generated intentionally for therapeutic purposes, but they also must be evaluated because any nuclease platform has the potential to produce any off-target effects (McCaffrey *et al*., 2016).

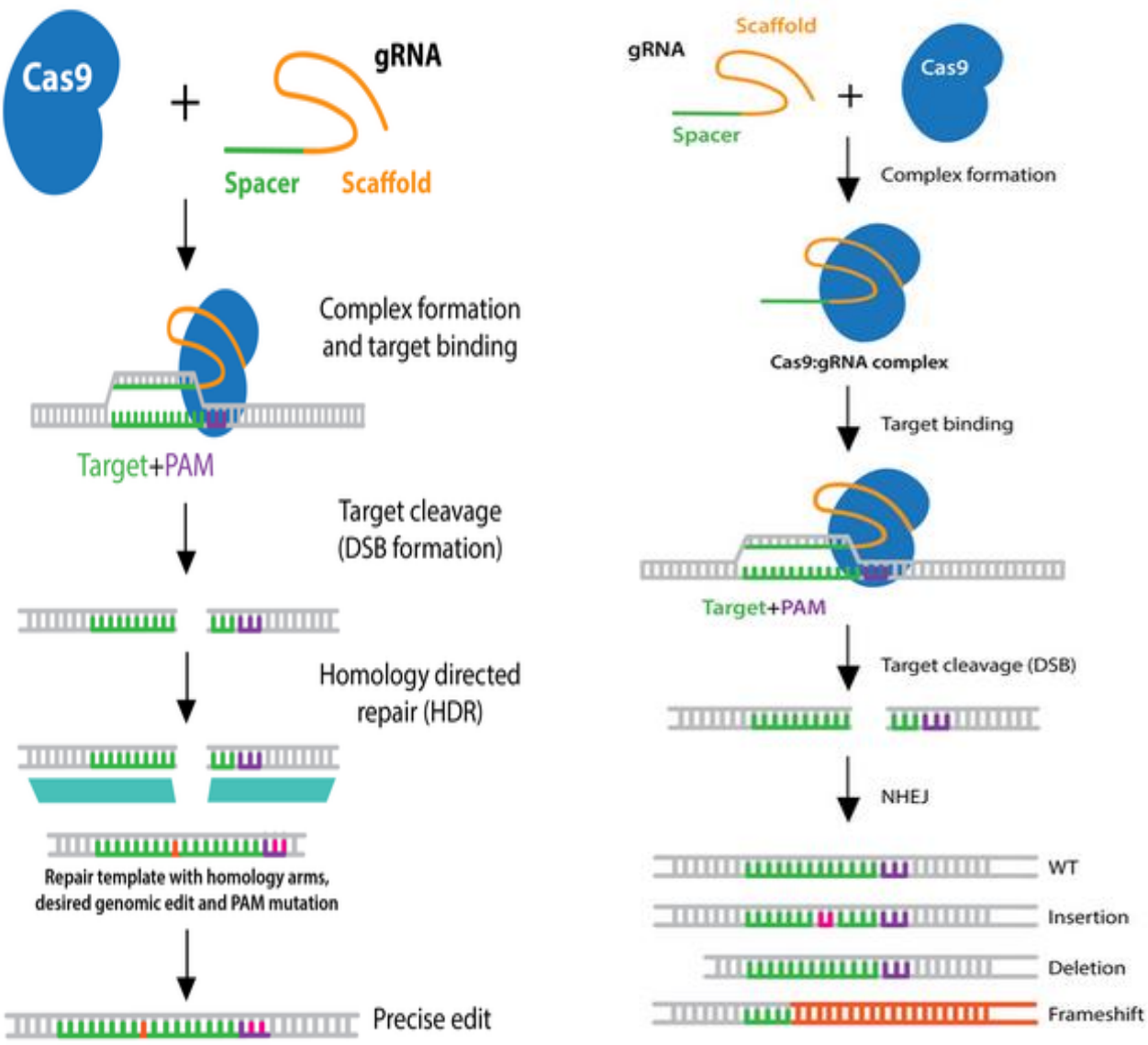


Fig. 1.3 Nuclease induced Double-stranded breaks (DSB) and its repair by Non-homologous end joining (NHEJ) and Homology directed Repair (HDR)

2.2 GENE EDITING TOOLS:

The genome editing tools has developed with continuing improvements in our understanding of the genetic contribution to nonmonogenic diseases, the principle of genome editing is being developed not only to cure monogenic diseases but also to cure more common diseases that have originated from multifactorial origins. The use of genome editing techniques to cure monogenic disease is conceptually simple and as well as to correct the underlying genomic typographical errors in the genome of the cell, but the power of genome editing tools which provides a mechanism that can do more than simply modify single nucleotides. It is a method that can make more sophisticated and nuanced genomic changes within the cell, which can be used to cure more common diseases or to modify their course. Genome editing tools having a great implications in the field of biotechnology, agriculture, medicine, etc., The latter applications are achieved in promoting the 'gene drive' of an introduced genetic element (such as a meganuclease) within an interbreeding population that can distort sex ratios (daughterless generations), or target genes related to fertility or pathogenicity. The exact nature of the therapeutic edit has to be driven by a solid understanding of the interplay between the underlying genetics and the specific pathophysiology of the disease. That is, one editing strategy might be appropriate for one disease but not applicable to another. Despite the variety of the genome editing tools available to us, the concerns of genomic errors introduced by genome editing approaches still remain (Varshney and Burgess, 2016).

Mainly four basic and two hybrid engineered nuclease platforms in gene editing systems that includes,

1.Engineered meganucleases,

2.Zinc Finger Nucleases (ZFNs),

3.Transcription Activator Like Effector Nucleases (TALENs),

4.Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas9 nucleases).

These are the important basic engineered nuclease gene editing techniques having wide range of applications.

1.Mega-TAL nucleases,

2.Cas9-FokI nucleases

These techniques have minor applications when compared to that of above mentioned engineered techniques and they come under the roof of hybrid techniques (Guha *et al*., 2017).

There are some differences between each of these nuclease platforms. For example, the type of break that is created by each nuclease is different from each other, Meganucleases and Mega-TALs generate breaks with 3′overhangs whereas Zinc Finger Nucleasess create breaks in 5′overhangs and the TALENs that create breaks in position that are usually 5′overhangs but not always, and FokI nuclease (Fn) and CRISPR/Cas9 nucleases create blunt breaks in the genome. In general, however, each of these platforms mediates their editing effects through the creation of a DSB and thus they share a fundamental mechanism of action. Non homologous end joining mediated genome editing requires engineered nuclease tool whereas the homologous recombination method need an engineered donor vector for their mechanistic actions. Donor vectors can be designed to template single bp changes or to insert large multi-gene cassettes into the genome. The homology arms for nuclease-mediated gene editing can be much shorter than those required for Homologous Recombination mediated gene targeting in murine embryonic stem cells instead of having to be 10 kilobases or greater, they can be as short as 400 bp. Because shortening the homology arms to below 400 bp which may decrease the overall editing efficiency. Single-stranded oligonucleotides (ssODNs) have also been used to template small nucleotide changes after the induction of a DSB. The ease with which ssODNs can be synthesized makes this approach relatively accessible to the researcher, but the mechanism by which ssODNs create a targeted change in the genome does not rely on the classic HR pathways. Moreover, ssODNs induce a replication and cell cycle arrest even in cancer cell lines (Barrangou, 2015).

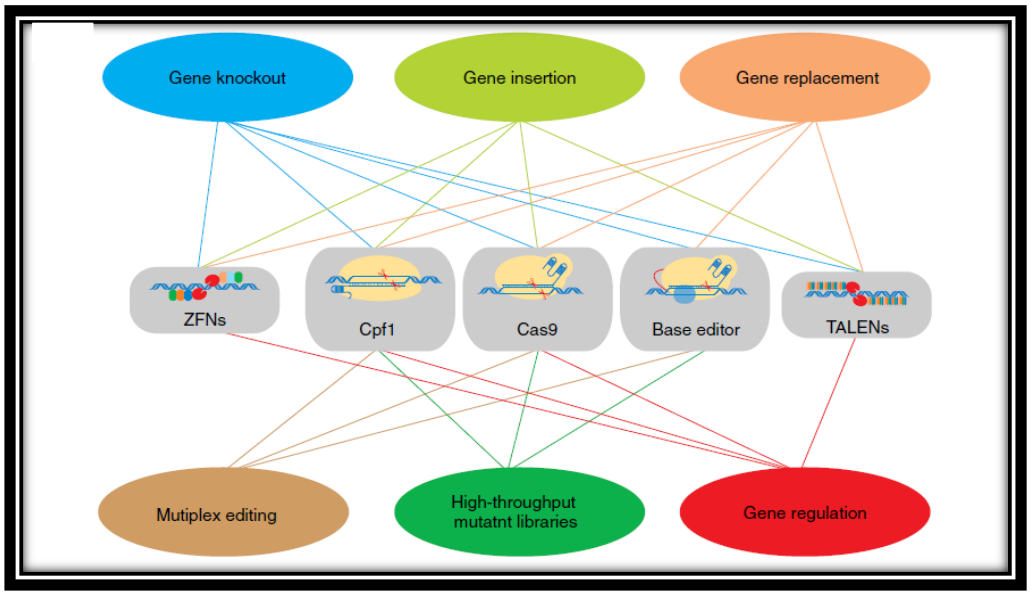


Fig 2: Network of genome editing tools

2.3 MEGANUCLEASES:

Engineered meganucleases are derived from the large family of natural homing endonucleases. A small number of these endonucleases have been designed to recognize natural target sites in the genome using a variety of strategies, including structure-based design and yeast surface display. Natural meganucleases have been recognized as the gold standard for specificity, but the challenge of engineering meganucleases for novel target sites has limited their translational development. Furthermore, the specificity of engineered meganucleases has not been fully evaluated. Meganucleases, or homing endonucleases are highly site-specific dsDNA endonucleases that can be reengineered to expand their target site repertoires using various strategies, such as computational structure-based design, domain swapping, combined with yeast surface display for efficient detection of HEases with desired sequence specificities. The LAGLIDADG family of meganucleases have been extensively studied and applied as genome editing tools. Unless otherwise mentioned, we are referring to LAGLIDADG enzymes as Meganucleases for simplicity (Christian *et al*., 2010).

One essential drawback for this class of enzyme is its non-modular configuration. The DNA recognition and cleavage functions can be, in part, intertwined in a single protein domain. Therefore, engineering of meganucleases has been challenging and has resulted in the development of other editing tools. However, a recent study suggests that there are multiple points across the LAGLIDADG protein that can be involved in holding metal ions in suitable positions to facility cleavage. These findings along with technologies, such as yeast surface display-SELEX, still hold promise for meganucleases to be engineered more efficiently in the near future. Moreover, single-chain modular nuclease architecture, termed ‘megaTAL’, was designed in which the DNA-binding region of a transcription activator-like (TAL) effector is appended to a site-specific meganuclease for cleaving a desired genomic target site. The latter synthetic version of a meganuclease provides a modular design, separating the endonuclease and DNA binding activities (Wang *et al*., 2016).

Therapeutic applications that demand precision with regards to gene modification activity can be addressed by these engineered variants of Meganucleases, as they are considered to be highly target-specific ‘molecular scissors’. Meganucleases are also in demand as components of vector/cloning systems (e.g. HomeRun vector assembly system) and synthetic biology applications (e.g. iBrick) that require rare-cutting enzymes. Even though the non-homologous end joining pathway is usually exploited to introduce mutations at the DSBs within the genome, sometimes, DSBs possess compatible “sticky” ends that can

be repaired without any introduced mutation. Recently, the 'MegaTev' architecture has been generated which involves fusion of the DNA-binding and cutting domain from a meganuclease (Mega, I-OnuI)with another nuclease domain derived from the GIY-YIG HEase (Tev, I-TevI). This protein was designed to position the two cutting domains ~30 bp apart on the DNA substrate and generate two DSBs with non-compatible single-stranded overhangs for more efficient gene disruption. More recently, similar to the MegaTev concept, Wolfs *et al*., have designed another dual nuclease, in which the Tev endonuclease domain is attached to the Cas9 nuclease domain, known as TevCas9. This hybrid nuclease, when introduced within human embryonic kidney cells (HEK293) along with appropriate guide RNAs, has been shown to delete 33 to 36 bp of the target site, thereby creating two non-compatible DNA breaks at moderately higher frequencies (40%). Therefore, this newly designed dual active endonuclease also promises to help genome editing events (i.e. introduce mutations) by avoiding the creation of compatible "sticky" ends which lead to a failed attempt of genome editing (Algasaier *et al*., 2016).

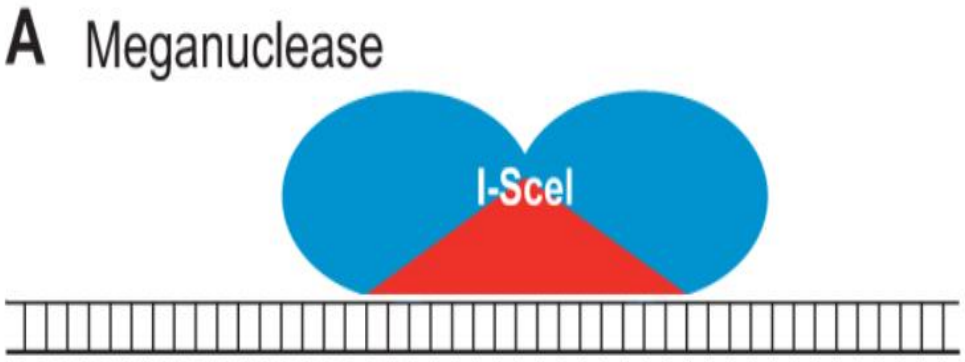


Fig 3: Meganuclease

2.4 Zinc Finger Nucleases:

More recently developed genome editing tools to be more flexible with regards to retargeting the reagent to different sequences by having a modular design such as a DNA-cutting domain that can be nonspecific and a distinct programmable DNA-binding domain which is more specific. Zinc Finger Nucleases are the first one to work as an efficient and targeted nuclease among others. Zinc Finger Nucleases are artificial proteins in which a zinc-finger DNA-binding domain is fused to the nonspecific nuclease domain from FokI. Briefly, The ZFNs are artificial endonucleases that have been generated by combining a small zinc

finger (ZF; ~30 amino acids) DNA-binding/recognition domain (Cys2His2) to a type IIS nonspecific DNA-cleavage domain from the FokI restriction enzyme. However, the cleavage activity of the FokI endonuclease demands dimerization. The ZFmodule which has been recognizes a 3 bp sequence, there is a requirement for multiple fingers in each ZFN monomer for recognizing and binding to longer DNA target sequences. In the past, using structure-based design, two ZFN variants were engineered that efficiently cleaved DNA only when paired as a heterodimer, thereby providing a potential avenue for improving the specificity of ZFNs as gene modification reagents. In a different structure-based study, using 3D protein modeling and energy calculations through computer-based software, researchers have been identified potential residues within the FokI dimer interface that are responsible for ZFN dimerization. These newly designed ZFNs were considered significantly less genotoxic which means they can cause cleavage at on-target sites in the cell-based recombination studies because the homodimerization could be prevented by lowering the dimerization energy, hence prevent activation of the dimeric FokI (Mino *et al.*, 2014).

The dimerized Zinc Finger Nucleases has recognizing the two adjacent sites in the genome and cleaves the DNA, then it initiating homologous recombination. Recently, ZFNs have been used as a potent antiviral therapy in the inactivation of specific coreceptors, thereby protecting cells from the viral entry in order to establish infection. Even though ZFNs showed impressive results in modifying the HIV CCR5 coreceptor surface protein in the autologous CD4 T lymphocytes of persons infected with HIV, there is still the risk of cleavage at ectopic sites due to the modular architecture of ZFNs and the non-specific nature of FokI. Apart from implementing ZFNs as genome editing tools, recently, the artificial zinc-finger protein (AZP)-staphylococcal nuclease (SNase) hybrid was designed (AZP-SNase) for potential antiviral therapies. This artificial nuclease can bind and cleave a specific origin of replication sequence of the human papillomavirus type 18 (HPV-18) thereby inhibiting viral replication in mammalian cells. However, the limitation of using this Zinc Finger Nucleases involves long term process of synthesis and nonmodular assembly processes. Even though some computational tools may help to improve targeting, but it is not possible to design a suitable ZFN pairs for every genomic locus. Another major disadvantage of this reagent is that the SNase has been shown to cleave both single and double-stranded RNA as well as the host DNA (single or double-stranded). Further modification involving switching of the SNase moiety in the AZP-SNase to the single-chain FokI dimer (scFokI) cleaved the viral DNA. Therefore, this newly designed hybrid ZFN is expected to serve as a novel antiviral reagent for inactivating human DNA viruses with fewer side effects (Singer *et al*, 2015).

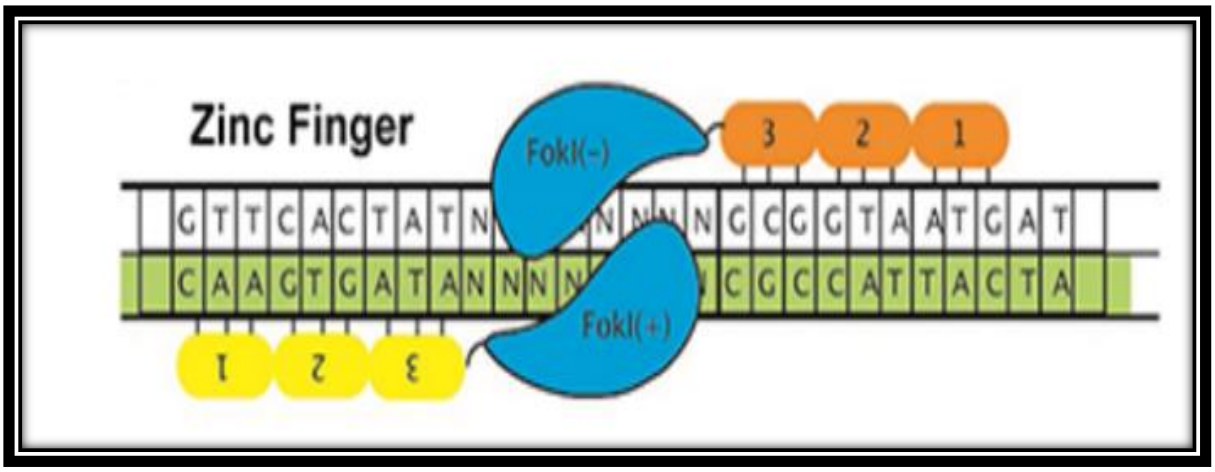


Fig 4: Zinc Finger Nucleases

2.5 TRANSCRIPTION ACTIVATOR LIKE EFFECTOR NUCLEASES (TALENS):

This Transcription Activator Like Effector Nucleases (TALENs) are first reported in the year of 2011 and it represents as an important tool in the genome engineering. Like Zinc Finger Nucleases, TALENs are also artificial endonucleases that can be designed by fusing the DNA-binding domain; it requires multiples of nearly identical repeats each comprised of approximately 34 amino acids obtained from TAL (transcription activator-like) effector (TALE) protein to the cleavage domain of the FokI endonuclease. In shortly, TALEN modular system is the fusion of TAL effector DNA binding protein which was isolated from *Xanthomonas spp,* AND FokI endonuclease. Each TALE repeat independently recognizes its corresponding nucleotide base with two variable residues and it is termed as Repeat Variable Di residues (RVDs) such that the repeats linearly represent the nucleotide sequence of the binding site. TAL repeats using engineered RVDs are currently being used to create a TALEN. These engineered RVDs might have increased specificity over natural RVDs. As for ZFNs, a pair of TALENs needs to be engineered to recognize a single target site. Even TALENs that use TAL repeats containing natural RVDs may have better specificity than ZFNs. Despite the tolerance to mismatches of longer TALENs *in vitro,* they seem to have higher genome editing activity and considered less genotoxic than ZFNs. TALENs can be redesigned to bind user-defined sequences by simply joining appropriate repeat units. Like ZFNs, TALENs are dimeric in nature; this is the important requirement for the design of two independent DNA-binding modules to target a single sequence. One advantage of the requirement for dimerization is enhanced specificity over monomeric enzymes. Although the

FokI enzyme is useful in terms of flexibility in the choice of various target sites, its nonspecific activity also increases the probability for more frequent cleavage at off-target sites in the genome. As an alternative approach to the FokI based architecture, monomeric Tev-TALE nucleases (Tev-mTALENs) were created. Here, the sequence-specific, monomeric nuclease domain from the I-TevI HEase is fused with TALEs. Thus, only a single DNA binding module is needed to target a sequence for cleavage. However, the use of a domain with predetermined recognition requirements, like TevI, significantly limits the range of genomic targets. The major disadvantage of using TALENs is the construction of TALE repeats remains big challenge and the efficiency of gene targeting with TALEN is variable (Han *et al*., 2018).

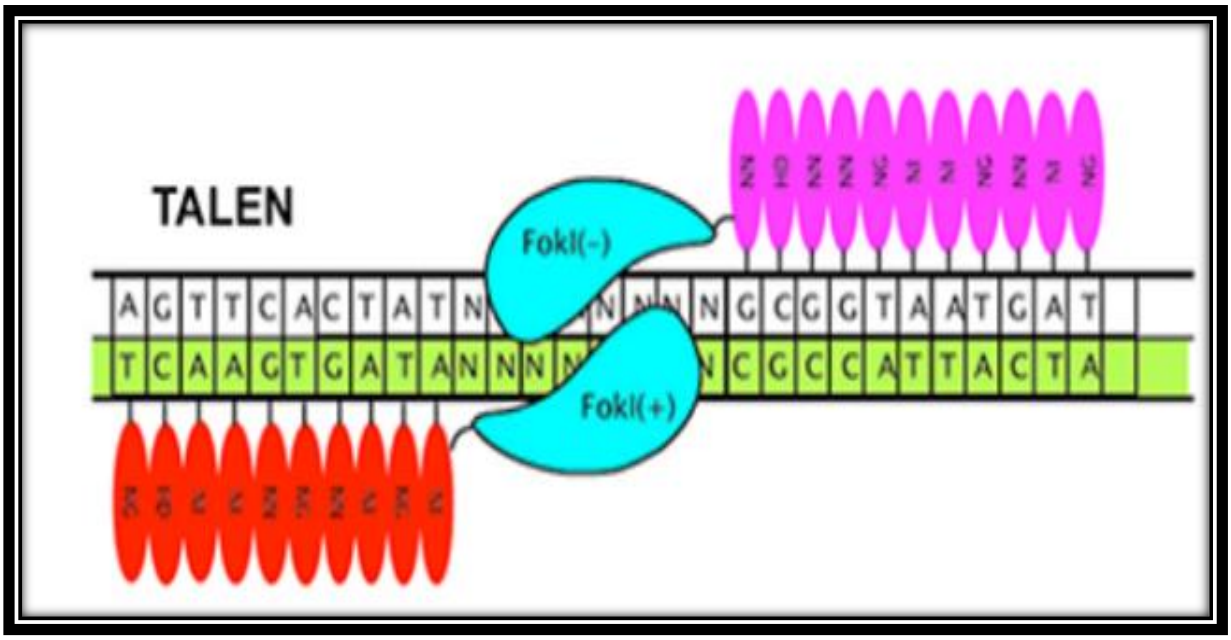


Fig 5: Transcription Activator Like Effector Nucleases (TALENs)

2.6 CRISPR:

For several years, scientists have been using gene targeting or gene editing to introduce new changes into a specific site in the genome either by removing or adding single bases or whole genes. In the current decade, there have been tremendous developments in the field of genetic engineering and related technologies. Bacterial CRISPR spacers are short, variable sequences derived from the genomes of viruses that previously invaded the bacteria. Such sequences provide ‘genetic memory’. During viral attacks, the CRISPR defence

mechanism of bacteria shears viral genome sequences analogous to spacer sequences and it is depicted in fig 5. If the invading virus is new, a new spacer is formed and archived into the sequences of spacers (Duan *et al*., 2018).

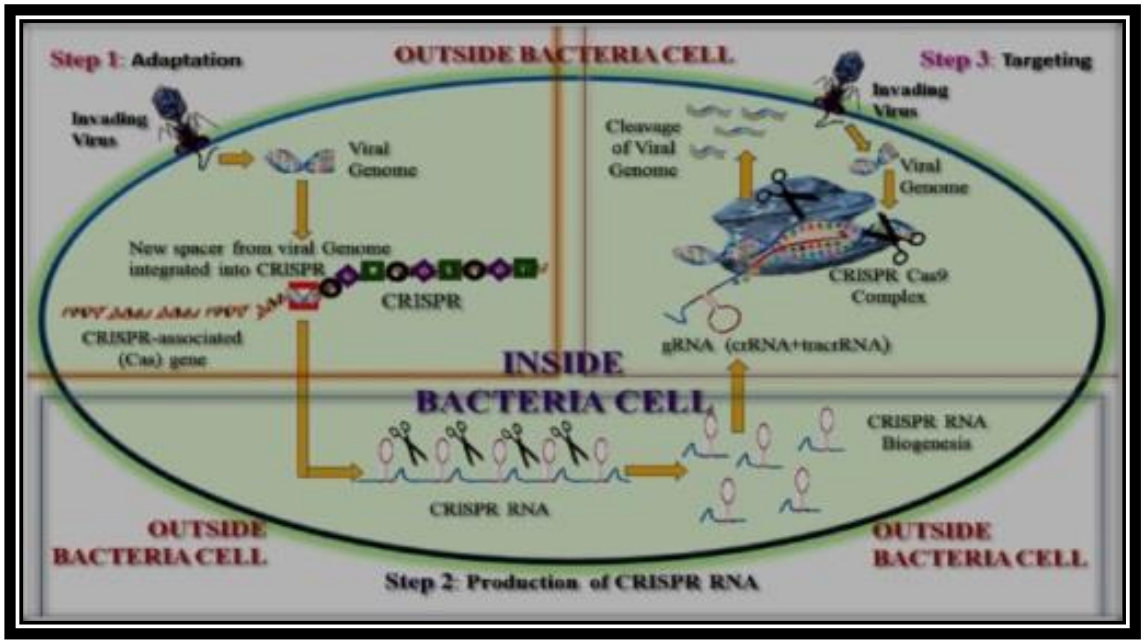


Fig 6: CRISPR/Cas9 System

Step 1 Adaptation involves getting DNA from the invading virus which is processed into short segments and these segments are inserted into the CRISPR sequence to function as new spacers.

Step 2 Production of CRISPR RNA involves the DNA undergoes a transcription process that copies DNA into RNA and the single-stranded RNA is cut into short pieces called CRISPR RNAs.

Step 3 Targeting involves CRISPR RNAs are programmed to destroy the viral material. Here, the 'RNA sequences' are copied from the viral DNA sequences (Thakore *et al*., 2015).

The critical breakthrough in the genome engineering tools is the development of CRISP/Cas 9 technology. The CRISPR/Cas9 system has been developed in recent years for genome editing, and it has been rapidly and widely adopted by the scientific community. The RNA-guided enzyme Cas9 originates from the CRISPRCas adaptive bacterial immune system. CRISPRs (Clustered Regularly Interspaced Palindromic Repeats) are short repeats interspaced with short sequences in bacteria genomes. CRISPR-Cas9 genome editing exploits

the CRISPR-Cas system to modify a genome in a targeted manner. Guided by RNA, the Cas9 endonuclease breaks DNA at a target sequence. Imprecise repair of the double strand break can result in insertion or deletion mutations, while repair pathways can be engineered to introduce specific point mutations or insertions. CRISPR-encoded RNAs have been shown to serve as guides for the Cas protein complex to defend against viral infection or other types of horizontal gene transfer by cleaving foreign DNA. Major progress has been made recently to modify the natural CRISPR/ Cas9 process in bacteria for applications in mammalian genome editing (Guilinger *et al*., 2014).

Compared with other genome editing methods namely meganucleases, ZFNs and TALENs, the CRISPR system is simpler and more efficient, and can be readily applied to a variety of experimental systems. The natural CRISPR/Cas9 system in bacteria has two important essential RNA components namely mature CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA). These two RNAs have partial sequence complementarity to each other and together form a well-defined two-RNA structure that only directs Cas9 to target invading viral or plasmid DNA. Some of the recent work indicates that it is feasible to engineer a single RNA chimera otherwise known as single guide RNA, or sgRNA by combining the sequences of both crRNA and tracrRNA. The sgRNA is functionally equivalent to the crRNA–tracrRNA complex, but is much simpler as a research tool for mammalian genome editing. In a typical CRISPR study, an sgRNA is designed to have a guide sequence domain and it currently designated as gRNA at the 5′end, which is complementary to the target sequence. The rationally designed sgRNA is then used to guide the Cas9 protein to specific sites in the genome for targeted cleavage (Chari *et al*., 2015).

The gRNA domain of the sgRNA determines both the efficacy and specificity of the genome editing activities by Cas9. Considering the critical roles of gRNA there are some multiple bioinformatics tools have been developed for rational design of gRNAs for the CRISPR/Cas9 system. Experimental analysis indicates that Cas9-based genome editing system has widespread off-target effects which results in a significant level of non-specific editing at other unintended genomic loci. Thus, most existing design tools have focused primarily on selection of gRNAs with improved specificity for genome targeting. However, more recent studies have been demonstrated that the off-target effects of the CRISPR-Cas9 system is not as extensive as previously speculated, and random targeting of the noncoding regions in the genome has little functional consequences in general. Furthermore, novel experimental systems have been developed to improve the targeting specificity of CRISPR/Cas9. Besides targeting specificity, another important aspect of bioinformatics

design is to select gRNAs with high targeting potency. Individual gRNAs vary greatly in their efficacy to guide Cas9 for genome editing. Thus, the design of potent gRNAs is highly desired, as inefficient genome editing by Cas9 will inevitably lead to significant waste of resources at the experimental screening stage. The importance of gRNA efficacy has only been appreciated very recently, with multiple studies attempting to identify sequence features that are relevant to functionally active sgRNAs (Butt *et al*., 2017).

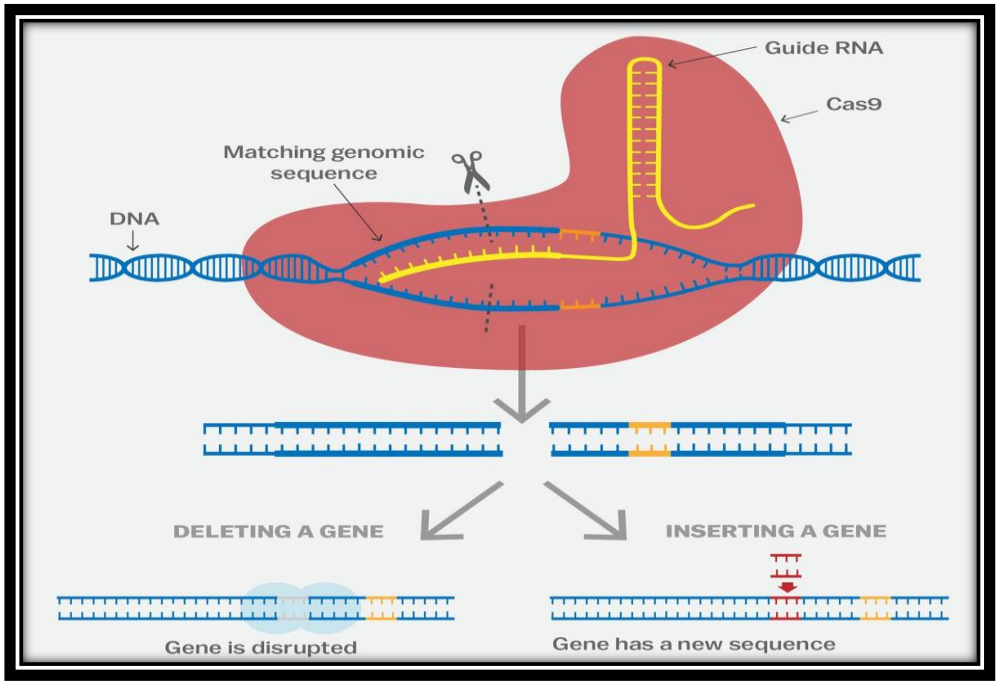


Fig 7: CRISPR/Cas 9 genome editing tool

CRISPR-Cas9 is a new molecular gene editing technique that enables scientists to make specific changes to the DNA of humans, animals as well as in plants. This technique is an analogous to a pair of molecular scissors that enable scientists to cut, silence or replace specific DNA sequences. The CRISPR-Cas9 gene editing technology can be used in both *in vitro* and *in vivo* which means it can be used to modify cells outside the body as well as it can be delivered into the body to modify specific target cells. Both applications have demonstrated effectiveness and ease of use in animal models. However, there is insufficient data to allow researchers to evaluate the risks and benefits of using this technology in humans (Burkard *et al*., 2018).

Dr. Schmuck-Henneresse explained that the Cas9 protein, which is derived from Streptococcus bacteria, forms an integral part of the CRISPR-Cas9 system. As streptococcal infections are common in humans, we hypothesized that there might be a pre-existing

immunological memory to Cas9. T cells from human immune cells that react to Cas proteins were found in almost all of the healthy humans. Cas molecules derived from other bacteria, such as staphylococci or gastrointestinal bacteria, also generate this type of immune reaction and this phenomenon that may be due to a high degree of similarity between these enzymes. These types of immune cells can produce unfavourable immune responses during gene therapy, thus potentially impairing the technology's safety and efficacy. The idea of the CRISPR-Cas technique has been adapted from the bacterial immune system. The CRISPR-Cas9 system has been widely adopted all over the world and successfully applied to target essential genes in different organisms and cell lines, including bacteria, zebrafish, monkeys, rabbits, mice and even humans (Jiang *et al*., 2013).

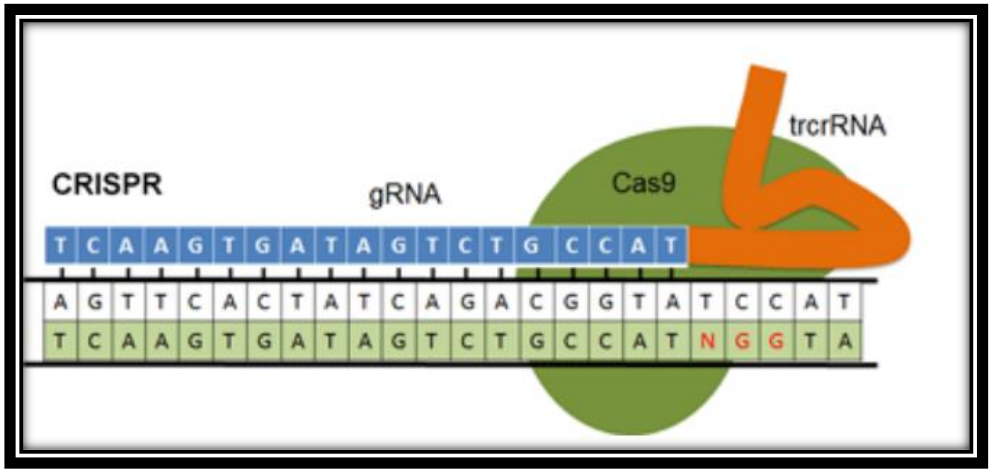


Fig 8: CRISPR/Cas9 engineered nuclease

2.7 BASE EDITING:

Homologous Recombination repair of DSBs using template donor DNA has been found to be much less efficient than template- free nonhomologous end joining, making it difficult to induce single nucleotide substitutions rather than indels. However, genome-wide association studies have shown that single-base changes are usually responsible for variations in elite traits. So there is a need of efficient techniques for producing precise point mutations in a specific target sites. CRISPR/Cas9-mediated base-editing technology is a new genome-editing approach that can accurately convert one DNA base into another, without the use of a DNA repair template (Sander and Joung, 2015).

The base-editing technologies employ Cas9 nickase (nCas9) or dead Cas9 (dCas9) fused to an enzyme with base conversion activity. For example, cytidine deaminases convert

cytosine (C) to uracil (U), and the latter is treated as thymine (T) in subsequent DNA repair or replication processes, so creating a C•G to T•A substitution. Likewise, adenine deaminases convert adenine (A) to inosine (I), which is treated as guanine (G) by polymerases, creating A•T to G•C substitutions. Cytidine-deaminase-mediated base editing (CBE) has been used in rice, Arabidopsis, wheat, maize, and tomato. Recently, this technology has been used in watermelon and wheat to create herbicide-resistant plants. Adenine deaminase-mediated base editing (ABE) is more complicated than CBE because no known naturally occurring cytidine deaminases catalyze adenine deamination in DNA rather than RNA. Fortunately, Gaudelli and colleagues, using several rounds of directed evolution and protein engineering, were able to develop an efficient ABE. In addition to generating point mutations, CBE can also be used to produce nonsense mutations that disrupt genes of interest and knockout their gene functions. CBE is much more specific than conventional SSN-mediated knockout, causing few if any indels. All-in-all, base-editing tools have given genome editing a new dimension, broadening its potential applications by means of nucleotide specific modifications at specific genomic sites (Kuscu *et al.*, 2018).

System fuses an inhibitor of this process to Cas9, raising editing frequency is maximum of ~20% and the double stranded break formation is very low that is <0.1%, since the DNA is not directly cleaved. To increase base editing efficiency more than 50% copy the edits into the opposite strand of DNA. Base editing efficiency will increase the improvement of CRISPR delivery. For base editing ribonucleoprotein delivery is the good option and it is a big achievement in the CRISPR field (Addgene, 2017).

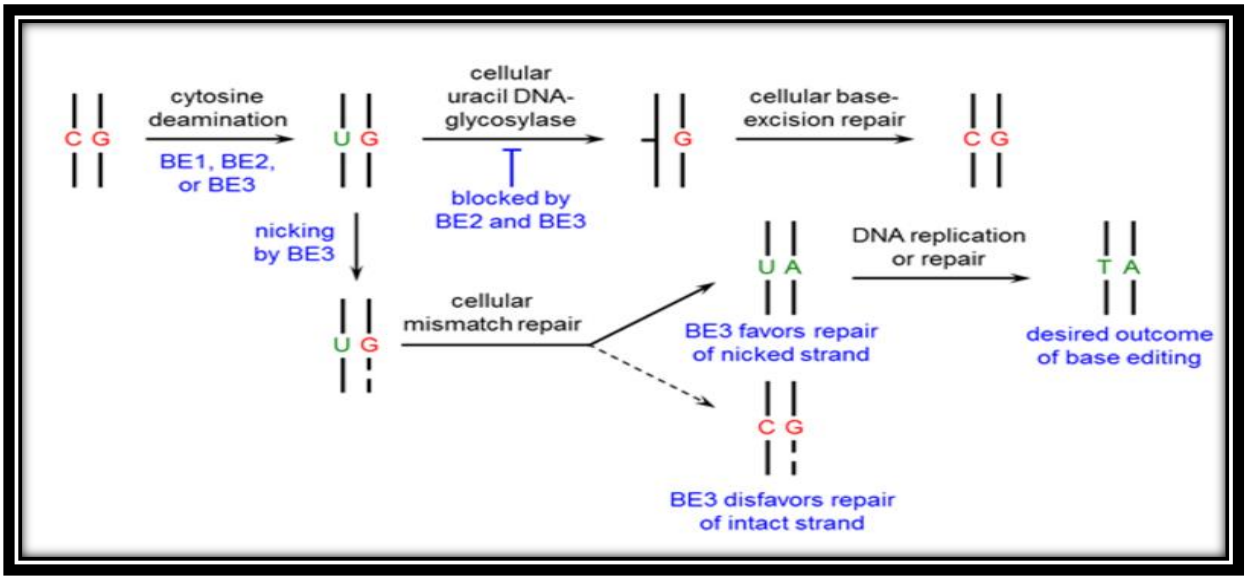


Fig 9: Base editing technique

2.8 Sequencing of Constructs :

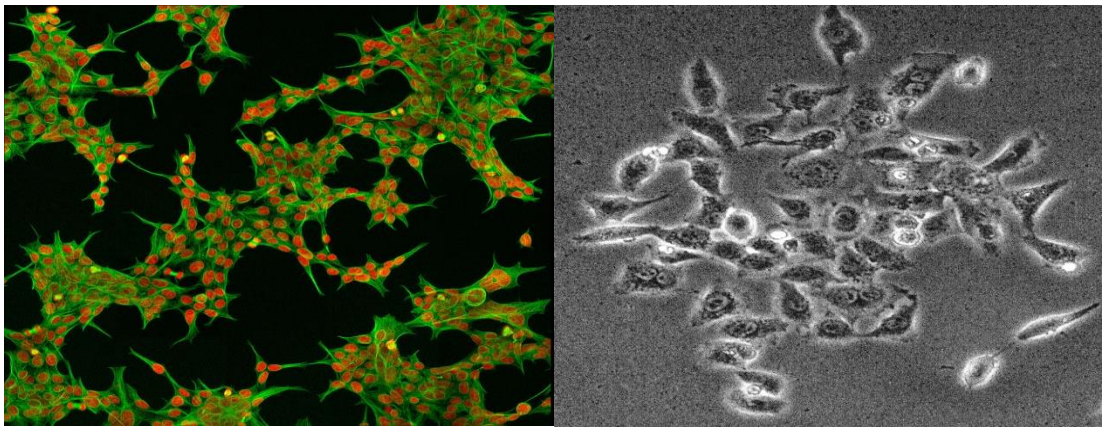
An automated fluorescence Sanger sequence detector machine is used for the process of sequencing. It was developed by Fredrick Sanger and first commercialized by Applied Biosystems. The basic requirements of this method are the DNA template, DNA primer and the DNA polymerase. De-deoxynucleoside triphosphates (dNTPs) and dideoxynucleoside triphosphates (ddNTPs) are also used in this process. The basic principle of Sanger sequencing involves the use of dNTPs and a low concentration of one of the four ddNTPs. These ddNTPs are generally fluorescently tagged to enable automated sequence detector. The working method of fluorescence Sanger sequencing basically involves dividing the template DNA into four equal parts and to each, only one of the four ddNTP (A/C/G/T) is added. For the other three nucleotides, dNTPs are added. The ddNTPs are at a concentration of about 100 times lower than that of dNTPs and when ddNTPs are encountered in the sequence, polymerization ceases to occur. Hence for each of the four reactions, after denaturation, DNA strands will be present with the fluorescently tagged ddNTP as its last nucleotide with which one can find out at which position each nucleotide is present.

2.9 HEK 293T Cell line

Cell culture studies provide a valuable complement for *in vivo* experiments; however, the necessary condition to use this strategy is that the immortalized cells display and maintain functional features as close as possible to related primary cells. In 1977 it was reported, for the first time, that human embryonic kidney cells, obtained from an aborted fetus, were transformed by exposure to sheared human adenovirus type-5 DNA, and therefore they became an immortalized cell line. However, these renal cell strains express only a limited number of typical markers of the kidney development, such as those of the mesenchymal-epithelial transition, but not others, such as those of the differentiated tubular segments, proximal tubules, and collecting ducts. Surprisingly, HEK 293 cells express many neuro markers, such as those typical of the neurofilament and other neuro specific metabolic enzymes (Archilli *et al*., 2018).

HEK 293T cell line is frequently and mostly used cell line in the laboratory to study the transfection, gene expression and virus production, cell cycle research, metabolism, receptor binding and different molecular biological techniques. This HEK 293T cell line is preferred mostly due to its merits that includes, adaptability and their growing capacity in

both serum or serum free media, stability, high gene expression rate, yielding large amount of proteins, growing capacity in both adherence and suspension culture conditions and so on (Sah *et al*., 2015)



(a) (b)

Fig 10: (a) HEK immunofluorescent image (b) HEK in culture medium

2.10 HUDEP CELL LINE

The recently established human umbilical cord blood-derived erythroid progenitor (HUDEP) cell lines have equipped red blood cell researchers with valuable in vitro models of erythroid development. Of the three established HUDEP cell lines, HUDEP-2 cells express predominantly adult β-globin and most closely resemble adult erythroid cells.

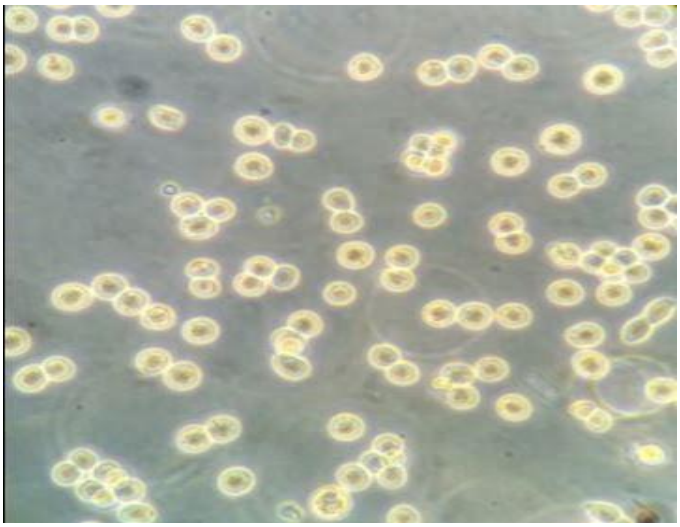


Fig 11: HUDEP cell line in growth phase

2.11 ANTIBODY TITRATION:

To optimize staining in flow cytometry antibody titration is recommended. Whilst antibodies will bind to high affinity targets on a cell, if they are present in excess they will also bind to low affinity targets. This results in an increase in background fluorescence and consequently a reduction in your ability to resolve populations, especially if there are subtle differences. Furthermore if the antibody concentration is too high, it may result in a false negative prozone effect. A titration experiment starts by selecting a fixed incubation time, cell type and experimental conditions. The last two should preferably match your final experiment. The cells are then stained in a series of dilutions of the antibody. It is good practice to add a viability dye even when titrating antibodies as dead cells can bind antibodies non-specifically, making your results hard to interpret. To determine the best antibody concentration, the stain index can be used as a guide.

AIM AND OBJECTIVE

This study is designed to validate the base editing efficiency of cytosine deaminase with guides of varying length and to determine the specificity of the guides to make specific edits in a series of possible editable bases. Also to optimize the antibody by titration process for staining and animal studies.

3.0 METHODOLOGY

The protocols adopted to achieve the mentioned objectives along with the materials, chemicals, instruments and equipment used are elucidated in the upcoming sections of this chapter.

3.1 Design of Guide Oligos for Target Genomic Sequence for Lentiviral Cloning

Sequence is retrieved for identified genomic targets through NCBI or UCSC Genome Browser using either search options for the gene or by genomic locations. The sequence is copied and stored. Now, the browser was directed to CHOP CHOP (http://chopchop.cbu.uib.no) and pasted the sequence. The genome assembly was selected as CRISPR CAS9. After the Target sites were found, three guides were selected. The guides are in the form of 20bp guide followed by an NGG PAM site. The guide sequences were chosen and the primers were designed for Lentiviral cloning of single guides in the format below.

Forward Oligo: 5’ CACCGNNNNNNNNNNNNNNNNNN 3’

Now the compliment of the guide sequence was taken, denoted by Xs.

Reverse Oligo: 5’ AAACXXXXXXXXXXXXXXXXXXC 3’

3.2 Cloning Strategy of Guide Plasmids

Guide RNA directs the Cas9 endonuclease to bind to a specific portion of the genomic DNA that is complimentary to the guide and to induce DSB. Traditional CRISPR strategies make use of plasmid constructs with the Guide cloned into it, driven by a promoter so that they can be manufactured by transcription within the cells. In most cases, the gRNA scaffold is at the end of the crRNA bit of the guide plasmid and also Cas9 gene with a suitable promoter is also present. This section will describe a cloning strategy used for cloning all the guides used in the project.

3.2.1 Ordering and Dilution of Oligo Guides

The guide oligos once designed, were synthesized at Bio Serve – Hyderabad. The obtained primers were diluted with around 500µl of nuclease free water to obtain a concentration of around 100µM.

3.2.2 Annealing of Oligos

When the primers for guide RNA is procured, it is obtained as two separate complementary strands which need to be annealed for cloning into specific vectors. The main aim is to allow both the individual complementary strands to denature. There are two main factors that influence the annealing efficiency. First, the salt concentration and the other being the rate of temperature decrease. Annealing happens most efficiently when the temperature is slowly decreased after denaturation.

For the annealing process, 100ng/µl of both the forward and reverse primers is added along with T4 DNA ligase buffer, T4 Polynucleotide kinase and nuclease free water. The total volume is made up to 10µl. The quantity of reagents for one reaction and their set up is given in table 1 and 2.

Table 1: Oligo annealing protocol and reagents required

REAGENT	**QUANTITY**
Forward primer	1µl
Reverse primer	1µl
T4 DNA ligase buffer	1µl
T4 Polynucleotide kinase	0.5µl
Nuclease free water	6.5µl
TOTAL VOLUME	**10µl**

The ramp temperature is dependent on the rate of decrease of temperature which is the reason behind the sudden temperature decrease from 95° to 25° C.

Table 2: Oligo annealing PCR set up

Temperature(°C)	Time (min)
37	45
95	5
25	15

3.2.3 Restriction digestion:

Restriction enzyme digestion takes advantage of naturally occurring enzymes that cleave DNA at specific sequences. There are many restriction enzymes, allowing scientist to target a wide variety of recognition sequences. Table 3 represents the reagents and their quantity used for one reaction.

Table 3: Reagents used for restriction and digestion

REAGENT	QUANTITY
Template	8 µl
10X NEB Buffer	5 µl
BsmB1	2µl
Water	35µl
TOTAL VOLUME	50 µl

3.2.3 Gel elution

Excise the DNA fragment from the agarose gel using a razor blade, scalpel or other device and transfer it into a 1.5ml tube. Add 3 volumes of ADB to each volume of agarose excised from the gel. Incubate at 37-55^{0}C for 5-10 min until the gel slice is

completely dissolved. Transfer the melted agarose solution to a Zymo-Spin™ Column in a collection tube. Centrifuge for 30-60 seconds. Discard the flow through. Add 200 µl of DNA wash buffer to the column and centrifuge for 30 sec. Discard the flow-through. Repeat the wash step. Add 6 µl DNA elution buffer or water directly to the column matrix. Place column into a 1.5ml tube and centrifuge for 30-60 sec to elute DNA.

3.2.4 Ligation of the annealed oligos into appropriate vectors

DNA ligation is the process by which DNA of strands are covalently stitched in pursuit of making recombinant DNA or plasmid. The DNA ligase that is purified from E. coli infected by T4 bacteriophage virus catalyzed the entire process of ligation and the reagents used for this process is given in table 4.

Table 4: Ligation of annealed oligos

REAGENT	**QUANTITY**
Vector	1µl (50ng)
Insert	6µl (25ng)
T4 DNA ligase buffer	2µl
T4 DNA ligase	1µl
Water	10µl
TOTAL VOLUME	20 µl

For this process, 50ng of vector and 25ng of insert is added along with 2µl of T4 DNA ligase buffer, 1µl T4 DNA ligase enzyme and the total volume is made up to 20µl along by adding nuclease free water. This whole reaction mixture is left at a 16°C water bath overnight. To do this, the annealed primers are diluted in nuclease free water in a ratio of 1:200.

3.2.5 Bacterial Transformation

Transformation can be defined as the uptake of exogenous DNA through its plasma membrane. For this to be possible, the bacteria must be in a state of competence. Competent cells can be prepared in the laboratory for experiments on recombinant technology. Hence transformation has two main parts to it. First, the preparation of competent cells and then heat shock for uptake of prepared plasmid.

Competent Cell Preparation

The Zymo competent cell kit was used for the process of preparing competent cells. First 0.5ml of fresh overnight grown E. coli cells of Xl1blue strain were inoculated in 50ml Zymo broth and is shaken vigorously (150-250 rpm) at 37°C till the OD reaches 0.4-0.6. The cells are then centrifuged at high speeds 3000-3700 rpm for ten minutes to obtain a pellet. The cells are to be kept in ice all the time and centrifugation must happen at 4°C.

The buffers that follow are available at 2X concentrations in the kit and are diluted to 1X dilution before use. The supernatant is removed and the pellet is re-suspended in 5ml 1X wash buffer that is provided along with the kit. Another round of centrifugation at 3500 rpm at 4°C is done. The supernatant is then discarded and the pellet is re-suspended in 5ml of competent buffer. The cells are then aliquoted as 50µl into single 1.5ml tubes for further use.

Transformation by heat shock

The competent cells that are prepared are taken out of -80°C and are thawed in ice. Around 100ng of the ligated DNA sample is taken and added to about 50µl of competent cells. Mixing process is performed by slight tapping with fingers. The mixture is then incubated in ice for about 30 minutes. Heat shock is then performed by placing the bottom half of the tube (containing the mixture) in a hot water bath that is maintained exactly at 42°C for 45-60 seconds. The tubes are immediately put back in ice for 5 minutes.

Making LB Agar plates

To make the agar plates, first 4g of L.B agar was dissolved in 100ml of MilliQ water. It was then autoclaved at 121°C and 1500 lbs pressure. After cooling the media, 1ug/ml of ampicillin antibiotic is added to avoid contamination. Then the complete media is added to the plates. After the media has cooled and solidified, the transformants that are kept in ice are transferred on the plates and streaked. The plates are then incubated at 37°C overnight. After 12-16 hours, colonies will be clearly visible.

3.2.6 Colony Pick and Inoculation

Clones are picked carefully using a 2µl pipette and added into the labelled 2ml tubes. For each plate, three clones are picked and inoculated in these tubes. The tubes are then placed in a shaking incubator rotating at 250 rpm at a temperature of 37°C for 14-18 hours.

Glycerol Stock

After 16 hours, 500ul of the samples are added to 500ul of 50% glycerol mix well and kept at -80°C. Then the remaining culture was centrifuged at 14000 rpm for 5 minutes. The supernatant is discarded and cell pellet was obtained. To the pellet add 500ul 1x PBS and resuspend well.

3.2.7 COLONY PCR

The reagents and colony PCR set up required for the one reaction is shown in table 5 and 6.

Table 5: The reagents used in Colony PCR

REAGENT	**QUANTITY (µl)**
DNA TEMPLATE	2
GFP RFP FORWARD PRIMER	1
GFP RFP REVERSE PRIMER	1
NUCLEASE FREE WATER	6
EMERALD MASTER MIX	10
Total	20

Table 6: Colony PCR set up

TEMPERATURE(°C)	**TIME (min)**
95	10
95	15sec
55	15sec
72	30sec
72	7
4	∞

3.2.8 Plasmid isolation using Mini-Prep kit

LB broth is first made by adding 2.5mg of Himedia LB broth to 100ml of water. The mixture is then autoclaved. After cooling 100ul of ampicillin antibiotic is added. 5ml of the media is then aliquoted into the required number of 50ml tubes and labelled. From the glycerol stock pinch of samples were added to respective tubes. The tubes are kept for bacterial growth in the shaking incubator for 16-18hrs at 37°c. Then the tubes are taken and kept for centrifugation at 4200rpm for 10min.

For the mini prep – Isolation of plasmid DNA, Machery – Nagel plasmid DNA purification kit is used. This kit contains a resuspension, lysis, neutralization, ethanol wash and elution buffers.

First 500µl of resuspension buffer is added to the cell pellet obtained in the previous steps. The mixture is then transferred to 2ml tubes. The resuspension buffer contains RNase to lyse any RNA present.

After resuspension, 500µl of lysis buffer is added to lyse the cells and bring out the plasmids present in the cytosol. The cells are incubated with the lysis buffer for an optimum time of 5 minutes. Leaving it for more time would result in shearing of DNA and lesser time would result in lesser yield of DNA. 600 µl of neutralization buffer is then added and mixed thoroughly to stop the lysis. The 2ml tubes are then centrifuged at 11,000 rpm for 10 minutes to obtain a clear supernatant. 750µl of the clear supernatant is added to the labelled NucleoSpin column that is placed on the collection tube (both are given along with the kit) and is centrifuged at 11,000 rpm for 1 minute. The flow through is discarded and repeated the step until the supernatant gets over. 500 µl of wash buffer is added to the column and centrifuged at 11,000 rpm for 1 minute. The flow through is again discarded and 600 µl of the ethanol wash buffer is added to the column and centrifuged at 11,000 rpm for 1 minute. The flow through is again discarded.

Finally, a dry spin is performed to remove all excess alcohol to prevent alcohol contamination. The empty column is spun at 11,000 rpm for 2 minutes. The collection tube is then discarded and replaced with labelled 1.5 ml vials. To the column, 30 µl of elution buffer that is maintained at 70°C is added and centrifuged at 11,000 rpm for 2 minutes.

The DNA is now eluted and the concentration should be anywhere greater than 400ng/ µl.

4.2.9 Sequencing for detection of positive clones

Sequencing PCR Set up

Sequencing was carried out with the BigDye® Terminator v3.1 Cycle Sequencing Kit. For sequencing one sample, 200 ng of template DNA is taken along with 1 µl of

Big Dye Sequencing buffer, 1.6 µl primers, 0.5 µl Ready Reaction Mix (RR mix 2x) and is made upto 10µl with nuclease free water and is kept in the PCR for the following cycle. The sequencing PCR set up is briefly given in table 7.

Table 7: Sequencing PCR set up

Temperature (°C)	Time
96	10s
50	5s
60	4 min
4	∞

Post Clean-up

Prior to sequencing, post PCR clean-up was carried out in the following way. Post cleanup is done by adding 10µl of DTR mixture to the PCR vials containing 10 µl of the PCR product along with 40µl of ethanol (80%). The mixture is then transferred to labelled 1.5ml tubes and placed on a magnetic stand to which only the DNA that is bound to the magnetic beads is attracted and the rest of supernatant is pipetted out. The pellet is then subjected to two washes by adding 100µl of 80% ethanol. After the washes, 40µl of injection solution was added that resulted in the elution of the DNA into the solution this sample can then further be sequenced.

Sequencing and Verification

Sequencing was carried out by an Illumina sequencer. The resultant sequence files are opened on Snapgene viewer and the presence of the insert and its sequence is confirmed.

3.2.10 Plasmid Isolation using Midi-Prep

One of the positive and control guide whose sequence is confirmed is selected and glycerol stocks are used for their inoculation into 100ml LB broth with ampicillin (preparation indicated in previous section).

The conical flasks are then incubated in shaking incubator at 250 rpm and 37°C for 16 – 18 hours. After the flasks are removed from the incubator, each flask is aliquoted into two labelled 50 ml falcon tubes. The tubes are then centrifuged at 4200 rpm at 4°C for 20 minutes. The supernatant is discarded.

For Midi prep plasmid isolation, Machery – Nagel plasmid isolation kit was used. It consisted of resuspension, lysis, neutralization, equilibration, elution and wash buffers - fil EF, endo EF and wash EF. It also contains a NucleoBond Column filter and a NucleoBond Column. To the pellet, 4ml of resuspension is first added and is dissolved in it by using a pipette. Similar to the mini pre, the resuspension buffer contains RNAse degrading RNA present. 4ml of lysis buffer is then added and mixed thoroughly by inverting the tubes few times to get a clear blue solution. This solution is incubated for 5 minutes. Leaving it for more time would result in shearing of DNA.

Post the 5 minute incubation, 4ml of neutralization buffer is added and the mixture is inverted 6 – 8 times till the solution turns completely white. This solution is incubated in ice for 5 minutes. Tangentially, the filter is fit inside the column and the entire column is fit inside a 250 ml conical flask using a washer that is provided along with the kit. To check if the filter is efficient and also to keep it moist, 15ml of equilibration buffer is added and allowed to run down. After the equilibration buffer has completely run down, the tubes with the neutralized solution are taken and mixed by inverting and then added gently into the filter column. When these samples run down, the DNA gets bound to the silica membrane that is present in the column. After all the samples have run down, the first wash is performed by adding FIL-EF to clear out all the contents from the filter. Then, two washes were done using 35ml of ENDO-EF and 15 ml

of WASH-EF. After both the washes are complete and all the reagents have run down, the DNA is eluted using 5 ml elution buffer.

3.5ml of isopropanol is added to the eluted mixture and is vortexed for 5 minutes to ensure high precipitation of the plasmid DNA. To obtain DNA pellet, the above solution is centrifuged at 4500 rpm for 30 minutes. After the pellet is obtained, the supernatant is discarded and the pellet is transferred to a labelled 2ml tube and centrifuged at 11,000 rpm for 10 minutes. After centrifugation, the ethanol is fully discarded and the pellet is allowed to dry till no traces of ethanol are seen which may otherwise lead to ethanol contamination. The dried pellet is then re-suspended with 200µl of T.E buffer. After the identification of target sites, designing of specific guide RNA and their consequent cloning and plasmid extraction, the plasmid DNA concentration would be about 2500ng/µl. The next process would be to deliberately insert this DNA into living mammalian cells.

3.3 CULTURING 293T

3.3.1 Revival of cells:

The cryovials were taken from the liquid nitrogen and thawed. The contents of the cryovial were transferred to a 15ml tube containing the expansion media and centrifuged at 1000 rpm for 5 minutes. Supernatant was removed and to the pellet expansion media was added and resuspended. The cell suspension was seeded in a T25 flask containing 5ml of expansion medium.

3.3.2 Subculture of cells

The media preparation for HEK and HUDEP-2 cells for subculture has given in table 8 and 9 respectively.

Table 8: Media preparation for HEK Cell

CELL LINE	CHARACTER	EXPANSION MEDIA	MEDIA PREPARATION
293T	Adherent	DMEM	DMEM + 10%FBS + 1% PENSTREP

HUDEP-2 Expansion Media

The StemSpan in used as the HUDEP-2 Expansion media which is prepared by adding the following cytokines in various concentrations

Table 9: Media preparation for HUDEP-2

S. No.	Components	Company	Stock Concentration	Working Concentration	Volume for 10 ml of medium
1	Rh-SCF	Immunotools	50 ng/µl	50 ng/ml	10 µl
2	Erythropoietin	Immunotools	2 U/ µl	3 U/ml	15 µl
3	Dexamethasone	Sigma	100 µM	1 µM	100 µl
4	Doxycycline	Sigma	1 mg/ml	1 µg/ml	10 µl
5	Glutamine	Hyclone	100 x	1 x	100 µl
6	Penstrep	Pan/Gibco	100 x	1 x	100 µl

3.4.1 Lipofection of Viral Plasmids

Cell Seeding:

Approximately 18–24 hours before transfection, 0.5million cells were seeded in 2.0ml complete growth medium per well in a 6-well plated and incubated overnight. Ideally cells should be 60-70% confluent prior to transfection. The reagents used for cell seeding have explained in table 10.

Table 10: Modified Protocol for MirusTransIT Lipofection

Reagent	**2 well of a 6 well plate (1 sample)**
OptiMEM	300µl
pMD2.G (PP1)	2µg
psPAX2 (PP2)	2µg
Transfer Plasmid	2µg
MirusTransIT LT1	7µl

The TransIT-LT1 Reagent: DNA complex was prepared (Immediately before transfection). The TransIT-LT1 Reagent was warmed to room temperature and vortexed gently before using. 300µl of Opti-MEM I Reduced-Serum Medium was added in a sterile tube. The packaging, envelope and the plasmid of interest was added in the desired concentration and pipetted gently to mix completely. Then, 7µl TransIT-LT1 Reagent was added to the diluted DNA mixture and gently tapped to mix thoroughly. It is then incubated at room temperature for 15–30 minutes. Now added this drop by drop to cells that were plated. The virus was collected at 24hrs and 48hrs, then centrifuge at 1500 rpm for 5mins at 4°C. Add the supernatant in new 15ml tube and add 1ml of 4x LentiX concentrator and kept in 4°C for one hour or 24 hours. 15ml tubes containing virus are centrifuged at 4200rpm for 45 mins and the pellet were obtained. The pellet was then resuspended in 150ul of 1X PBS to get a 50X concentrated virus. Then aliquoted 40 µl into each cryovial and stored in -80°C.

3.4.2 Transduction into HUDEP Cells

For transduction, HUDEP cells were counted and seeded at a density of 10^5 in each well of a 12 well plate in the presence of 0.75ul/ml of 1X Polybrene. The concentrated or unconcentrated virus was added into each well slowly. The plates were spinfected at a speed of 2250rpm for 1.5 hrs at room temperature. The entire medium was changed after 24hrs.

3.4.3 Selection of Cells

After 72 hours, fluorescent microscope was used to visualise for the selectable marker fluorescent protein as well as analysis by flow cytometry to know the percentage of transduced cells. These cells were grown out till a large number of cells obtained, then the cells were sorted by flow cytometry and the stable cell line was established.

3.5 First, the genomic DNA from the cells under consideration was extracted using Quick DNA Extraction protocol

3.5.1 GENOMIC DNA ISOLATION

Pipette 25µl Proteinase K and 200µl of blood into a 1.5ml tube. Add 200 µl Buffer B3 to the samples and vortex the mixture vigorously. Incubate the samples at 70^0C 10-15min. add 210µl ethanol to each sample and vortex again. For each preaparation take one nucleospin column placed in a collection tube and load the sample. Centrifuge for 1 min at 11000 x g. place the nucleospin column in a new collection tube and add 500µl buffer BW. Centrifuge for 1 min at 11000 x g and disard the collection tube with flow through.

Place the nucleospin column in a new collection tube and add 600µl buffer B5. Centrifuge 1 min at 11000 x g. discard the flow through and reuse the collection tube. Place the nucleospin blood column back into the collection tube and centrifuge for 1 min at 11000 x g. place the nucleospin blood column in a 1.5ml centrifuge tube and add 100µl preheated buffer BE(70^0C). Dispense buffer directly into a silica membrane. Incubate at room temperature for 1 min. Centrifuge for 1 min at 11000 x g.

3.5.2 PCR

A 50µl PCR reaction was set up using ~100ng of genomic DNA as a template. For each amplicon set up 3 PCR reactions using the following: i) template sgDNA from targeted cells (e.g. Cas9 transfected cells), ii) gDNA from negative control cells (e.g. non-specific DNA transfected cells), iii) water (i.e. no template control). The reagents required and PCR set up for the one reaction is briefly given in table 11 and 12.

Table 11: Reagents for PCR set up

REAGENT	**QUANTITY**
Forward primer (10 µM)	1µl
Reverse primer (10 µM)	1µl
Q5 Hot Start High-Fidelity 2X Master Mix	25µl
Template DNA	100ng
Nuclease free water	To 50 µl
TOTAL VOLUME	50µl

Table 12: PCR set up

STEP	**TEMPERATURE**	**TIME**
Initial Denaturation	98°C	30 seconds
35 cycles	98°C	5 seconds
	*50-72°C	10 seconds
	72°C	20 seconds
Final Extension	72°C	2 minutes
Hold	4-10°C	

3.5.3 RNA ISOLATION

Add 350µl buffer RA1 and 3.5µl β-mercaptoethanol to the cell pellet and vortex vigorously. {Reduce the viscosity and clear the lysate by filtration through nucleospin filter (violet ring)}. Place the nucleospin filter in a collection tube, apply the mixture, and centrifuge for 1 min at 11,000 x g. discard the nucleospin filter and add 250µl ethanol (70%) to the homogenized lysate and mix by pipetting up and down. For each preparation take one nucleospin RNA column (light blue ring) placed in a collection tube. Pipette the lysate up and down 2-3 times and load the lysate to the column. Centrifuge for 30s at 11000 x g. place the column in a new collection tube. Add 350µl of membrane desalting buffer and centrifuge at 11000 x g for 1 min to dry the membrane. Apply 95µl DNase reaction mixture directly onto the center of the silica membrane of the column. Incubate at room temperature for 15 min. add 200µl buffer RAW2 to the nucleospin RNA column. Centrifuge for 30s at 11000 x g. place the column into a new collection tube. Add 600µl buffer RA3 to the nucleospin RNA column. Centrifuge for 30s at 11000 x g. discard the flow through and place the column back into the collection tube. Add 250µl buffer RA3 to the nucleospin RNA column. Centrifuge for 2 min at 11000 x g to dry the membrane completely, place the column into a nuclease free collection tube. Elute the RNA in 60µl RNase free H20 and centrifuge at 11000 x g for 1 min.

3.5.4 ANTIBODY TITRATION

Add 100µl of whole blood (about 1 million cells) into each of the above tubes containing appropriate antibody concentration and unstained. Incubate in dark at room temperature (RT) for 10-15 minutes. Add 2ml FACS lyse (1x) solution.vortex. Incubate in dark at RT for 10-12 minutes. Centrifuge cells at 200-300g for 5 minutes. Discard supernatant and break the pellet by gently tapping the bottom of tube. Add 2ml of sheath fluid .Vortexed and centrifuged at 200-300g for 5 minutes, discard supernatant. And break the pellet. Resuspend cells in 0.5 ml of stain Buffer or PBS with 0.5 % paraformaldehyde. Acquire in pre-calibrated flow cytometer. Data analysis and plotting signal/noise graph. Open the BD FACSDiva software for acquisition and analysis.

4.0 RESULTS AND DISCUSSION

Results obtained for the experiments performed as described in previous chapters for our objectives are discussed in this chapter along with appropriate images. The interpretations of the observed data are discussed with scientific basis as per literature which share data concordance. The genome editing tools may offer many advanced biotechnological methods which provides efficient modification in the genome of the organism. The transformation of genome editing technologies in the last few years has also taught us a tremendous amount about scientific progress. The experimental analysis indicates that CRISPR/Cas9-based genome editing could have widespread off-target effects, resulting in a significant level of non-specific editing at other unintended genomic loci.

4.1 SELECTION OF APPROPRIATE VECTOR

Appropriate vector selection is essential for the study to make it possible and the basic architecture of pL-CRISPR.EFS.GFP is depicted in fig 9.

This vector backbone pL-CRISPR.EFS.GFP expresses the Lentiviral CRISPR-Cas9 delivery for spcas9 and sgRNA. It also coexpresses egfp via P2A cleavage site. Also, EFS Promoter driven and includes an Ampicillin resistant gene needed to obtain the transformant colonies. It is a high copy number plasmid containing a BSMB1 digestion site which is ordered from ADDGENE plasmid repository.

4.1.1 GUIDE OLIGO DESIGN

The guides were designed using the online software, CHOP-CHOP and adding to it appropriate overhangs to make it ligatable at the BsmbI site. The exon sequence and the guides designed (along with their oligos) are indicated in table 13. PAM site (NGG) and the overhangs are denoted by highlighted bases (in the guide) and lowercase bases (in the oligos) respectively.

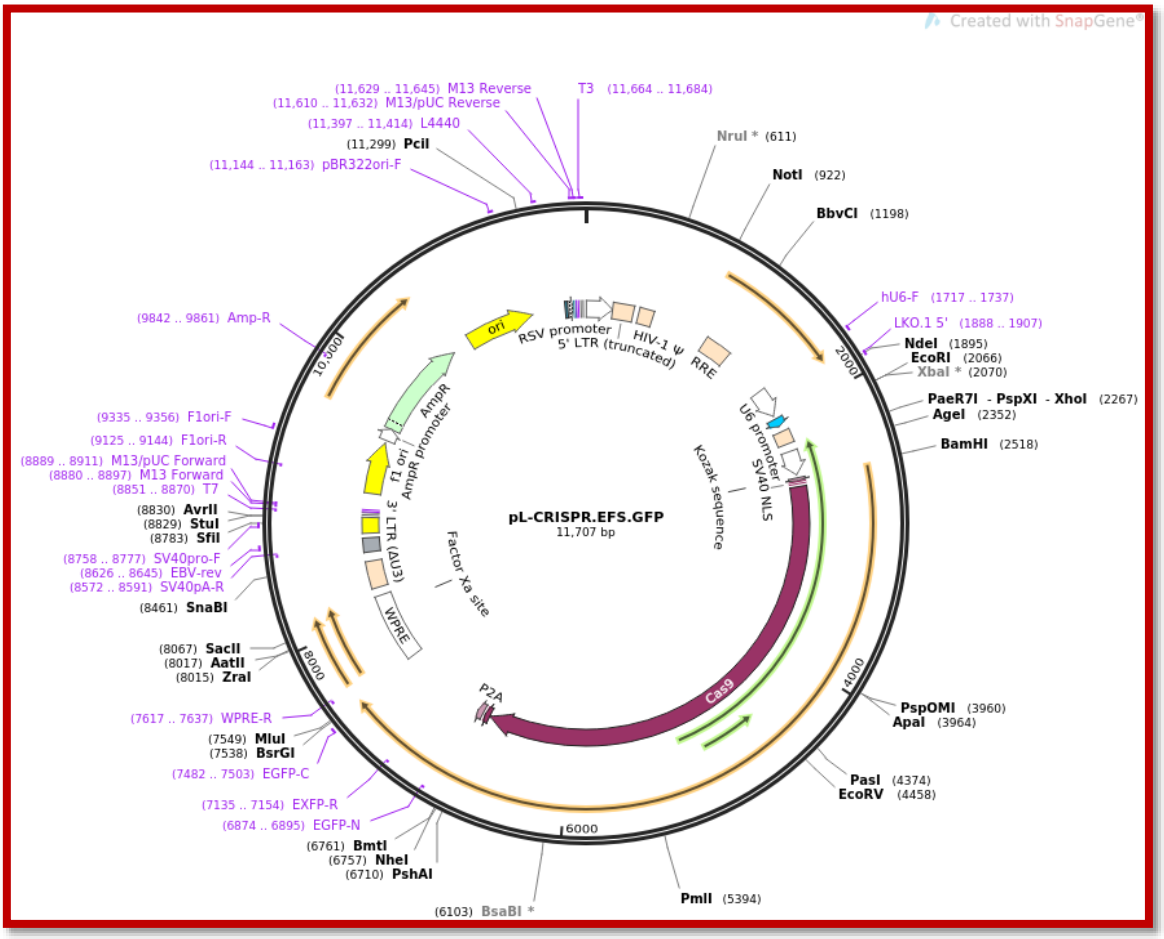


Fig 9: Vector Backbone

Table 13: Designed primers for cloning

Sequence Name	Sequence Text	TM	Concen-tration	Volume in ul to be added to get 100pm/ul or 100uM CONC
BE_16_19BP_F 1083	CACCGCCTGGCTAAACTCCACCCA	63	213.05	426.1
BE_16_19BP_R 1084	AAACTGGGTGGAGTTTAGCCAGGC	59	209.66	419.32
BE_16_18BP_F 1085	CACCGCTGGCTAAACTCCACCCA	61	206.54	413.08
BE_16_18BP_R 1086	AAACTGGGTGGAGTTTAGCCAGC	57	229.14	458.28
BE_16_17BP_F 1087	CACCGTGGCTAAACTCCACCCA	59	237.74	475.48
BE_16_17BP_R 1088	AAACTGGGTGGAGTTTAGCCAC	55	219.99	439.98
BE_16_16BP_F 1089	CACCGGGCTAAACTCCACCCA	58	205.1	410.2
BE_16_16BP_R 1090	AAACTGGGTGGAGTTTAGCCC	54	217.14	434.28
TARGET11_C6G_F 1081	CACCGTTGGCGTCTGATTAGGGTGG	63	229.07	458.14
TARGET11_C6G_R 1082	AAACCCACCCTAATCAGACGCCAAC	59	229.32	458.64

4.1.2 RESTRICTION DIGESTION

Restriction enzymes specifically cleave the nucleic acids at specific nucleotide sequences called restriction sites to generate a set of smaller fragments. Restriction enzymes form part of a restriction modification system of bacterial cell that provides protection against invasion of the cell by foreign DNA – especially bacteriophage DNA. But the cells own DNA is not cleaved by these restriction enzymes. This self-protection is achieved by the help of the specific DNA methyl transferase enzyme which will methylate the specific DNA sequence for its respective restriction enzymes by transferring methyl groups to adenine or cytosine residues to produce N-6-methyladenine or 5- methylcytosine. The selected appropriate vector was subjected to restriction and digestion in order to determine its number of basepairs and size of the vector.

The digested product was run on agarose gel for 2 hours 45 min at 90V. Then the gel was viewed under gel doc. The bands were observed around 6000 bp. The vector size was found to be around 8000bp. Thus, the vector of interest was fragmented and gel elution was performed to obtain the backbone for further cloning procedures. The result is given in fig 10.

Fig 10: restriction digestion

There are many literatures available on restriction digestion studies.

Fomenkov *et al*., 2017 have done restriction digestion on newly developed *ecoBLmcrX* vector where it specifies the modification dependent restriction activity.

Tsai *et al*., 2017 have reported that phage 9 g DNA sensitivity to >200 Type II restriction endonucleases (REases). Among the REases tested approximately 29% generated complete or partial digestions, while the remaining 71% displayed resistance to restriction.

4.1.3 SELECTION OF COLONIES AFTER TRANSFORMATION

Transformation process allows a bacterium to take up genes from its surrounding environment; that is transformation involves the direct uptakes of fragments of DNA by a recipient cell and the acquisition of new genetic characteristics.

NEB 10-beta Competent *E. coli* is a derivative of the popular DH10B. It is T1 phage resistant and endonuclease I (*endA1*) deficient for high- quality plasmid preparations. The ligated product transformed with the help of B10competent cell. Numerous colonies were observed after 16 hours of incubation at 37^{o}C which is shown in the plate 1. Single colonies in triplicates from each plate were picked and inoculated in LB broth for plasmid isolation and identification.

Several literatures have done the transformation studies.

Anand *et al*., 2018 have reported that the ternary vector system facilitated screening different origins of replications on the accessory plasmid and T-DNA vector and four combinations were identified that have high raw transformation frequency of about 86-103%.

Miyango *et al*., 2018 have explained that novel transformation strategies for transforming *Geobacillus kaustophillus* using Pls20cat-mediated plasmid mobilization.

In a study conducted by Hasegawa *et al*., 2018 reported that the contribution of transformation type HGT to genetic dynamics in the environment may be underestimated and the study indicated HPTT in *E.coli* occurs at substantial transfer frequencies at 10-5 to 10-10 under the conditions that can be feasibly encountered in the environment.

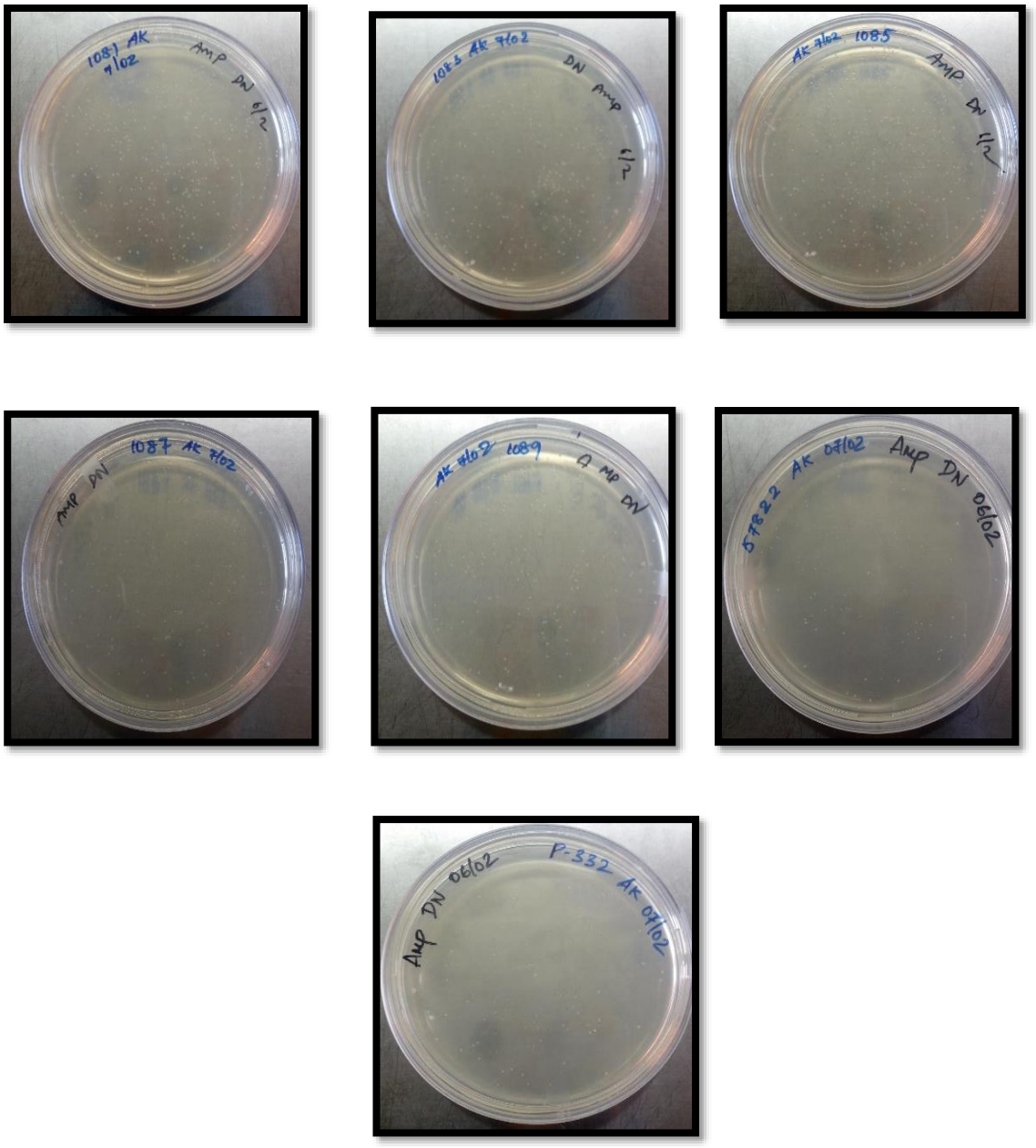


Plate 1: Colony growth after transformation

4.1.4 SCREENING FOR RECOMBINANTS USING COLONY PCR

Colony PCR involves lysing the bacteria and amplifying a portion of the plasmid. Insert-specific primers or vector-specific primers were used to screen for recombinant plasmids. Now we have transformed our DNA and allowed the colonies to grow overnight. To determine if they contain the insert of interest, they are screened by colony PCR. Followed by gel electrophoresis was performed. The bands were observed around 750bp (Fig 11) which confirms that the cells are transformed.

Several literatures have done colony PCR studies.

Huang *et al*., 2017 concluded that cloning accuracy is maintained constantly at 95% under various cloning conditions.

Jamal *et al*., 2017 have reported that the colony PCR method is significantly fast, less expensive, less laborious and most importantly, as reliable and precise as other methods to extract bacterial DNA.

4.1.5 SANGER'S SEQUENCE ANALYSIS

After the screening for recombinants were confirmed, the colony PCR product was taken for sequencing PCR. Sanger's sequencing was performed using BE16 forward primer revealed that each sequence obtained was matched with respective primers. The results are depicted in fig 12.

The results revealed that the sequencing by colony PCR which confirmed the sequencing of transformants were done using BE16 forward primer and the obtaine sequences are perfectly matched with appropriate primers.

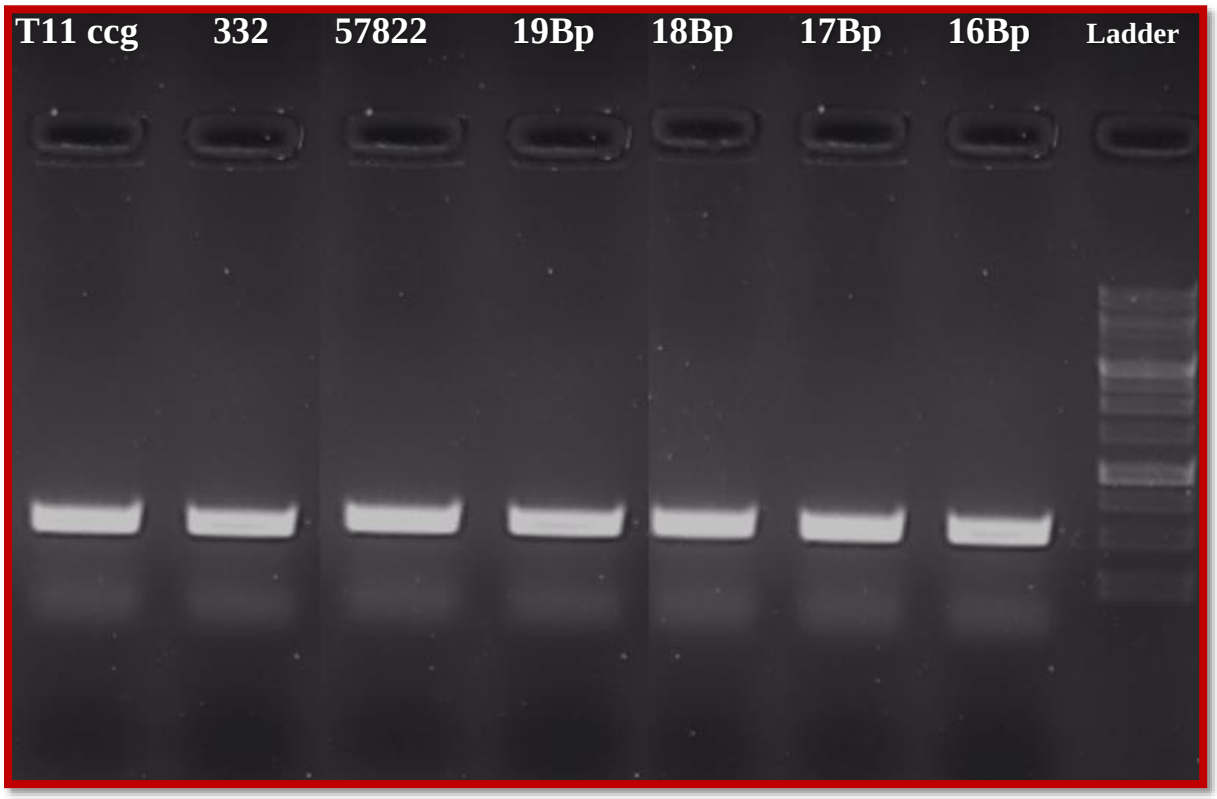


Fig 11: Gel electrophoresis of colony PCR product

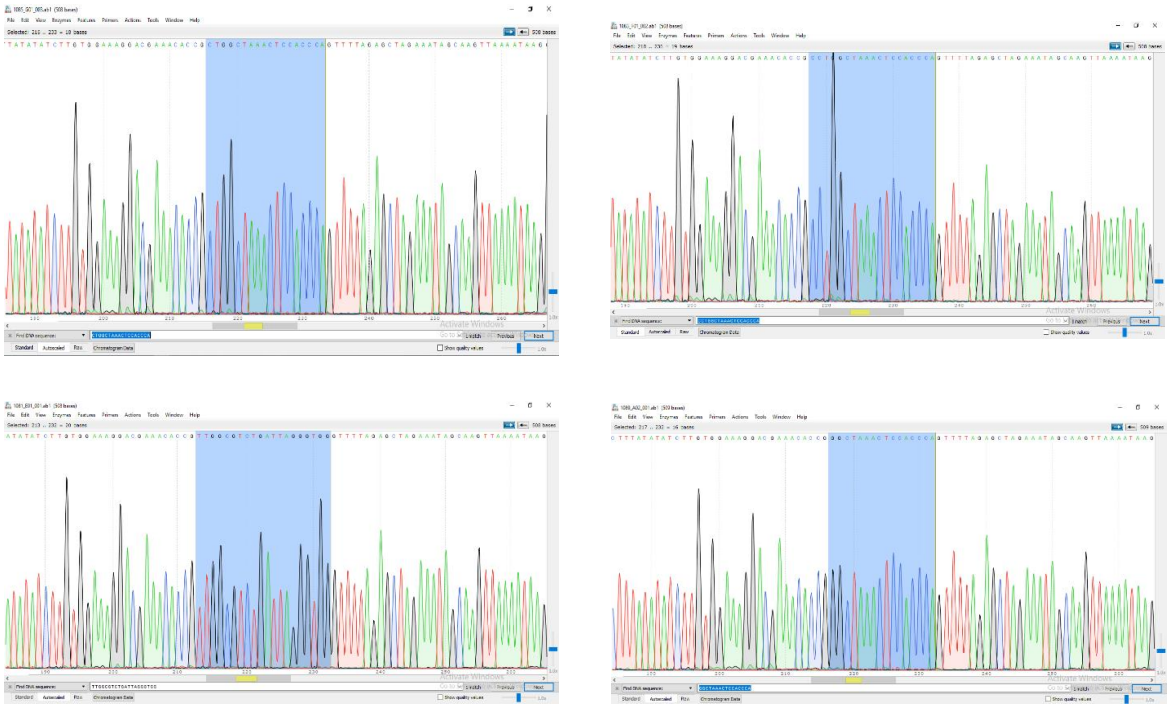


Fig 12: Sanger's sequence

4.1.6 PLASMID DNA EXTRACTION OF POSITIVELY SEQUENCED CLONES USING MINI KIT

The plamid DNA was extracted from positively sequenced clones and the results are given in table 14.

Table 14: Plasmid concentration of Base editing guide clones

CLONE NAME	CONCENTRATION (ng/µL)	260/280	260/230
57822	6800	1.91	2.25
332	564	1.90	2.22
19Bp	605	1.88	2.15
18Bp	873	1.90	2.05
17Bp	772	1.91	2.23
16Bp	621	1.87	2.18
T11 CCG	771	1.90	2.10

Plasmid was isolated from bacterial culture using MN kit. The concentration and degree of purity was obtained as mentioned in the table 5.1. The ratio of approximately 1.8 is generally accepted as “pure” for DNA. The ratio of approximately 2.0 is generally accepted as “pure” for RNA.

4.2 TRANSFECTING THE GENE OF INTEREST VIA LENTIVIRUS

The isolated plasmid was transfected into HEK cells. The 293T cells expressing the green fluorescence when the lentivirus starts to produce and integrating into host chromosome expressing GFP. The fluorescence is captured using Fluorescent Microscopy and the results are depicted in fig 14.

The results showed that the green color indicated that the lentivirus starts to produced and integrated with the host chromosome.

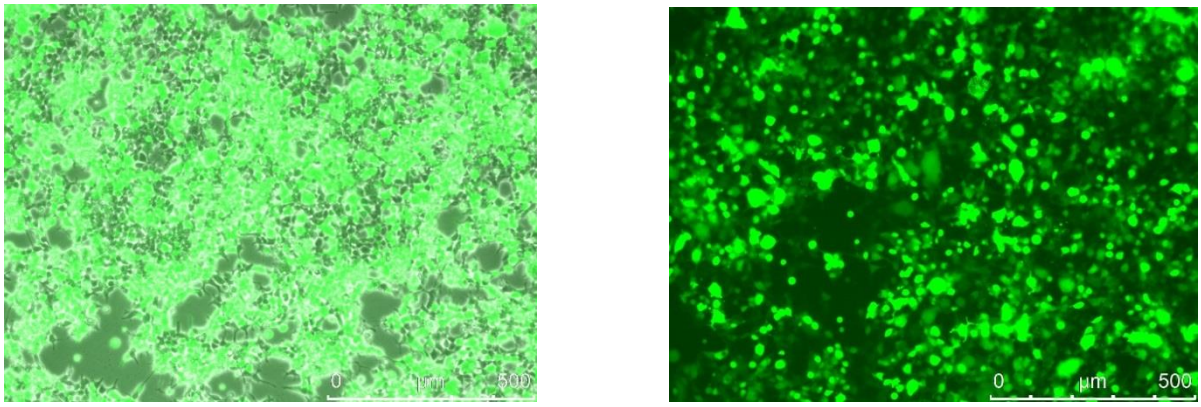


Fig 14: Fluorescent images of transfected HEK 293T Cells

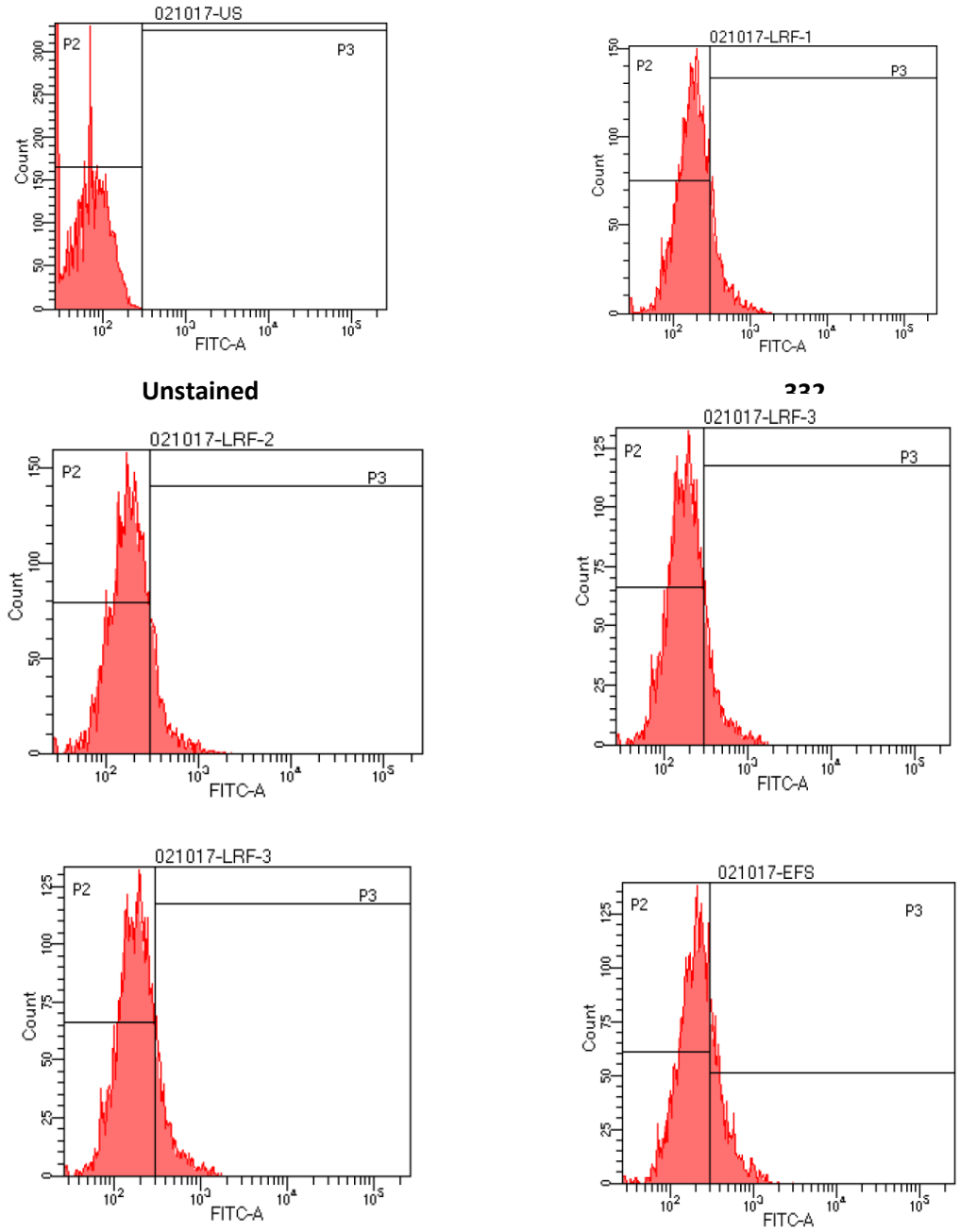


Fig 15: Flow Cytometry Analysis of Base editing guide transfected cell populations

4.2.1 FACS analysis of Transfected cells:

The FACS analysis helps to evaluate the percentage of transfected cells which predominantly expresses GFP in cell. The results of FACS analysis of transfected cells are given in fig 15.

The FACS analysis reveals the percentage of transfected cells which predominantly expresses GFP which is conjugated to FITC. The normal HEK cells which does not express any GFP is used as an Unstained for gating to avoid auto fluorscence. After the analysis, the transduced or GFP-positive cells were sorted from non-transduced cells. These sorted populations contain only the transduced cells which are maintained in culture to produce a stable cell line. These transfected cells showed very low GFP hence the procedure follows virus collection and transduction.

4.3 TRANSDUCTION IN HUDEP-2 STABLE CELL LINE EXPRESSING Cas9

The transduction in HUDEO-2 stable cell line expressing Cas9 is given in plate 2 and 3. The results revealed that the HUDEP 2 cell line showed significant GFP when observed under Phrase contrast microscope.

4.3.1 FACS ANALYSIS OF TRANSDUCED CELLS

Transduced cells were analysed by FACS in different tubes namely unstained tube and tubes containing 20bp, 19bp, 18bp, 17bp and 16bp. The results are given in fig 16, 17,18,19,20 and 21.

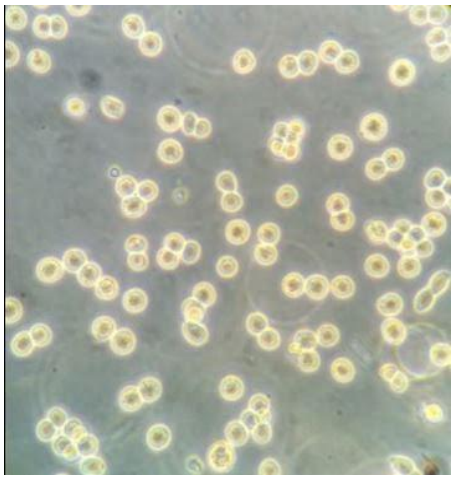


Plate 2: The HUDEP-2 stable cell line expressing the Cas-9

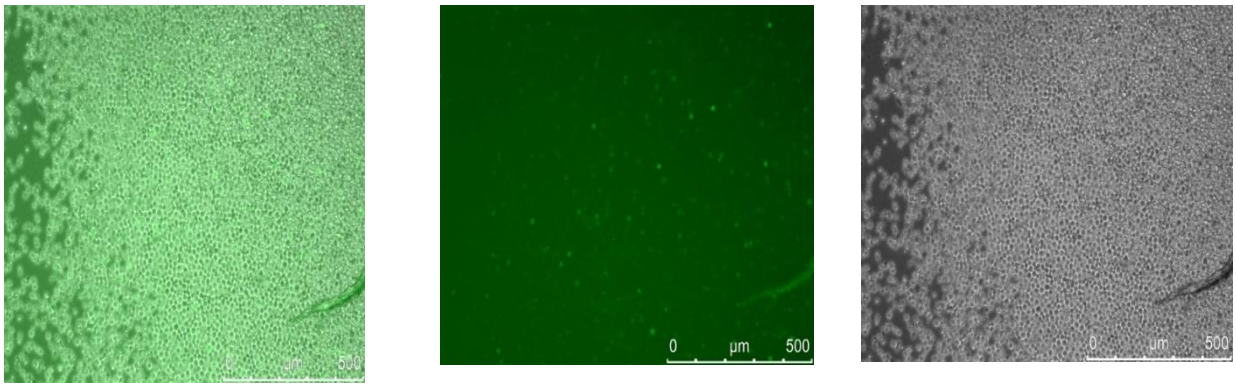
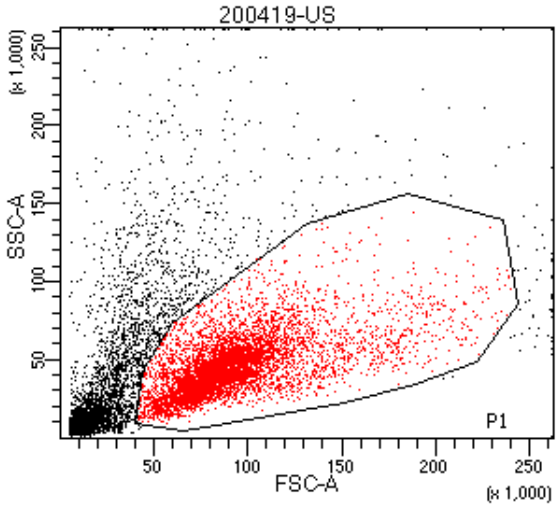
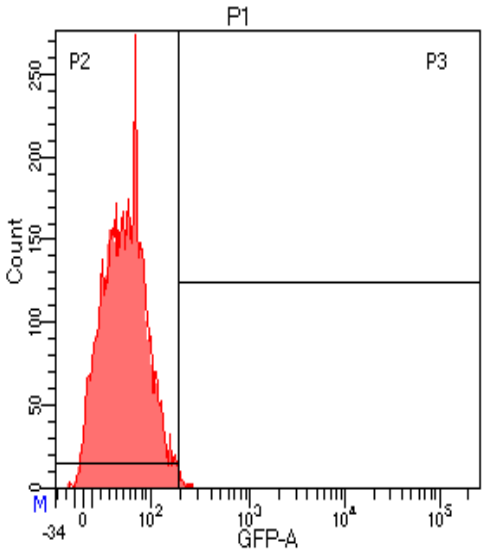


Plate 3: transduced HUDEP 2 cell line expressing CAS 9 stable gene

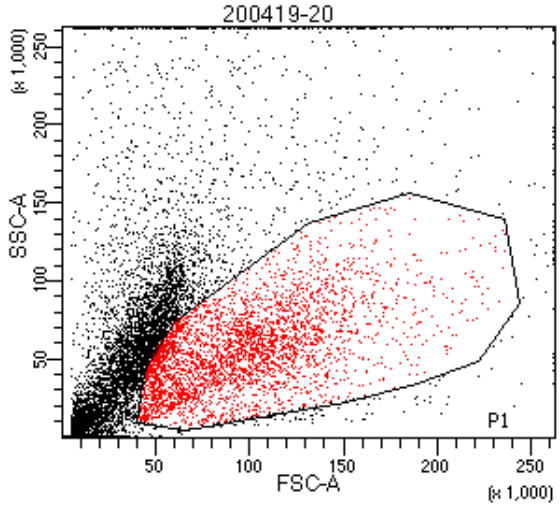


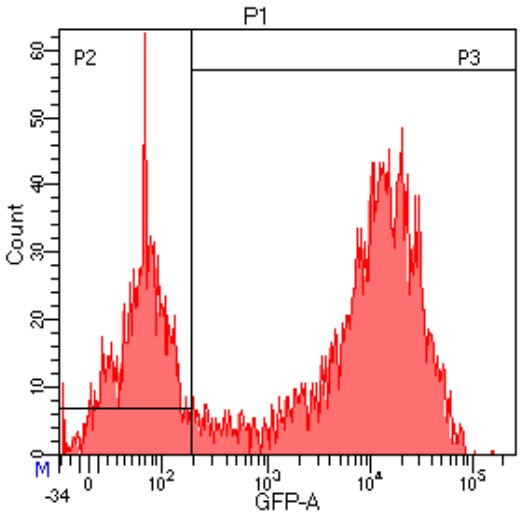


Tube: US

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
P1	6,105	61.1	61.1
P2	6,071	99.4	60.7
P3	34	0.6	0.3
P4	0	0.0	0.0
P5	6,105	100.0	61.1

Fig 16: FACS analysis of unstained tube

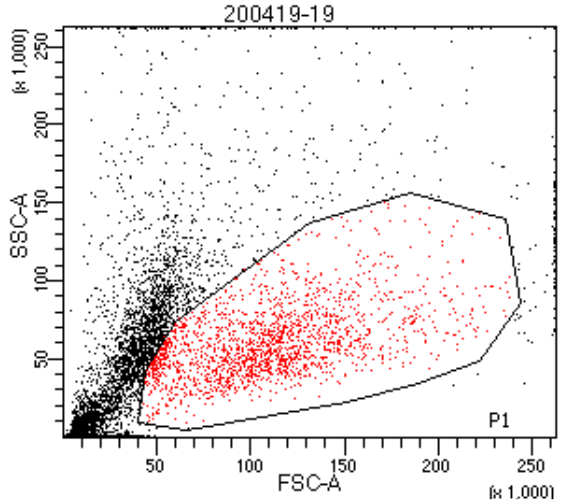


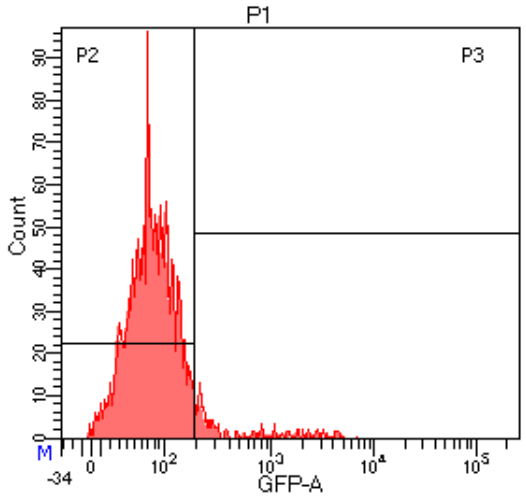


Tube: 20

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
P1	3,376	33.8	33.8
P2	1,071	31.7	10.7
P3	2,305	68.3	23.0
P4	45	1.3	0.4
P5	3,331	98.7	33.3

Fig 17: FACS analysis of 20Bp Tube

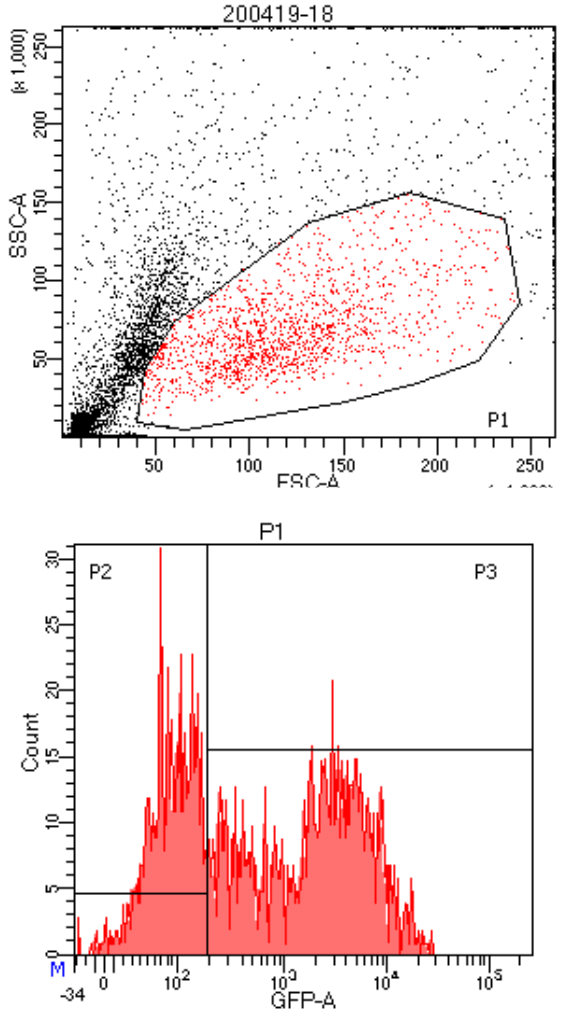




Tube: 19

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
P1	1,882	18.8	18.8
P2	1,721	91.4	17.2
P3	161	8.6	1.6
P4	0	0.0	0.0
P5	1,882	100.0	18.8

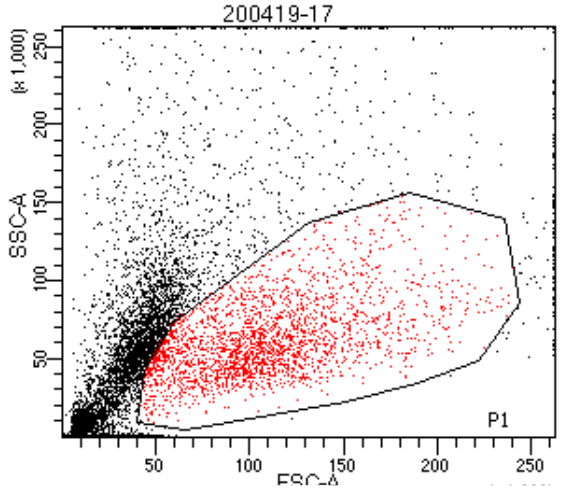
Fig 18: FACS analysis of 19bp tube

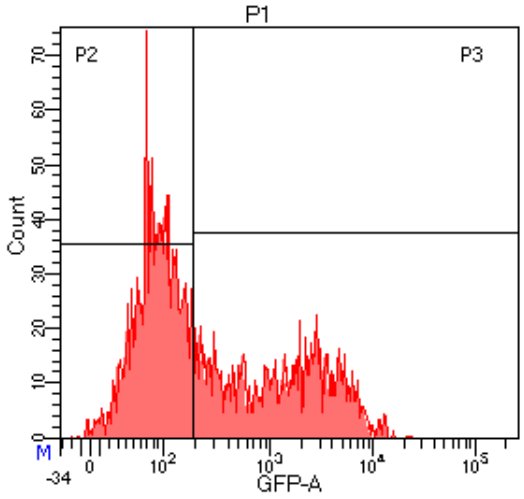


Tube: 18

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
P1	1,535	15.4	15.4
P2	575	37.5	5.8
P3	960	62.5	9.6
P4	1	0.1	0.0
P5	1,534	99.9	15.3

FIGURE 19: FACS analysis of 18 bp tube

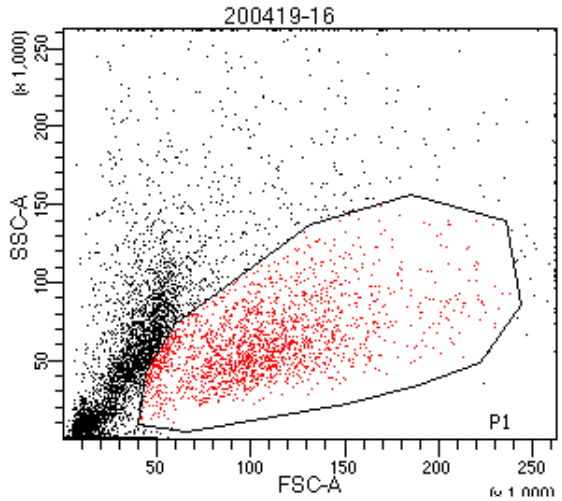


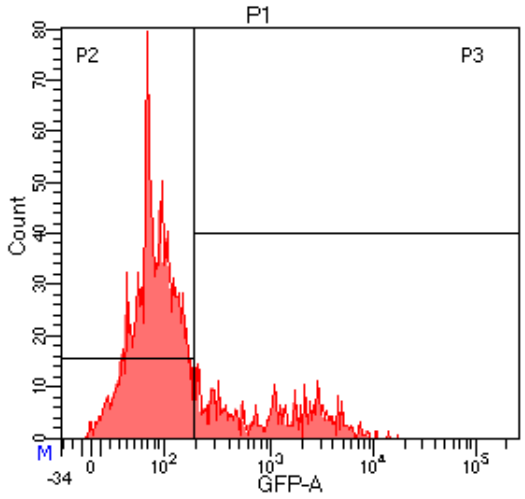


Tube: 17

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
P1	2,428	24.3	24.3
P2	1,313	54.1	13.1
P3	1,115	45.9	11.2
P4	0	0.0	0.0
P5	2,428	100.0	24.3

Fig 20: FACS analysis of 17bp tube





Tube: 16

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
P1	1,866	18.7	18.7
P2	1,382	74.1	13.8
P3	484	25.9	4.8
P4	1	0.1	0.0
P5	1,865	99.9	18.6

Fig 21: FACS analysis of 16bp tube

The transduced cells were taken for FACS analysis after 48 hours. Untransduced cells taken as unstained. 68.3% GFP was observed in 20Bp tube as shown in figure. 8.6% GFP was observed in 19Bp tube as shown in figure. 62.5% GFP was observed in 18Bp tube as shown in figure. 45.9% GFP was observed in 17Bp tube as shown in figure. 25.9% GFP was observed in 16 Bp tube as shown in figure. 13.8% GFP was observed in T11CCG tube as shown in figure. 70% GFP was observed in 57822 tube as shown in figure The observed gfp after transdufction was very low in specificity gene, 19 Bp, 17BP, 16Bp tubses. Significant transduction was observed in 20Bp, 18Bp and 57822 tube.

4.4 GENOMIC DNA ISOLATION

The genomic DNA was isolated from different clones and the data are represented in table 15.

Table 15: Genomic DNA isolation

CLONE NAME	CONCENTRATION (ng/µL)	260/280	260/230
57822	420	1.91	2.2
332	449	1.90	2.1
19Bp	520	1.92	2.2
18Bp	408	1.91	2.1
17Bp	492	1.91	2.1
16Bp	411	1.91	2.1
T11 CCG	442	1.90	2.1

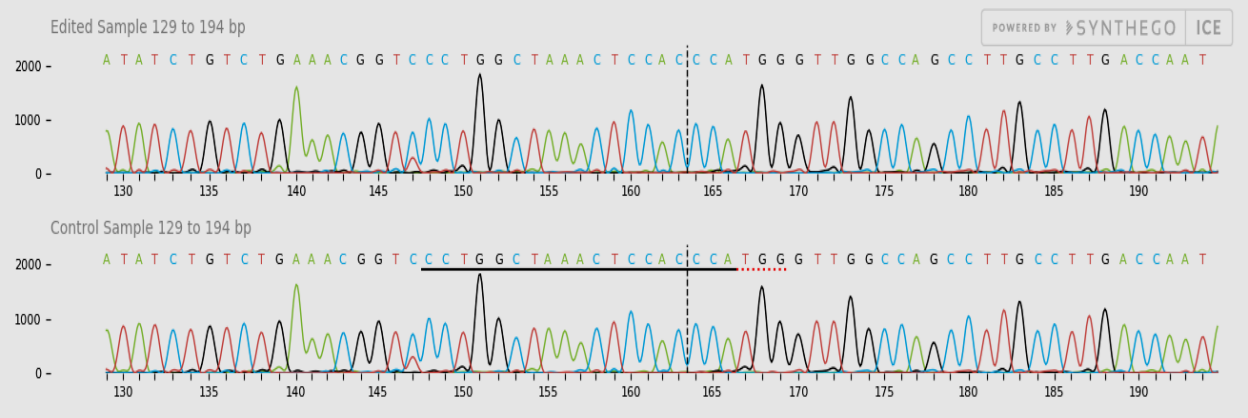
Genomic DNA was isolated using MN kit for the transfected cells. The concentration was obtained using nanodrop. Table shows the concentration and DNA purity of each isolated Genomic DNA.

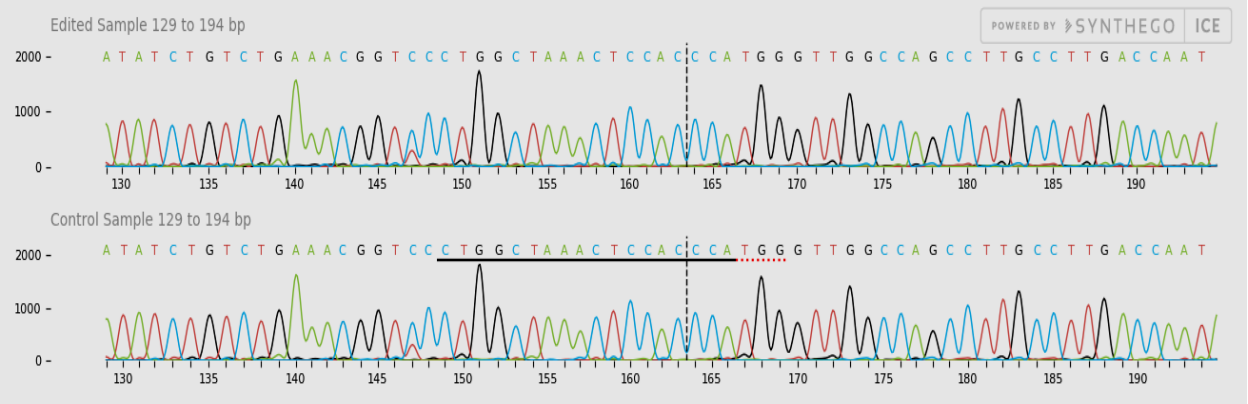
4.4.1 SYNTHEGO ICE ANALYSIS

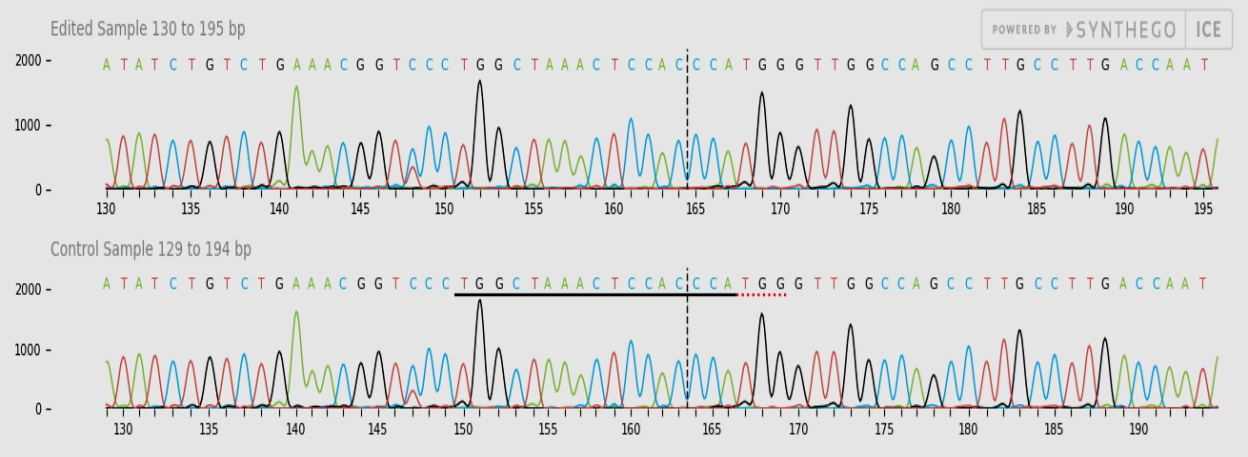
After 48 hours the transduced cells were sorted using FACS. The sorted cells were kept for expansion for 72 hours. After 72 hours, the genomic DNA was isolated and sequencing PCR was set up for the same.

Inference of CRISPR Edits (ICE) is a free and open source software tool that offers fast and reliable analysis of CRISPR editing data. Simply upload your Sanger sequencing data into ICE to view total indel frequency, relative frequency of each indel, a knockout (KO) score that reveals the proportion indels likely to generate complete loss-of-function mutations, and total differences between the edited and wild type sequences.

ICE offers accurate results that correlate strongly with next-generation sequencing (NGS)-based analysis, and is compatible with data generated by both single guide and multiplex editing strategies. This software does not require manually changing setting or parameters, making it incredibly easy to use. The results are depicted in fig 22.







ICE analysis of 20bp, 19bp,18bp,7bp,16bp

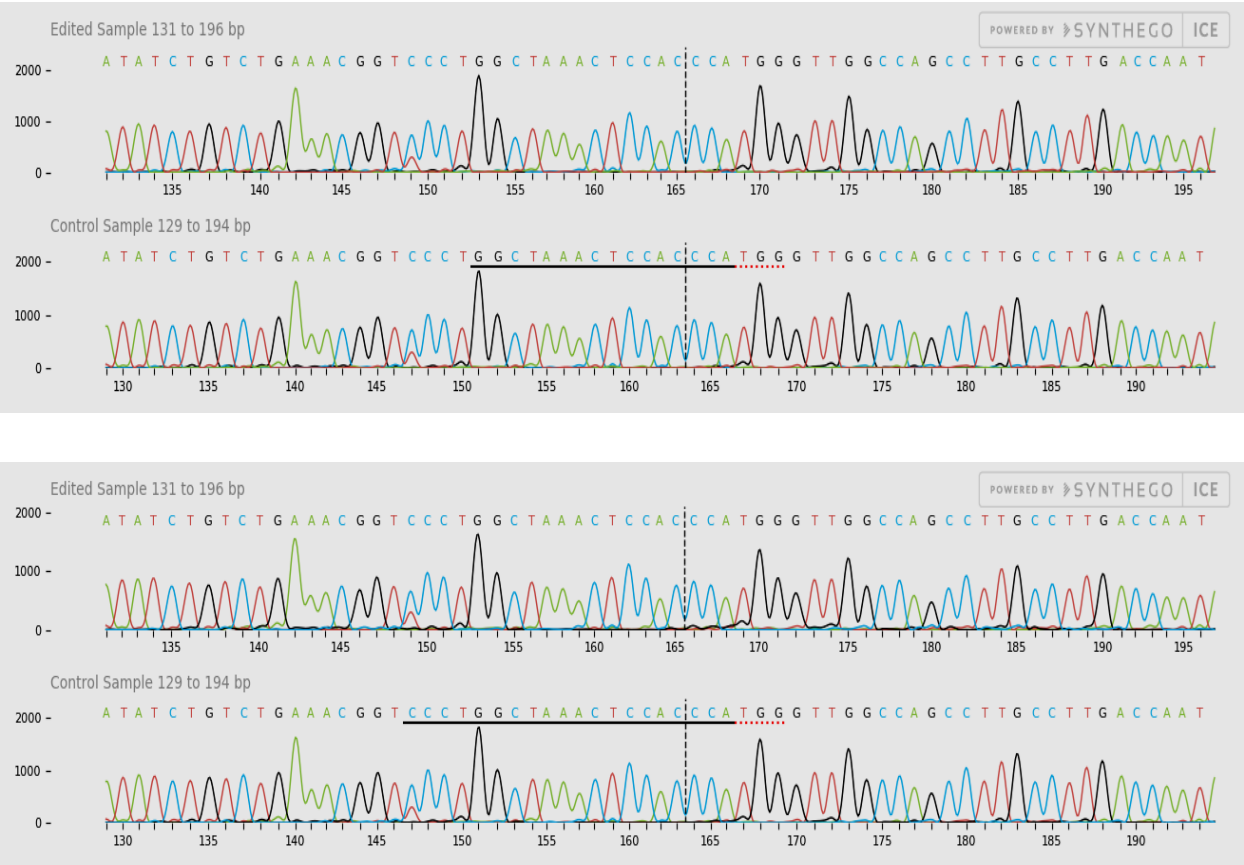


Fig 22: ICE analysis of 20bp, 19bp,18bp,7bp,16bp

To examine the editing efficiency of cytosine deaminase base editor by varying guide length and specificity, the BE16 stables do not show significant C to T conversion in 20bp, 19bp, 18bp, 17bp, and 16bp. similar studies were performed by Li, Z *et al.*, (2018) confirmed that the BE3 system could achieve C-to-T (G-to-A) conversions in cell culture both for individual genes and multiple copies of genes (up to 20). The pig models created via BE3 closely reproduced the phenotypes of human diseases. Besides, a genome-wide analysis showed that early stop codons could be introduced by BE3 in 16,677 pig genes, suggesting a broader range of potential applications of this technology. With the development of new technologies, which have wider PAM compatibility, even more genes could be base edited.

4.5 RNA ISOLATION

The RNA was isolated from different clones and the data are given in table 16.

Table 16: RNA ISOLATION

CLONE NAME	CONCENTRATION (ng/µL)	260/280	260/230
57822	620	1.9	2.2
332	598	1.9	2.2
19Bp	554	1.8	2.1
18Bp	655	1.9	2.0
17Bp	589	1.8	2.2
16Bp	621	1.8	2.1
T11 CCG	675	1.8	2.1

After transduction some of the cells were taken for RNA isolation. RNA was isolated using MN kit and the final concentration was measured using Nanodrop. Table shows the concentration and the RNA purity of each sample.

4.6 ANTIBODY TITRATION

The antibody titration is the process to obtain a necessary cells and reagents. Antibody titration is important to maximize the signal while reducing the noise. This antibody titration is used to minimize the noise due to non-specific binding of the antibodies to low-affinity targets. The antibody titration in flow cytometry were done for CD45 anti-Human PerCP Cy5.5, CD 19 PE, CD19 BV421, CD13 PE, CD 13 PE (MONOCYTES), CD 16 PE and CD 16 BV421 cells to find out the concentration of antibody that gives the brightest signal with lowest background in the flow cytometry. The single cell for each antibody was gated and the positive and negative populations were identified which has been shown in fig 23, 25, 27, 29 and 31. The antibody titrations are graphically represented in fig 24, 26, 28, 30 and 32.

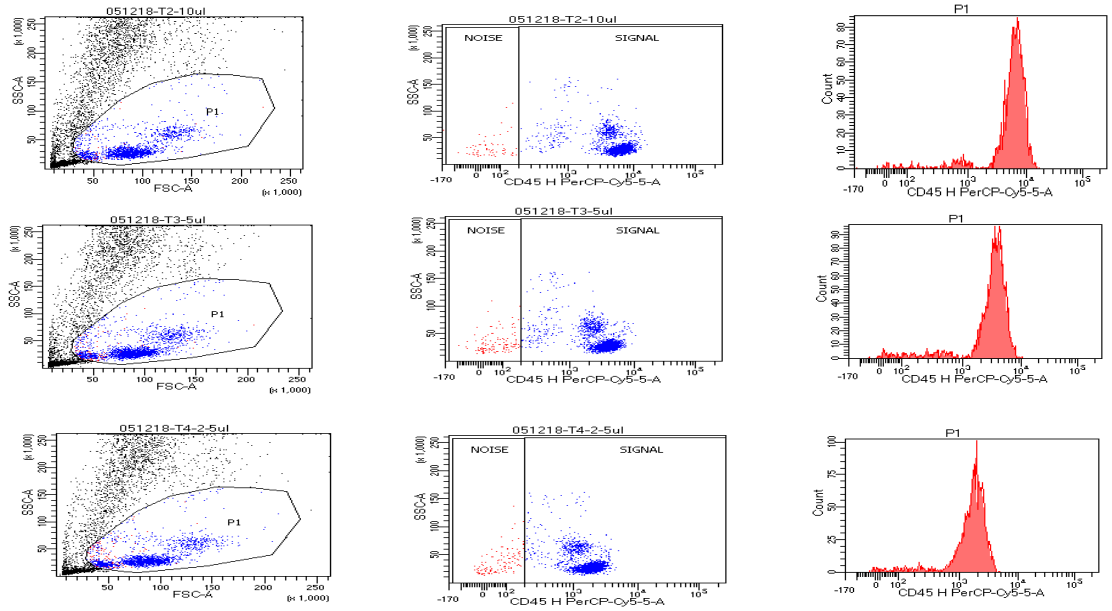


Fig 23: Identification of positive and negative populations in CD45 anti-Human PerCP Cy5.5 antibody by FACS

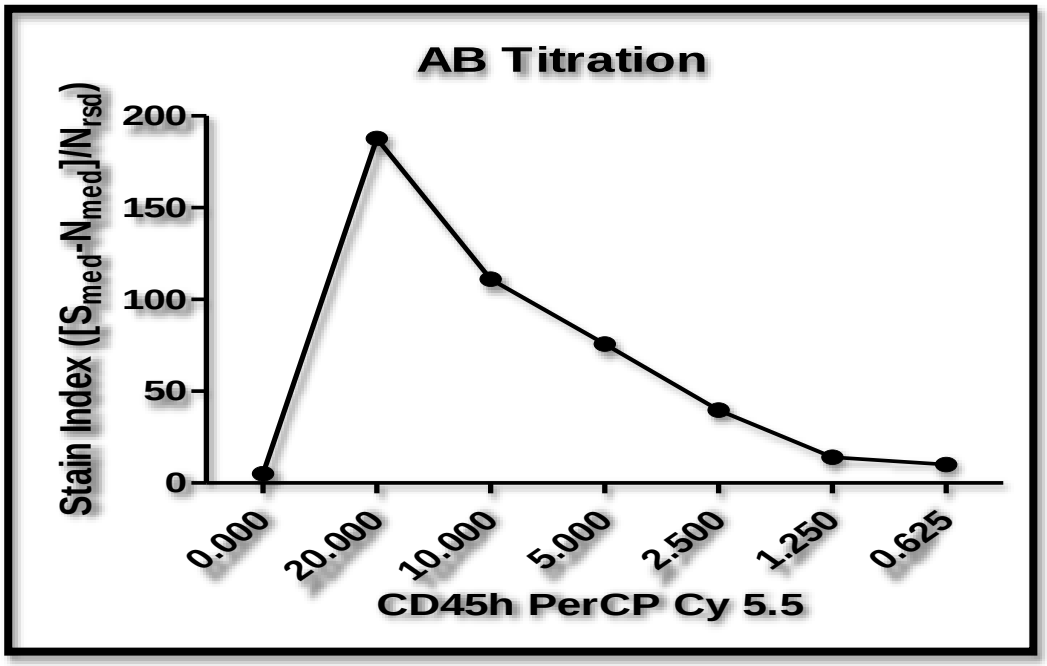


Fig 24: Antibody titration of CD45 anti-Human PerCP Cy5.5 antibody.

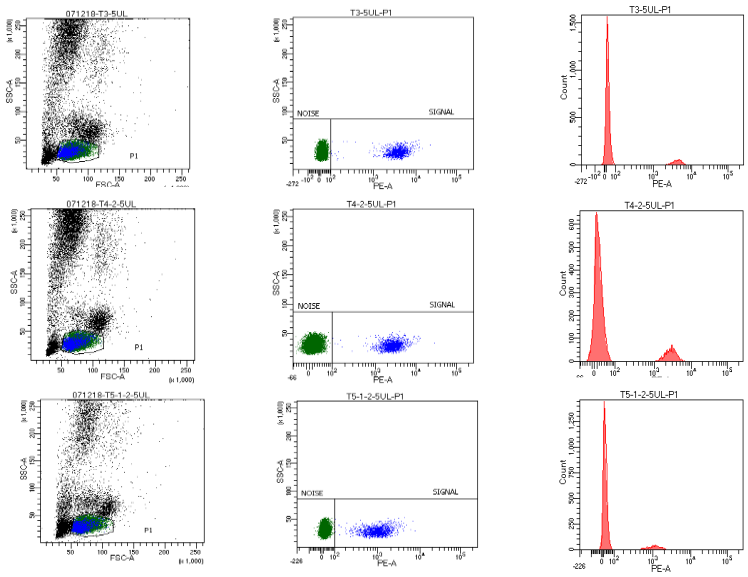


Fig 25: Identification of positive and negative populations in CD 19 PE antibody by FACS

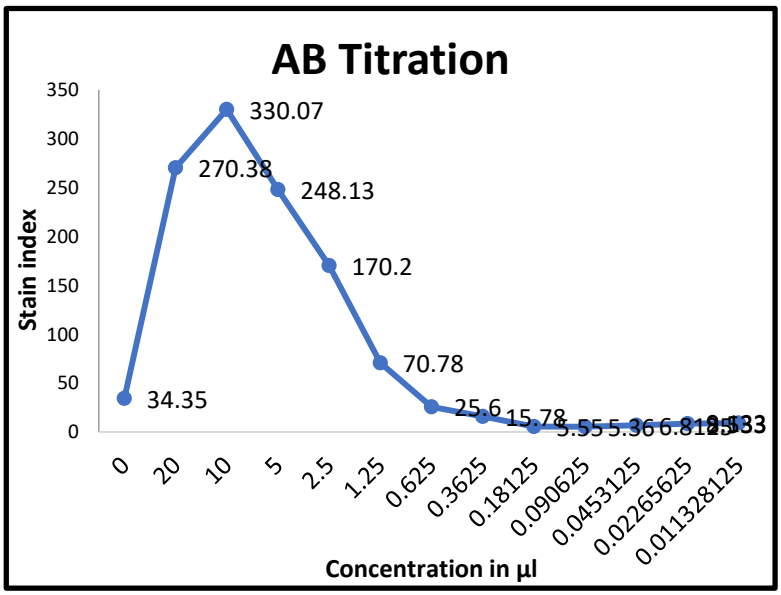


Fig 26: Antibody titration of CD 19 PE antibody

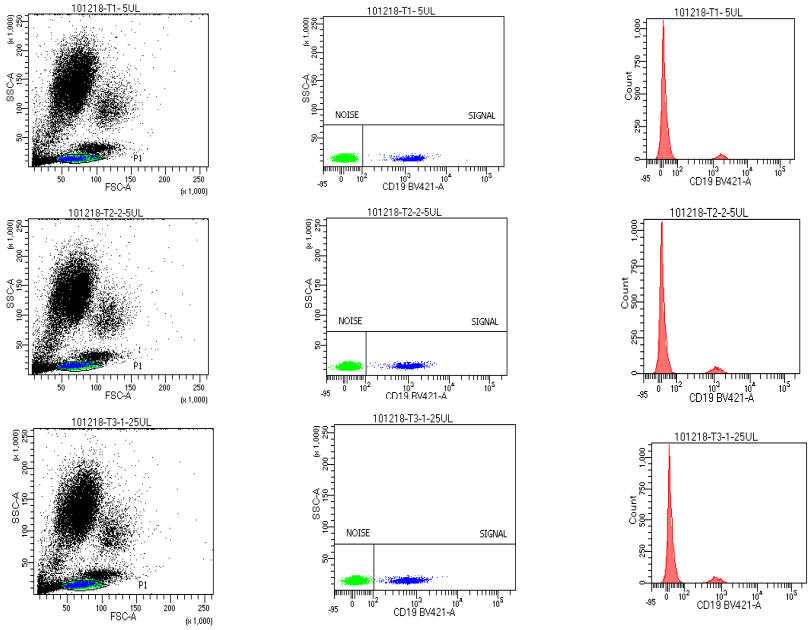


Fig 27: Identification of positive and negative populations in CD19 BV421 antibody by FACS

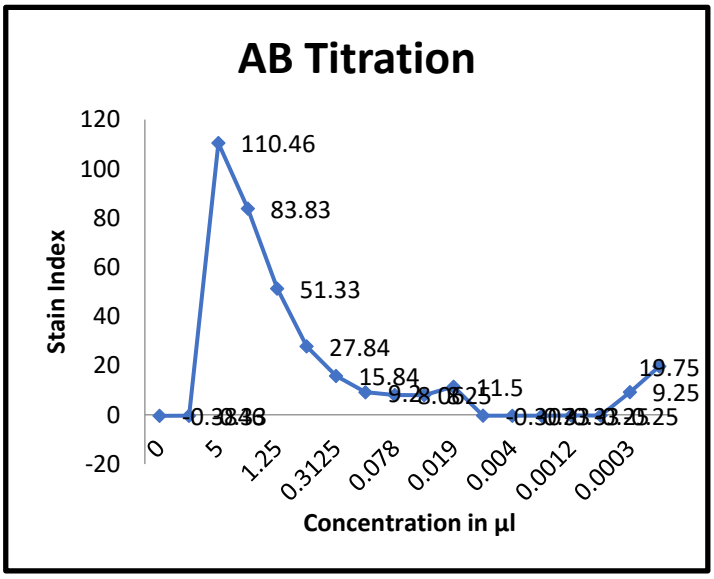


Fig 28: Antribody titration of CD19 BV421 antibody

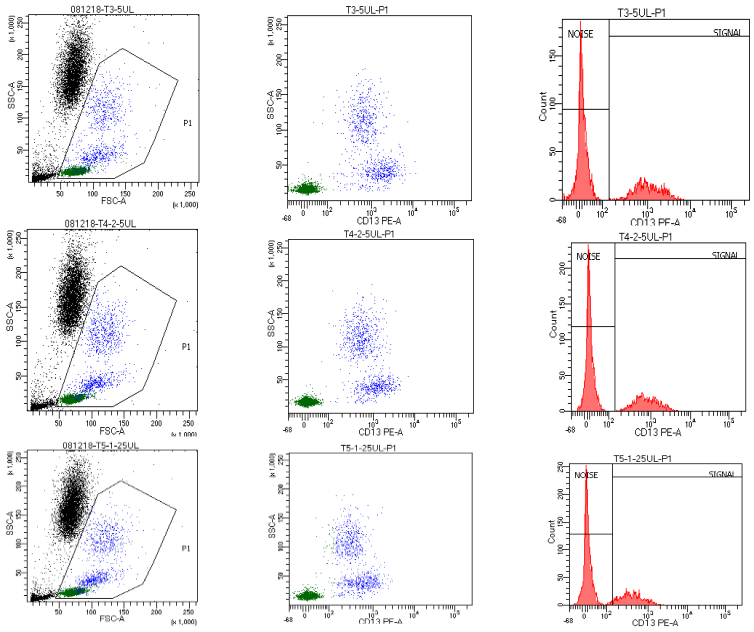


Fig 29: Identification of positive and negative populations in CD13 PE antibody by FACS

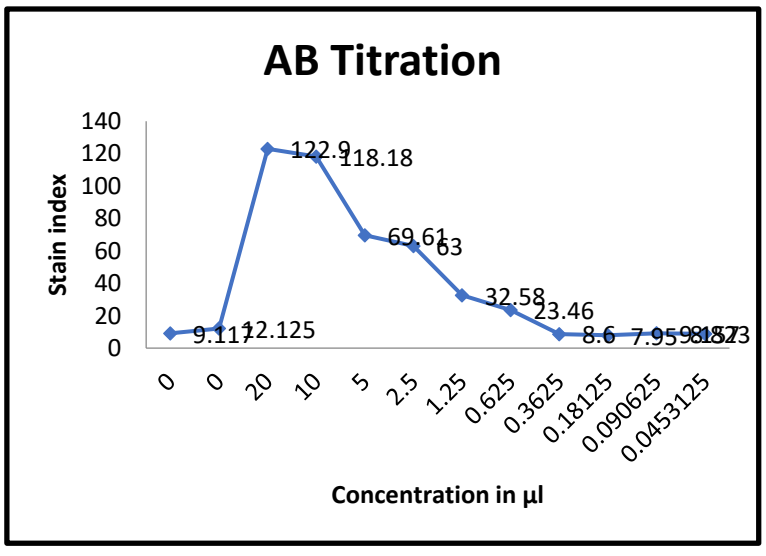


Fig 30: Antibody titration of CD13 PE antibody

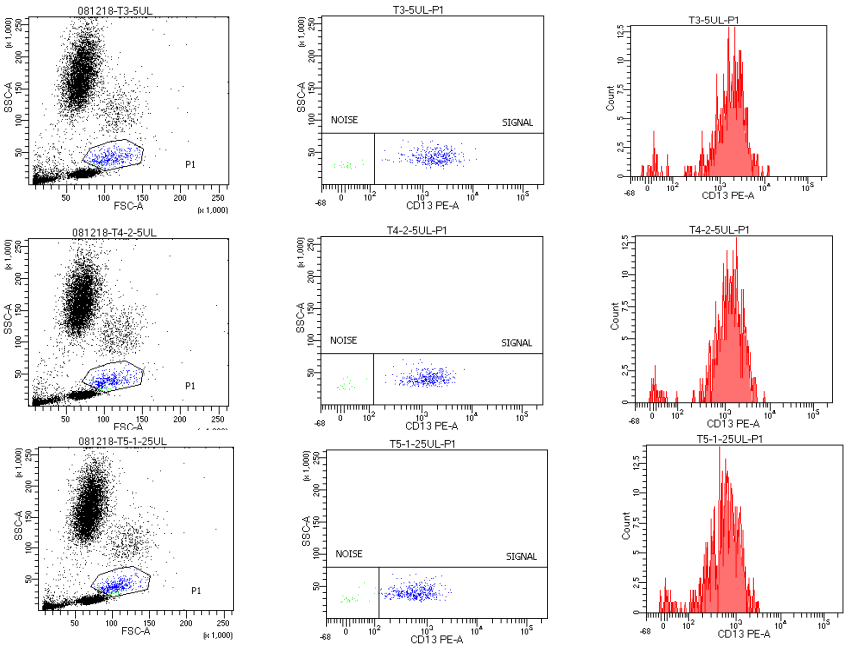


Fig 31: Identification of positive and negative populations in CD 13 PE (Monocytes) in FACS

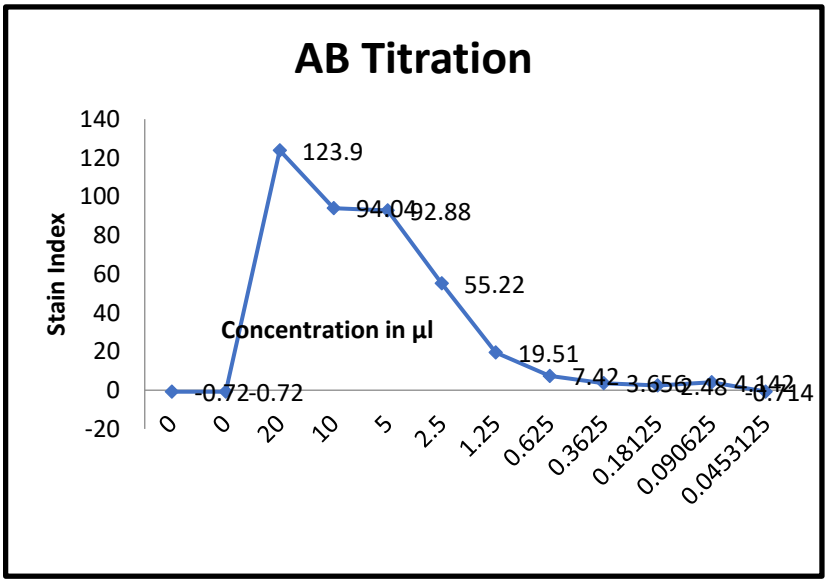


Fig 32: Antibody titration of CD 13 PE (Monocytes)

The results showed that all the concentrations of antibody tested were given a separation between the positive and negative populations. The best concentration of antibody was found to bind with the target 2.5µl/100µl of blood for every cell in the lymphocytes. If there is an excessive amount of antibody was produced it will go and bind to low-affinity targets. This data explained that the increasing the amount of antibody increases staining of the negative population without improving the staining of positive populations.

The antibody titration was carried out by many researchers to identify the best concentration of antibody that produces best separation based on staining index.

Tendulkar *et al*.,2017 have done the antibody titration in group O platelet donors where the maximum titre value for this was found with 128µl among selected group of peoples. The best titre value was obtained in females rather than males.

Rampengan *et al*., 2018 have reported that there is no significance difference of hepatitis B antibody titre value in 1, 2 and 3 months after hepatitis B vaccination.

Another study by Cheang *et al*., 2016 have performed antibody titration in people with hepatitis infection. The study proved that the gender did not associate with seroprotection Hepatitis B and the immune response after primary hepatitis B immunization.

Hatchette *et al*., 2017 have done the antibody titre experiment to screen the large number of specimens for antibodies to measles, mumps, rubella and varicella where thye best titre value was found with the range between ≥0.13 and <1.10AU/ml to achieve 100% specificity.

In another experiment were done by Robbiani *et al*., 2017 determine whether a population infected with ZIKV also displays a range of antibody responses where he proved that the serologic reactivity to dengue 1 virus (DENV1) EDIII before ZIKV exposure is associated with increased ZIKV neutralizing titers of antibody after exposure.

5.0 SUMMARY AND CONCLUSION

The aim of the project has been to highlight the potential of editing efficiency of Cytosine deaminase base editor by varying guide length and specificity. Also to optimize the amount of antibody used for staining of cells for FACS analysis and animal studies. Recent research studies reported that efficient RNA- guided base editing has promising potential in gene therapy.

- The project started with the designing the guides. Retrieval of CDS and using CHOP CHOP to find potential guides based on available PAM sites in the first exon. The primers were annealed with the PCR. For the further cloning the vector has been digested to require backbone for ligation and so on with the help of the BsmB1 restriction enzyme. A clear band was observed during gel electrophoresis. The gel was eluted and the concentration of backbone was measured using spectrophotometer.
- The digested product was ligated with the annealed guides and kept for transformation with the help of B10 competent cells for about 16 hrs. 20 to 30 colonies were observed in each transformed plate.
- The single colonies were picked and inoculated for screening of transformation. The bacterial growth in the LB broth was observed in each test tube after 16 hours of inoculation. Colony PCR was set up and the PCR product was loaded in 1% gel. The bands were observed. Hence transformation of the ligated product into the bacterial cell was confirmed.
- Prior to sequencing pre clean up was done for every samples followed by sequencing PCR and post clean up. The sequence were confirmed with the SNAPGENE application.
- After the sequence confirmation the plasmid was isolated with the help of MN kit. The concentration of isolated plasmids were recorded by means of absorbance. The plasmid was transfected in HEK 293 cells with the help of mirus and packaging plasmids. After 48 hours of transfection GFP were observed under microscope.

- Virus were collected after 48 hours and 72 hours and kept in -80ºC. virus concentarion was done with the help of Lenti X concentrator. 60% - 70% confluent HUDEP BE stables were transduced with concentrated virus. After 48 hours GFP were checked with the help of FACS. 20bp showed better GFP compared to other samples. Then the cells were sorted by single cell sorting procedure using FACS and the cells were kept for expansion with the STEMSPAN media for 72 hours.
- After 72 hours of expansion the genomic DNA was isolated using MN kit (quick extract) and the absorbance were measured with the help of spectrophotometer. RNA was isolated from transduced cells and absorbance were measured. Using SYNTHEGO ICE ANALYSIS, base editing were tested. There was no efficient C to T conversion in the sequence.
- Antibody CD45, CD19 PE, CD19 BV421 and CD13 PE were titrated and analysed with the help of FACS.

BIBLIOGRAPHY

1. Guha, T.K., AlvanWai, and Hausner, G (2017) Programmable Genome Editing Tools and their Regulation for Efficient Genome Engineering, Computational and Structural Biotechnology Journal, 15:146–160.
2. Wong, N., Liu, W., and Wang, X (2015) WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system, Genome Biology:1-8.
3. Xi, L., Schmidt, J.C., Zaug, A.J., Ascarrunz, D.R., and Cech, T.R., (2015) A novel two-step genome editing strategy with CRISPR-Cas9 provides new insights into telomerase action and TERT gene expression, Genome Biology, 16(231): 1-17.
4. Kim, D., Kim, J., Hur, J.K., Been, K.W., Yoon, S.H., and Kim, J,S,. (2016) Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells, Nat Biotechnol, 34(8):863–868.
5. Komor, A.C., Yongjoo, B.K., Packer, M.S., Zuris, J.A., and Liu, D.R., (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature, 533:420–424.
6. Wolfs, J.M., Hamilton, T.A., Lant, J.L., Laforet, M., and Zhang, J., (2016) Biasing genome-editing events toward precise length deletionswith an RNA-guided TevCas9 dual nuclease, Proc Natl Acad Sci U S A.
7. Xu, H., Xiao, T., Chen, C.H., Li, W., Meyer, C.A., and Wu, Q., (2016) Sequence determinants of improved CRISPR sgRNA design, Genome Res, 25:1147–1157.
8. McCaffrey, J., Sibert, J., Zhang, B., Zhang, Y., and Hu, W., (2016) CRISPR-CAS9 D10A nickase target-specific fluorescent labeling of double strand DNA for whole genome mapping and structural variation analysis, Nucleic Acids Res, 44(2).
9. Varshney, G.K., and Burgess, S.M., (2016) DNA-guided genome editing using structure-guided endonucleases, Genome Biology, 17(1):187-194.
10. Barrangou, R., (2015) Diversity of CRISPR-Cas immune systems and molecular machines, Genome Biology, 16(247):1-11.
11. Wang, L., Li, F., Dang, L., Liang, C., and Wang, C., (2016) In vivo delivery systems for therapeutic genome editing, International Jouranal of Molecular Sciences, 17(5).

12. Algasaier, S.I., Exell, J.C., Bennet, I.A., Thompson, M.J., and Gotham, V.J., (2016) DNA and protein requirements for substrate conformational changes necessary for human flap endonuclease-1-catalyzed reaction, J Biol Chem, 291(15):8258–8268.

13. Mino, T., Mori, T., Aoyama, Y., and Sera, T,. (2014) Inhibition of DNA replication of human papillomavirus by using zinc finger-single-chain FokI dimer hybrid, Mol Biotechnol, 56(8):731–737.

14. Porteus,M.H., (2015) Towards a new era in medicine: therapeutic genome editing, Genome Biology, 16(286):1-12.

15. Fagerlund, R.D., Staals, R.H.J., and Fineran, P.C., (2015) The Cpf1 CRISPR-Cas protein expands genome-editing tools, Genome Biology, 16(251):1-3.

16. Zhang, Y., Massel, K., Godwin, L.D., and Gao, C., (2018) Applications and potential of genome editing in crop improvement, Genome Biology, 19(210): 1-11.

17. Singer, M., Marshall, J., Heiss, K., Mair, G. R., Grimm, D., Mueller, A., and Frischknecht, F., (2015) Zinc finger nuclease-based double-strand breaks attenuate malaria parasites and reveal rare microhomology-mediated end joining, Genome Biology, 16(249): 1-18.

18. Carroll, D., and Charo, R.A., (2015) The societal opportunities and challenges of genome editing, Genome Biology, 16(242): 1-9.

19. Yao, Y., (2018) Genome editing: from tools to biological insights, Genome Biology, 19(186):1-2.

20. Han, R., Li, L., Ugalde, A.P., Tal, A., Manber, Z., Barbera, E.P., Chiara, V.D., Elkon, R., and Agami, R., (2018) Functional CRISPR screen identifies AP1-associated enhancer regulating FOXF1 to modulate oncogene-induced senescence, Genome Biology, 19(118): 1-13.

21. Chari, R., Mali, P., Moosburner, M., and Church, G.M., (2015) Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach, Nat Methods, 12:823–826.

22. Guilinger, J.P., Thompson, D.B., and Liu, D.R., (2014) Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification, Nat Biotechnol, 32(6):577–582.

23. Sander, J.D., and Joung, J.K., (2014) CRISPR-Cas systems for editing, regulating and targeting genomes, Nat Biotechnol, 32:347–355.

24. Achilli, C., Ciana, A., Minetti, G., (2018) Immortalized HEK 293 Kidney Cell Lines as Models of Renal Cells: Friends or Foes? Journal of Controversies in Biomedical Research, 4(1): 6–9.

25. Sah, J.P., Salahuddin, M., Sayed, M.A., Salauddin, A.S.A., Sayfullah, M., Rashid, M., Saclain, S., Washim, M.R., Zeb, M.A., Hossain, D., Chowdhury, R., and Momin, M.A., (2015) Transfection optimisation in HEK-293 cell line, Asian Journal of Medical and Biological Research, 1 (1):121-129.

26. Burkard, C.T., Wilson, A.D., McGrew, M.J., Archibald, A.L., Sang, H.M., Houston, R.D., Whitelaw, C.B., and Watson, M., (2018) Livestock 2.0 – genome editing for fitter, healthier, and more productive farmed animals, Genome Biology, 19(204):1-11.

27. Duan, J., Lu, G., Hong, Y., Hu, Q., Mai, X., and Guo, J., (2018) Live imaging and tracking of genome regions in CRISPR/dCas9 knock-in mice, Genome Biology:19.

28. Kuscu, C., Parlak, M., Tufan, T., Yang, J., Szlachta, K., and Wei, X., (2017) CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations, Nat Methods. 14:710–712.

29. Butt, H., Eid, A., Ali, Z., Atia, M.A.M., Mokhtar, M.M., and Hassan, N., (2017) Efficient CRISPR/Cas9-mediated genome editing using a chimeric single-guide RNA molecule, Front Plant Sci, 8:1441.

30. Doudna, J.A., and Gersbach, C.A., Genome editing: the end of the beginning, Genome Biology, 16(292):1-3.

31. Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., and Kabadi, A.M., (2015) Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements, Nat Methods, 12(12):1143–1149.

32. Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A., (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems, Nat Biotechnol, 31(3):233–239.

33. Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., and Hummel, A., (2010) Targeting DNA double-strand breaks with TAL effector nucleases, Genetics, 186:757–761.

34. Hilton, I.B., and Gersbach C.A., (2015) Enabling functional genomics with genome engineering, Genome Res, 25(10):1442–1455.

35. Joung, J.K., Voytas, D.F., and Kamens, J., (2015) Accelerating research through reagent repositories: the genome editing example, Genome Biology,16:255.

36. Robbiani, D,F., Bozzacco, L., Keeffe, J.R., Khouri, R., Olsen, P.C., Gazumyan, P., Babajew, D.S., Rios, S.A.., Nogueira, L., Patel, R., Azzopardi, S.A., Uhl, L.F.K., Saeed, M., Reyes, E.E.S., Agudelo, M., Yao, K.H., Golijanin, J., Gristick, H.B., Lee,Y.E., Hurley, A., Caskey, M., Pai, J., Oliveira, T., Wunder, E.A., Sacramento, G., Nery, N., Orge, C., Costa,F., Reis, M.G., Thomas,N.D., Eisenreich, T., Weinberger, D.M., Almeida, R.P., West, A.P., Rice, C.M., Bjorkman, P.J., Teran, G.R., Ko, A.I., MacDonald, M.R. and Nussenzweig, M.C. (2017) Recurrent Potent Human Neutralizing Antibodies to Zika Virus in Brazil and Mexico, Cell, 169:597–609.

37. Cheang, H.K., Wong, H.T. and Ho, S.C. (2016) Immune response in infants after universal hepatitis B vaccination: A community based study World Health Organization, Vaccines, 28:589-592.

38. Tendulkar, A.A., Jain, P.A. and Velaye, S. (2017) Antibody titres in group O platelet donors, Asian Journal of Transfusion Sciences, 1-14.

39. Rampengan, N.H., Yustianto, A., Warouw, S., Wilar, R. and Salendu, P. (2018) The average of HBS antibody titre in different months after hepatitis B premier vaccination, Current Pediatric Research, 22(2):1-3.

40. Hatchette, T.F., Scholz, H., Bolotin, S., Crowcroft, N.S., Jackson, C., McLachlan, E. and Severini, A. (2017) Calibration and evaluation of quantitative antibody titres for measles virus by using BioPlex 2200, Clinical vaccines and immunology, 24(1):1-9.

41. Fomenkov, A., Sun, Z., Dila, D.K., Anton, B.P., Roberts, R.J. and Raleigh, E.A. (2017) EcoBLMcrX, a classical modification-dependent restriction enzyme in *Escherichia coli* B: Characterization *in vivo* and *in vitro* with a new approach to cleavage site determination, PLOS one, 1-21.

42. Tsai, R., Correa, I.R., Xu, M.Y. and Xu, S.Y. (2017) Restriction and modification of deoxyarcheosine (dg+)-containing phage 9 g DNA, Scientific Reports, 20-26.

43. Anand, A., Bass, S.H., Wu, E., Wang, N., McBride, K.E., Annaluru, N., Miller, M., Hua, M. and Jones, J.T. (2018) An improved ternary vector system for *Agrobacterium* – mediated rapid maize transformantion, Plant molecular biology, 97:187-200.

44. Miyango, M., Tanaka, K., Ishikawa, S., Mori, K., M., Arribas, A.M., Meijer, W.J.J. and Yoshida, K.I. (2018) A novel method for transforming the thermophilic bacterium *Geobacillus kaustophilus*, Microbial cell factories, 17(27):1-17.

45. Hasegawa, H., Suzuki, E. and Maeda, S. (2018) Horizontal plasmid transfer by transformation in *Escherichia coli* Environmental factors and possible mechanisms, Frontiers in microbiology, 9(2365):1-6.

46. Huang, F., Spangler, J.R. and Huang, A.Y. (2017) In vivo cloning of up to 16kb plasmids in *E.coli* is as simple as PCR, PPLOS ONE, 12(8):1-15.

47. Jamal, M.A.H.M., Sharma, S.P., Chung, H.J., Kim, H.J., Hong, S.T. and Lee, S. (2017) Ultra-High efficiency colonyPCR for high throughput screening of bacterial genes, Indian journal of microbiology, 57(3):365-369.