

Construction and Optimization of a ssDNA Vector Production System for Gene Therapy Platform

by

Rohini Prakash

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Biology

Waterloo, Ontario, Canada, 2023

© Rohini Prakash 2023

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of this thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Abstract

Gene therapy (GT) holds immense potential to treat and/or cure a myriad of diseases such as cancer, immunodeficiency, cardiovascular diseases, and more without involving pharmacological intervention, radiotherapy, or surgery. The success of GT however, depends greatly on the vector used in delivery. An ideal vector should confer high transfection efficiency, high transgene expression, should be stable, well-tolerated, cost-effective, and scalable. Both viral and non-viral vectors have demonstrated success with variable safety/efficacy profiles. However, each vector has its associated disadvantages. The use of many current viral vectors may lead to insertional mutagenesis, immune/proinflammatory responses, and are difficult/expensive to manufacture. On the other hand, non-viral vectors are comparatively safer and can be manufactured easily but have lower transfection efficiency and less effective transgene expression. Due to the associated disadvantages with each type of vector, there is a need for constructing and developing newer efficient vectors. Among other novel strategies, bacteriophages as delivery vehicles for genetic cargo present a current approach under investigation. Bacteriophage-mediated gene transfer is now achievable due to the adaptability of these entities that may provide a safer and scalable option for gene therapy since they can be engineered to deliver genetic material into mammalian cells and tissues in a targeted manner via phage display technology. Filamentous (Ff) phages such as M13 have demonstrated the potential to be used as a targeted GT vector. The present study focuses on the construction and optimization of a ssDNA phage replicated vector carrying a eukaryotic gene cassette for gene therapy. Using the principle for Ff phage replication and packaging of ssDNA in vivo in the *E. coli* host, the principle will be extended to encapsidation of phage-replicated ssDNA by the eukaryotic anellovirus capsid, also to be produced via an *E. coli* protein production system. This combination offers unexplored

territory for the packaging and production of anelloviruses and their subsequent potential as a functional GT delivery vector. Like Ff phages, anelloviruses are similarly ssDNA viruses, but with smaller genome sizes of between 2.0 and 3.8 kb. These viruses have not been associated with any diseases in humans and are found in high abundance in various cell types and tissues making them intriguing as non-immunogenic, potentially redosable GT vectors. The ability of these viruses to transfect different cell types and tissues and the potential to travel via exosomes makes them suitable and safer candidates for gene delivery, compared to other viral vectors currently employed. Although anelloviruses may prove to be an ideal vector, research in the field is very limited at present and there are no cell culture systems for growing anelloviruses in vitro. Thus, newer approaches for production need to be developed. This project aims to construct and preliminarily characterize an *E. coli*-based anellovirus-phagemid hybrid vector production system for use in GT applications.

Using an *E. coli* production system, TTV19 anellovirus capsid proteins were successfully expressed and purified. The ssDNA to be packaged by anellovirus capsid proteins was produced using the M13 replication system in tandem via a novel miniphagemid processing technology. The anellovirus capsid proteins were able to assemble ssDNA producing viral particles ex vivo in a cell extract system when combined, which was demonstrated using transmission electron microscopy, and real-time PCR. Overall, this approach offers promise in the production of newer, safer, and cost-effective viral vectors for gene therapy platforms.

Acknowledgements

First and foremost, I am extremely grateful to my supervisor Dr. Roderick Slavcev for his invaluable advice, continuous support, and his dedicated involvement in every step throughout the process. He has provided me with numerous opportunities that have helped me to grow academically and professionally and have been an excellent mentor. I would like to thank the members of my advisory committee, Dr. Marc Aucoin, Dr. Trevor Charles, and Dr. Valerie Ward, for their valuable input, and support. The work in this thesis was made possible by the University of Waterloo.

Without the assistance of a great number of people, I could not have completed this thesis, and I would like to thank them. I am incredibly grateful to Dr. Marc Aucoin for his invaluable input and guidance that helped shape the work in this thesis. I am also grateful to the entire Aucoin lab (Madhuja Chakraborty, Jacqueline Powichrowski, Scott Boegel, and Lisa Nielson) for the help and support they have given me. I am grateful to Dr. Nafiseh Nafissi for her advice and guidance. I am extremely grateful to the School of Pharmacy and Monica Tudorancea for all her help with the School of Pharmacy Core Research Facility equipment. I am grateful to Nekkar lab for providing me with TEM grids. I would like to thank Ho lab for providing chemicals for my experiments. I am grateful to Jin Wang for helping me with qPCR. I would like to thank Mishu Groh for transmission electron microscope training I would like to express my gratitude to the Department of Biology, Dr. Kirsten Muller, Dr. Moira Glerum, Dr. Mungo Marsden, Dr. Cheryl Duxbury, April Wettig, and Katelyn Fehrenbach for their continuous help and support over the years.

I would like to express my gratitude to those who helped me on an academic level as well as a personal level. I am grateful to Dr. Salma Jiminez for her guidance during my initial days in the lab. I am extremely grateful to Dr. Shirley Wong for her help and guidance throughout

my research journey. I am grateful to Jesse St. Jean for his constant help and motivation that helped me grow both professionally and personally. I am extremely grateful to have a friend like Deborah Pushparajah who taught me so many things and has been a great friend who stood by me through the thick and thins of my journey and help me grow academically and on a personal level. I would like to express my gratitude to the entire Slavcev lab (Dr. Salma Jiminez, Dr. Shirley Wong, Jesse St. Jean, Heba Alattas, Deborah Pusparajah, Merium Fernando, Andrew Ng, Harika Nagareddy, Nicholas Mourad, Jiayang Li, Nicholas Chen, Talia Fiaani) for their help, friendship, laughs, and beautiful memories.

I am grateful to the almighty for providing me with great opportunities in life. I have my deepest gratitude and appreciation to my parents (Nava Prakash Singh and Bina Singh) for their unconditional love and support. I am grateful to my siblings (Vishakha Singh and Rohit Prakash) for their support especially my sister who has been my rock throughout the journey. I am grateful to my brother-in-law (Deepak Kumar) who motivated me to go for research studies and my nephew (Vrishank Singh) whom I have missed a lot through the years. I am grateful to my friends (Mohita Chaudhary, Mahima Chaudhary, Bhavani Bandhu, Chitra Ramachandran, Deborah Pushparajah, Kaelan Horn, Harsha Rohila and Matilda D'souza) for their love and support.

Dedication

This thesis is dedicated to my parents, Nava Prakash Singh and Bina Singh, for their unconditional love and support.

Table of Contents

List of Figures.....	x
List of Tables	xii
List of Abbreviations	xiii
1. INTRODUCTION.....	1
1.1 GENE THERAPY.....	3
1.1.1 GENE THERAPY SYSTEMS	7
1.1.2 NON-VIRAL VECTORS vs VIRAL VECTORS	13
1.2 ssDNA EUCARYOTIC VIRUSES	33
1.3 ANELLOVIRUS	36
1.3.1 GENOME REPLICATION AND TRANSCRIPTION OF ANELLOVIRUS	38
1.3.2 POTENTIAL INTERACTION OF ANELLOVIRUS WITH THE HOST IMMUNE SYSTEM.....	41
1.3.3 ANELLOVIRUS AS A VECTOR FOR GENE THERAPY	43
1.4 FILAMENTOUS PHAGES	44
1.4.1 LIFE CYCLE OF M-13 FILAMENTOUS (Ff) PHAGE	46
1.4.2 USE OF FILAMENTOUS PHAGES FOR GENE THERAPY	50
1.4.3 ADVANTAGES AND LIMITATIONS OF FILAMENTOUS PHAGES APPLICATION TO GENE THERAPY	52
2. RATIONALE, HYPOTHESIS, AND OBJECTIVES	54
2.1 RATIONALE	54
2.2 HYPOTHESIS.....	55
2.3 OBJECTIVES	56
3. MATERIALS AND METHODS	57
3.1. Strains, vectors, and primers.	57
3.2. Plasmid design and construction.	57
3.3. Transformation of <i>E. coli</i> cells with plasmids.	60
3.4. Protein production and extraction.	60
3.5. SDS-PAGE and Western Blot.....	60
3.6. Protein Purification.	62
3.7. Protein quantification.....	63
3.8. Phagemid production.	63
3.9. Purification of phagemid.....	63
3.10. Quantification of phagemid through plaque assay.....	64
3.11. Phenol-chloroform extraction.....	64

3.12. <i>In vitro</i> assembly of viruses.	65
3.13. DNase enzymatic digestion assay.....	65
3.14. Iodixanol ultracentrifugation.....	65
3.15. Transmission Electron Microscopy (TEM).....	66
3.16. Polymerase chain reaction (PCR).....	66
3.17. Quantification of virion particles.	67
4. RESULTS	69
4.1. Selection of a suitable anellovirus candidate for capsid protein production and <i>in vitro</i> assembly of ssDNA.	69
4.2. Production of TTV19 capsid protein for <i>in vitro</i> virus assembly.	70
4.2.1. Production of TTV19 capsid protein.	70
4.2.2. Optimization of TTV19 Capsid protein production.	71
4.2.3. Purification of expressed TTV19 capsid protein.....	72
4.3. Production of ssDNA containing the TTV19 DNA binding signal using M13 replication system.....	76
4.3.1. Production, purification, and quantification of M13 phagemid containing the TTV19 DNA binding signal.....	76
4.3.2. Extraction of ssDNA from M13 phagemid for <i>in vitro</i> assembly.	77
4.4. <i>In vitro</i> assembly of ssDNA containing the gene of interest and TTV19 intergenic region, and TTV19 capsid protein.....	78
4.4.1. Iodixanol ultracentrifugation of crude extract containing TTV capsid protein and assembled crude extract containing TTV19 capsid protein and ssDNA.	79
4.4.2. Transmission electron microscopy (TEM) for <i>in vitro</i> assembled TTV19 capsid protein and ssDNA containing the gene of interest.	80
4.4.3. Detection of packaging of ssDNA by TTV19 capsid protein <i>in vitro</i> by PCR.	85
4.4.4. Confirmation of packaging of ssDNA by TTV19 capsid protein and determination of concentration of TTV19 virion particles.	86
5. DISCUSSION	89
6. SUMMARY	96
7. FUTURE RESEARCH.....	97
REFERENCES.....	99
APPENDICES	124
Appendix A: TTV19 DNA sequences	124
Appendix B: Media and buffer compositions.....	126
Appendix C: Supplementary qPCR Data.....	127
Appendix D: TTV19 ORF1 protein cleavage site	128

List of Figures

FIGURE 1. <i>In vivo</i> and <i>ex vivo</i> techniques related to gene delivery to target cells.....	9
FIGURE 2. <i>In vivo</i> and <i>ex vivo</i> techniques related to gene delivery to target cells.....	10
FIGURE 3: Key steps involved in gene therapy.....	13
FIGURE 4: Genome organization of anellovirus TTV isolate TA278.....	39
FIGURE 5: Possible mechanisms whereby TTV infection may modulate host defences.....	43
FIGURE 6: Structure of the filamentous bacteriophage.....	45
FIGURE 7: A) Schematic representation of M13 phage particle depicting the organization of coat proteins around the ssDNA genome. B) Map of the M13 phage genome identifying all genes, the intergenic region, promoters, and terminators.....	47
FIGURE 8: Life cycle of M13 phage.....	49
FIGURE 9. Plasmid maps.....	59
FIGURE 10. Agarose gel of pET-28_TTV19Cap and its restriction digestion.....	70
FIGURE 11. Production of TTV19 capsid protein.....	71
FIGURE 12: Production of TTV19 Capsid protein at different O.D.....	72
FIGURE 13: Purification of His-Tagged TTV19 capsid protein.....	74
FIGURE 14: Standard BCA graph for protein concentration determination.....	75
FIGURE 15: Phagemid Plaque assay using spot plate method.....	76
FIGURE 16: Calibration graph to estimate phage concentration.....	77
FIGURE 17: ssDNA obtained from M13 phagemid carrying the gene cassette.....	78
FIGURE 18: Iodixanol ultracentrifugation of crude extract containing TTV19 capsid protein.....	79
FIGURE 19: Iodixanol ultracentrifugation of <i>in vitro</i> assembled ssDNA and crude extract containing TTV19 capsid protein.....	80

FIGURE 20: TEM images for crude extract containing TTV19 extract samples.....	82
FIGURE 21: TEM images His-Tag purified TTV19 extract samples.....	84
FIGURE 22: 1.8% gel showing PCR amplification of gfp gene region.....	86
FIGURE 23: Calibration graph to estimate TTV19 virion concentration.....	87

List of Tables

TABLE 1: Gene therapy products and clinical trials.....	4
TABLE 2: Comparison of Non-viral and Viral vector systems.....	14
TABLE 3: Major viral vectors used in gene therapy.....	23
TABLE 4: Advantages and Disadvantages of non-viral vectors.....	32
TABLE 5: Properties of ssDNA viruses.....	34
TABLE 6: List of genera in the family <i>Anelloviridae</i>	37
TABLE 7: Bacterial strains used in the study.....	57
TABLE 8: Plasmids used in the study.....	57
TABLE 9: Primers used in this study.....	57
TABLE 10: PCR amplification protocol.....	67
TABLE 11: Genomic characteristics and locations of transcription elements of the <i>Anellovirus</i> species.....	69
TABLE 12: C _t Values and virion concentration for the qPCR amplified samples.....	88

List of Abbreviations

AAV	Adeno-associated virus
ADA	Adenosine deaminase
AML	Acute myeloid leukemia
Amp ^R	Ampicillin resistance
ASC	Adult stem cell
AV	Adenovirus
AAV	Adeno-associated virus
CT	Cell Therapy
CP	Coat Protein
CMV	Cytomegalovirus
CNS	Central nervous system
DLBCL	Diffuse large B-cell lymphoma
DMD	Duchenne Muscular Dystrophy
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
<i>E. coli</i>	<i>Escherichia coli</i>
EMA	European Medicines Agency
ESC	Embryonic stem cell
FDA	Food and Drug Administration
Ff	Filamentous phage
GARD	Genetic and Rare Diseases Information Center
GFP	Green fluorescent protein
GT	Gene therapy
GTMP	Gene therapy medicinal product
HA	Hemagglutinin
HIV	Human immunodeficiency virus
HNSCC	Head and neck squamous-cell carcinoma
HSV	Human Simplex virus
HSV-TK	Herpes simplex thymidine kinase
ICTV	International committee on Taxonomy of viruses
IL	Interleukin
iPSC	Induced pluripotent stem cell
ITR	Inverted terminal repeats
Kn ^R	Kanamycin Resistance
LB	Luria-Bertani
LPLD	Lipoprotein lipase deficiency
LV	Lentivirus
MDS	Myelodysplastic syndrome
miRNA	MicroRNA
MRI	Magnetic resonance imaging
NCATS	National Center for Advancing Translational Sciences
NCBI	National center for Biotechnology information
NFκB	Nuclear factor kappa-light-chain enhancer of activated B-cell
NHL	Non-Hodgkin's lymphoma

ORF	Open reading frame
PAMP	Pathogen-associated molecular pattern
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PMBCL	Primary mediastinal large B-cell lymphoma
PPR	Pattern recognition receptor
PSC	Pluripotent Stem Cell
PTM	Post-translational modifications
qPCR	Real-time polymerase chain reaction
RCR	Rolling Circle replication
RES	Reticuloendothelial system
RF	Replicative Form
RNA	Ribonucleic acid
RV	Retrovirus
RHR	Rolling Hairpin Replication
SAS	Sense anti-sense
SCID	Severe combined immunodeficiency
SDS PAGE	Sodium dodecyl sulphate–polyacrylamide gel electrophoresis
SMN	Survival motor neuron
ssDNA	Single-stranded DNA
TEM	Transmission electron microscopy.
TESAE	Treatment-emergent serious adverse events
TLR	Toll-like receptor
TMA	Thrombolytic microangiopathies
TTMDV	torque teno midi virus
TTMV	Torque teno mini virus
TTV	Torque teno virus
TEM	Transmission electron microscopy
UTR	Untranslated Region
VEGF	Vascular endothelial growth factor

1. INTRODUCTION

The human genome consists of approximately 25,000 genes that encode many gene products and genetic elements responsible for biological processes necessary for the maintenance of life. The genetic code remains largely unchanged and is passed on from one generation to another. However, sometimes a gene can be altered due to mutation which affects gene processing or expression. These genetic mutations lead to altered gene product expression, processing, or functions affecting the normal cell structures, functions, and their physiological roles, which can manifest as life-threatening diseases and disorders. According to the leading rare disease patient advocacy organizations in the world: Genetic and Rare Diseases Information Center (GARD) and Global Genes®, dysfunctional genes account for 80% of the total 7,136 diseases reported to date. More than 300 million people in the world are affected by genetic diseases. The National Center for Advancing Translational Sciences (NCATS) reports that only 500 human diseases are treatable with an estimated 10,000 drugs available to date, thus necessitating the need to develop new drugs and treatment options. With the advancement in modern medicine, therapies like chemotherapy, radiation, and surgery have offered some relief to combat diseases like cancer by alleviating disease symptoms and increasing lifetime. However, such treatments often involve the use of synthetic chemicals which over a period alters the body's chemistry and creates dependency over time. Such issues have been partly resolved by developing a wide array of protein-based therapies including transcription factors, signaling proteins, gene editing enzymes, growth factors, engineered protein scaffolds, hormones, blood factors, thrombolytes, antibodies, and antigens (Usmani et al., 2017; Fu et al., 2020), which have been largely incorporated into the treatment of various diseases such as cancer, diabetes, infectious diseases, and other disorders (Boohaker et al., 2012; Fu et al., 2020). Protein and peptide drugs have various

advantages compared to other drugs as they are highly versatile, provide a wide range of pharmaceutical targets, have high specificity, and demonstrate low toxicity levels (Fosgerau and Hoffmann, 2015). Over 60 peptide drugs have been approved in the United States and other major markets and several other protein therapeutics are currently being evaluated in different phases of clinical trials (Lau and Dunn 2018; Fosgerau and Hoffmann, 2015). Although protein and peptide-based therapies offer several advantages there are a few limitations associated with such therapies, such as perhaps low solubility, proteolytic degradation, short circulating half-life, physiochemical instability, and perhaps most importantly, immunogenicity (Antosova et al., 2009; Bruno et al., 2013), thus making the administration and compliance of protein and peptide-based therapeutics a highly complicated and difficult process. Along with these drawbacks, the production and manufacturing of protein and peptide drugs are highly complex processes; for instance, some protein and peptide drugs may involve multiple critical steps to produce a single protein therapeutic and involve high production costs (Walsh 2014; Lagasse et al., 2017). Due to the above-stated reasons, there is a great deal of unpredictability involved in the development of both chemical and protein-based therapeutics. These drawbacks and risks can be circumvented and dealt with by gene-based therapies that may enable long-lasting production of the desired therapeutic proteins and at the same time localize gene product expression to the affected site of the body (Ibraheem et al., 2014). Thus, gene therapy holds immense potential for treating a myriad of diseases such as cancer, immunodeficiency, cardiovascular diseases, and more without involving pharmacological intervention, radiotherapy, or surgery (Chilukuri et al., 2008; Aleti et al., 2013; Dunbar et al., 2018; Prondzynski et al., 2018).

1.1 GENE THERAPY

The basic concept of gene therapy involves the introduction of a gene into a target cell which, either leads to treating or the slowing down of the progression of the disease. The concept of GT arose in the 1960's when scientists first demonstrated the feasibility of incorporating new genetic functions in mammalian cells (Friedmann 1992; Verma and Weitzman, 2005). In 1972, the idea of gene therapy was proposed and received approval to be used as a treatment method (Friedmann and Roblin, 1972). The first clinical trial which involved the use of retrovirus labeling to detect tumors was subsequently approved in 1982 by the National Institute of Health (Rosenberg et al., 1990). This study also demonstrated that human cells can be genetically engineered in vitro and that these cells can later be re-introduced into the patient's body safely – an approach now referred to as ex vivo cell therapy (Rosenberg et al., 1990). In 1995, the first successful clinical study for gene therapy was reported when a patient suffering from severe combined immune deficiency syndrome (SCID) received treatment using retrovirus reinfusion carrying the adenosine deaminase (ADA) gene to correct the immunodeficient condition (Blaese et al., 1995). Over the years the definition of gene therapy has become broader. The Food and Drug Administration (FDA) defines gene therapy as products that “mediate their effects by transcription and/or translation of transferred genetic material and/or by integrating into the host genome and that are administered as nucleic acids, viruses, or genetically engineered microorganisms,” and the European Medicines Agency (EMA) describes gene therapy medicinal product (GTMP) as a “biological medicinal product that contains an active substance which contains or consists of a recombinant nucleic acid used in or administered to humans to regulate, repair, replace, add or delete genetic sequences and its therapeutic, prophylactic or diagnostic effect relates directly to the recombinant

nucleic acid sequence it contains, or to the product of genetic expression of this sequence” (Hanna et al., 2017; Wirth et al., 2013).

Even after early success with GT and its clear demonstration of potential to treat disease, the further development of the therapy was faced with issues due to the adverse effects caused by the virus-induced immune response and malignancy. These drawbacks were mitigated by the advent of viral vectors like adeno-associated viruses, and advancement in gene editing tools (Duan et al., 2021). Initially, GT was only used for treating genetic disorders which now has extended to treat both acquired and inherited diseases. To date, there have been 943 clinical trials launched worldwide all focusing on somatic GT (Duan et al., 2021). Table 1 lists some of the recent GT products and approved clinical trials (Duan et al., 2021).

TABLE 1: Gene therapy products and clinical trials.

PRODUCT	APPROVAL ORGANIZATION	APPROVAL TIME	DISEASES	VECTOR	WORKING PRINCIPLES	BOTTLENECK
Vitravene (Fomivirsen)	FDA/ EMA	1998. First approved gene therapy drug, later withdrawn from the market in 2002 as the number of CMV in HIV patient cases decreased after the development of HAART.	Treatment of CMV retinitis in immunocompromised patients	RNA, was delivered locally by IVT administration and hence did not require a delivery agent (Hammond et al., 2021).	Antisense oligonucleotide against UL123 gene of CMV	N/A
Gendicine	CFDA	2003	Head and neck squamous-cell carcinoma (HNSCC)	Adenovirus	Recombinant human p53 adenovirus	High immunogenicity
Oncorine (H101)	China FDA	2005	Nasopharyngeal	Adenovirus	E1B-deleted adenovirus	Self-relieving local injection,

carcinoma							mild to medium grade fever and influenza like symptoms
Neovasculgen (PI-VEGF165)	Russian Ministry of Healthcare	2012. First plasmid-based medicine.	Peripheral vascular disease and limb ischemia	Plasmid, directly injected into the target ischemic tissue (Goswami et al., 2019).	Plasmid DNA encoding the 165-amino-acid isoform of human vascular endothelial growth factor	Not recorded	
Glybera	EMA	2012. Due to the high production cost of Glybera and lack of demand led to the discontinuation of the therapy in 2017.	Lipoprotein lipase deficiency (LPLD)	AAV1	Adeno-associated virus (AAV) carries the functional LPL gene	Immune response, Low packaging capacity, and Random vector genome insertion (price was unaffordable)	
Imlygic	FDA	2015	Melanoma	HSV1	Using Oncolytic, engineered herpes virus to anticancer	Immune response	
Zalmoxis	EMA	2016. Received conditional approval which was later withdrawn in October 2019.	Leukemia and lymphoma	Retrovirus	Treatment with donor T cells were genetically modified to express the herpes simplex thymidine kinase (HSV-TK) suicide gene21	Immune response	
Strimvelis	EMA	2016	ADA-SCID	Retrovirus	Using retroviral vector to insert a functional copy of the gene ADA to CD34 + cells	Immune response, and Random vector genome insertion	
Exondys 51(Etepliren)	FDA	2016	Duchenne Muscular Dystrophy (DMD)	RNA	Antisense oligonucleotide for dystrophin	Not recorded	
Spinraza	EMA	2016	Spinal muscular	RNA	Using an antisense	Unstable molecules	

				atrophy Non-Hodgkin's		oligonucleotide to alter the splicing of the survival motor neuron (SMN) gene	
Kymriah	FDA	2017		Non-Hodgkin's lymphoma (NHL)	Lentivirus	Chimeric antigen receptor T-cell immunotherapy targets CD19 antigen	Immune response, Lack of targetable tumour-specific surface antigens, and Tumour cell heterogeneity
Luxturna	FDA	2017		Leber congenital amaurosis	AAV2	Using AAV to deliver normal human RPE65 gene to retinal cells	Immune response, Low packaging capacity, and Random vector genome insertion
EDIT-101	FDA	2017. Under phase-II clinical trial		Leber congenital amaurosis 10 Eliminating	AAV5	Eliminating a mutation in the CEP290 gene using CRISPR	NA
Yescarta	FDA EMA	2017 2018		Diffuse large B-cell lymphoma (DLBCL); primary mediastinal large B-cell lymphoma (PMBCL)	Retrovirus	Chimeric antigen receptor T-cell immunotherapy targets CD19 antigen	Immune response, Lack of targetable tumour-specific surface antigens, and Tumour cell heterogeneity
Tegsedi (Inotersen)	EMA FDA	2018		Hereditary Transthyretin-related Amyloidosis	RNA	Antisense oligonucleotide against transthyretin mRNA	Inflammatory and immune system problems, stroke, allergic reaction, and low levels of vitamin A leading to eye problems
Zolgensma	FDA EMA	2019. Most expensive drug worldwide.		Pediatric Spinal Muscular Atrophy	AAV9	Replacing abnormal SMN1 gene with a normal SMN1 gene	May lead to liver damage
Zynteglo	EMA	2019		Transfusion-dependent β thalassemia	Lentivirus	Autologous CD34 + cells encoding thalassemia β A-T87Q-globin gene in vitro using a	Immune response, and Random vector genome insertion

CTX001	FDA	2019. phase clinical trial.	Under I/II trial.	β - thalassemia and sickle anemia	Lentivirus	lentiviral vector Using the CRISPR gene- editing technology to increase the production of fetal hemoglobin in patients' red blood cells	NA
Libmeldy (Atidarsagene autotemcel)	EMA	2020		Metachromatic Leukodystrophy	Lentivirus	Autologous CD34 + cells encoding ARSA gene	Immune response
Breyanzi (lisocabtagene maraleucel)	FDA	2021		Relapsed or refractory diffuse large B cell lymphoma; follicular lymphoma	Retrovirus	CD19 CAR	Inflammatory response
Abecma (Idecabtagene vicleucel)	FDA EMA	2021		Multiple myeloma	Lentivirus	BCMA CAR	Inflammatory response, febrile neutropenia
Skysona (elivaldogene autotemcel/ Lenti-D)	EMA	2021. withdrawn 2022	Later in	Juvenile Cerebral Adrenoleukodystrophy	Lentivirus	<i>ABCD1</i> gene	Pancytopenia
Carvykti (ciltacabtagene autoleucel)	FDA	2022		Relapsed or refractory multiple myeloma	Lentivirus	BCMA CAR	Inflammatory response

(Adapted From: Shahryari et al., 2019; Duan et al., 2021; Arabi et al., 2022)

1.1.1 GENE THERAPY SYSTEMS

GT is based on the type of cell receiving the treatment and is divided into germline and somatic cell therapies (Goswami et al., 2019). Germline therapies are based on the genetic manipulation of reproductive cells, either sperm or egg, that are passed from one generation to another. Successful germline clinical GT has been demonstrated in mice, rats, rabbits,

sheep, cattle, goats, and pigs (Travis, 2015; Naldini, 2015; Cohen and Adashi, 2016). However, germline GT in humans is prohibited due to ethical issues and the lack of advanced tools (McCarthy, 2015). Therefore, in humans, GT is only carried out in somatic cells i.e., all cells except sperm and eggs. Somatic GT is considered a safe treatment as genetic changes remain in the patient and are not passed on to the next generation (Goswami et al., 2019). There are two ways by which GT may be carried out in somatic cells – *ex vivo* or *in vivo* (Figure 1) (Somia and Verma, 2000). The *in vivo* method directly involves the administration of desired genetic material into the targeted cells or tissue (Goswami et al., 2019). This approach is mostly employed for diseases such as cystic fibrosis, muscle dystrophy, and cancer but still requires more technological advancements to obtain successful results (Manunta et al., 2017; Nance et al., 2017; Guan et al., 2017). The therapeutic drug Neovasculagen® which is a plasmid vector carrying vascular endothelial growth factor (VEGF) is directly injected into the targeted ischemic tissue to stimulate blood vessel growth (Deev et al., 2017). Whereas an *ex vivo* method is the traditional cell therapy (CT) approach which involves the isolation of cells from the patient. The cells are modified to add or replace the therapeutic genetic element(s) and then re-inserted into the patient's body (Goswami et al., 2019). Cells such as hepatocytes in the liver, retina photoreceptors in the eye, stem cells in the bone marrow, and T-lymphocytes have been the focus of the *ex-vivo* method of treatment (Cohen and Adashi, 2016). For instance, the FDA-approved Kymriah™, a prescription cancer treatment, uses a patient's white blood cell or T-cells for inserting CD19 gene *ex vivo* (Morrow, 2017).

In vivo gene therapy

Ex vivo gene therapy

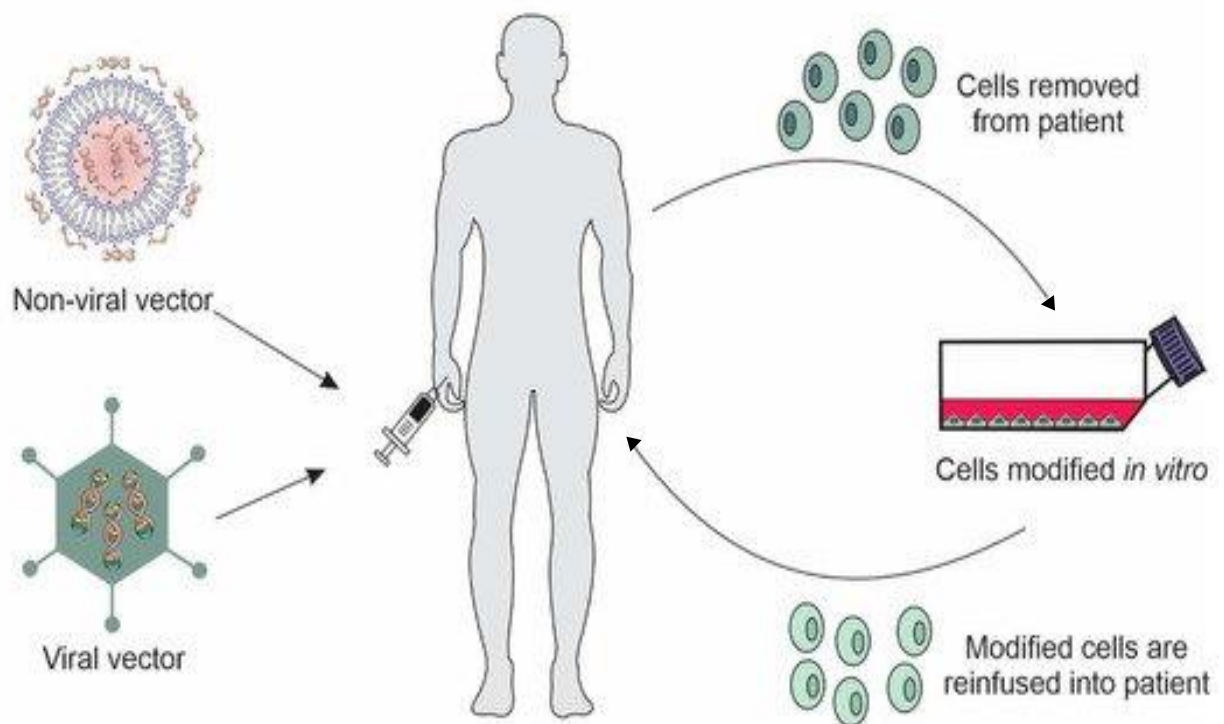


FIGURE 1. *In vivo* and *ex vivo* techniques related to gene delivery to target cells.
(Adapted From: Luiz et al., 2022).

With the advancement in science and technology, GT has become the most explored avenue for therapeutic as well as diagnostic purposes. The growing interest in stem cell research opened new avenues for these pluripotent cells in the field. Stem cells are self-renewing cells that are capable of undergoing differentiation into various cell or tissue types (Chang et al., 2019). Based on their differentiation capacity, they are divided into multipotent and pluripotent categories (Chang et al., 2019). The Pluripotent stem cells (PSC) include embryonic stem cells (ESC) derived from the embryo and induced pluripotent stem cells (iPSC) derived by gene transfer that may undergo indefinite proliferation and differentiation into different types of tissues depending upon the treatment conditions given (Chang et al., 2019). In contrast, multipotent stem cells include adult stem cells (ASC) which are obtained from tissue precursors (Chang et al., 2019). Figure 2 summarizes the procedure for obtaining

the stem cells and their characterization (Chang et al., 2019). Of these stem cell types; iPSC research has rapidly progressed as these cells, face fewer ethical concerns compared to ESCs, and are available for the development of new clinical applications (Robinton and Daley, 2012).

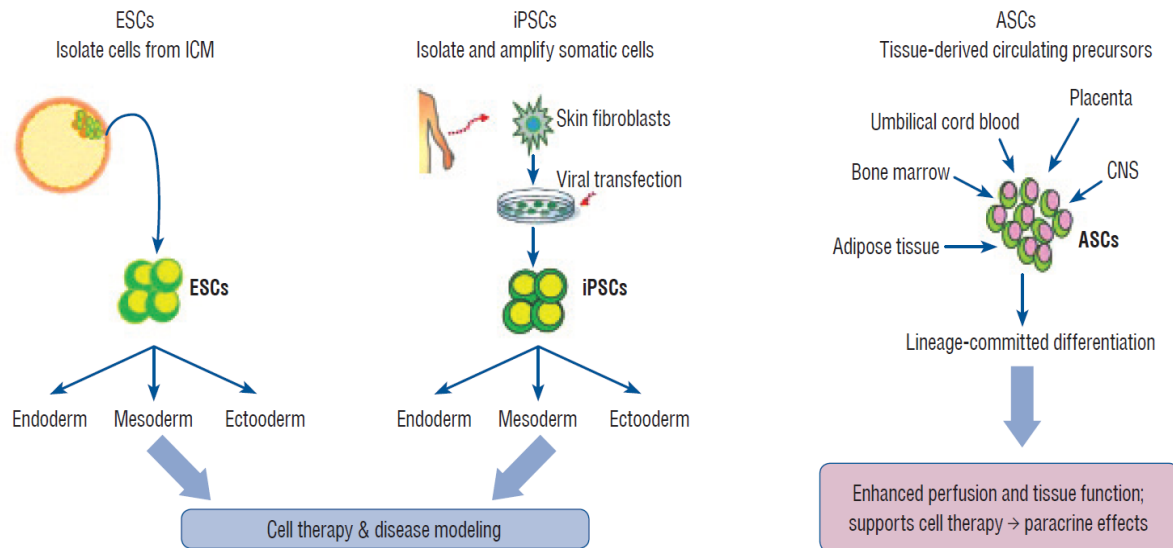


FIGURE 2. Isolation and characterization of different stem cells. ESC: embryonic stem cells, ICM: inner cell mass, iPSC: induced pluripotent stem cell, ASC: adult stem cell, and CNS: central nervous system (Chang et al., 2019)

iPSCs were first developed in 2006 by Yamanaka using mice models which led to the subsequent establishment of human iPSCs in 2007 (Takahashi et al., 2007). These stem cells are derived from somatic cells of the adult body through the expression of specific exogenous genes or proteins that resembles ESC morphologically, antigenically, and phenotypically in many ways (Chang et al., 2019). Contrary to ESCs, iPSCs can be generated from various somatic cell tissues like dental tissue, peripheral blood, and urine (Chang et al., 2019). Therefore, the generation of iPSCs has demonstrated fewer ethical concerns compared to ESCs and offers the advantage of customized treatment using the somatic cells of the patient with the same characteristics as ESCs (Turner et al., 2013; Yoshida and Yamanaka, 2011).

The pluripotency in somatic cells was successfully induced in adult cells by Yamanaka's team. They used four retrovirally transfected transcription factors, namely, Octamer 3/4 (Oct3/4), SRY-box containing gene 2 (Sox2), Krüppel-like factor 4 (Klf4), and the proto-oncogene cytoplasmic Myc protein (c-Myc), collectively known as “Yamanaka factors”, successfully dedifferentiating fibroblasts (Takahashi et al., 2006; Takahashi et al., 2007). Although, the retroviral infection technology delivered each of the “Yamanaka factors”, the technique was posed with two major disadvantages, limiting use towards clinical applications. First, the genomic integration of reprogramming factors could lead to unwanted outcomes such as tumourigenesis; and second, this technology could impair pluripotency (Chang et al., 2019). Therefore, reprogramming factors were required to establish pluripotency, but continuous activation of an exogenous gene might decrease the differentiation capacity into specific cell types or transform cells altogether (Yu et al., 2007; Meissner et al., 2007). Alternative induction methods have since been introduced to avoid direct alterations of host cell DNA where the pluripotency genes either remain separate from the host genome or can be completely removed. This was achieved by enabling transient expression of reprogramming factors by using adenoviruses, plasmids, minicircle vectors, Sendai viruses, synthetic mRNA, or directly with recombinant transcription factor proteins (Chang et al., 2019). The reprogramming technology of iPSC improved the safety and increased the efficiency of iPSCs to be used for clinical applications (Yoshida and Yamanaka, 2011). iPSCs may be used for various clinical applications such as the development of disease-specific autologous cell therapy, disease models to evaluate underlying mechanisms, and drug screening and toxicity tests (Chang et al., 2019). More intense research is required for iPSC technology for safer application and commercial utilization.

The success of gene therapy mostly depends on the vector. A vector is a carrier for the genetic materials such as DNA, mRNA, siRNA, miRNA, and anti-sense oligonucleotides to reach its targeted destination either to restore a specific gene function or turn off the gene responsible for the disease or disorder development (Goswami et al., 2019; Zu and Gao, 2021). Figure 3 demonstrates the key steps involved in gene therapy (Ramamoorth and Narvekar, 2015). The success of both *ex vivo* and *in vivo* gene therapy methods depends on the efficiency of vector transport and unloading into target cells, perseverance, activity, immune response, regulatory issues and ethical concerns, and commercialization (Wang et al., 2013). The choice of vector plays a pivotal role in the treatment approach. Two types of vectors can be used for GT: non-viral and viral vectors (Somia and Verma, 2000). The ideal gene delivery vector should be: i) very specific; ii) capable of efficiently delivering one or more genes of the size needed for clinical application; iii) unrecognized by the immune system; iv) scalable for production in large quantities at high concentration; v) non-allergenic; vi) able to express therapeutic transgene(s) for the required period of the treatment cycle (Somia and Verma, 2000).

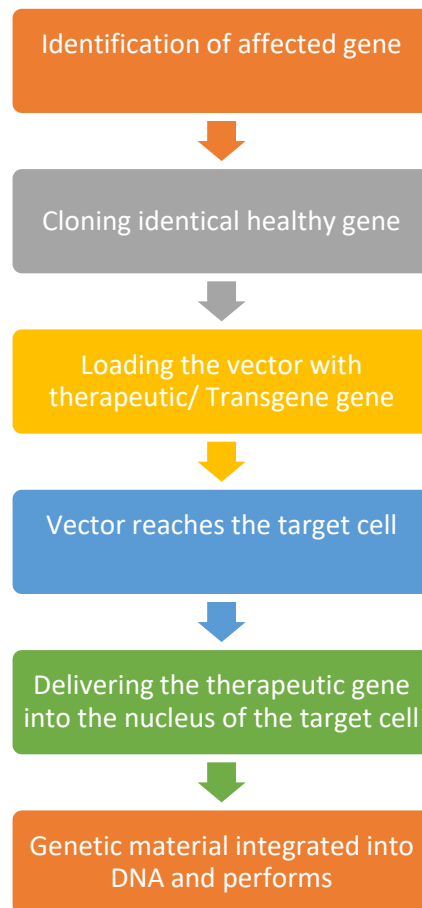


FIGURE 3: Key steps involved in gene therapy. (Adapted from Ramamoorth and Narvekar, 2015).

1.1.2 NON-VIRAL VECTORS vs VIRAL VECTORS

The success of GT highly depends on the vector system. Both gene therapy systems: non-viral and viral vectors offer advantages and disadvantages as discussed above. These attributes are summarized in Table 2. To date, viral vectors have dominated the industry and its clinical studies due to high transfection and gene expression efficacy. Most of the currently approved gene therapy uses viral vectors as shown in Table 1. However, the safety concerns associated with viral vectors are many and continue to plague GT clinical trials with human deaths and malignancies. Some of these limitations might be overcome by using genome editing tools by removing unwanted genes. A more recent approach employs bacteriophages such as Ff filamentous phages to be used as a non-viral vector. In combination with Ff phage display phages are designed to display specific peptide/protein

sequences on the surface that allows them to interact with mammalian cells for gene delivery (Ju and Sun, 2017). Although this approach is safe and more precise, it demonstrates low gene delivery efficiency and relatively poor gene expression. Each vector system will be discussed in detail in subsequent sections

TABLE 2: Comparison of Non-Viral and Viral Vector systems.

FEATURE	NON-VIRAL VECTOR	VIRAL VECTOR
Application	Therapeutics, theranostics and diagnostics (inherent controllable, optical, and electrical properties of materials used)	Therapeutics and diagnostics
<i>In-vivo</i> Efficiency	Dependent on size, shape, surface charge, surface chemistry and surface-to-volume ratio	Strong promoters for gene expression but could also revert to wild-type expression, long term effects.
Toxicity	Depends on the chemical composition. Some of the vectors use non-biodegradable agents which causes safety concerns.	High immunogenicity and possibility of reversion to virus of the wild type
Transfection Efficiency	Delivery is relatively low. Can be facilitated by external magnetic field, sonophoresis, temperature effects. Can be enhanced by attaching appropriate ligands	High
Nucleic Acid Incorporation	Hold the genes or biomolecules through the surface functional modifier. Loading is limited by the particle surface area and functionalization	Directly into capsid
Advantages	Easy to prepare, low cost involved, lower immunogenicity, high flexibility.	Ex vivo stable transfection of blood cells, high efficiency on particle basis
Limitations	Low efficiency on particle basis	Risk of insertional mutagenesis, immunogenicity, and inflammatory responses.

(Adapted from : Pranatharthiharan et al., 2013; Uludag et al., 2019)

1.2.2.A. VIRAL VECTORS

The most common viral vectors are Adenovirus, Adeno-associated virus (AAV), Retroviruses, Lentiviruses, and Herpes simplex virus. Additional vectors used in GT, but to a lesser degree, include vaccinia virus, togavirus, poxvirus, alphavirus, Foamy virus, and others (Lukashev and Zamyatnin, 2016; Lundstorm, 2003; Chira et al., 2015). Viruses have evolved to become highly efficient at nucleic acid delivery to specific cell types while avoiding immunosurveillance. In the past, there has been resentment toward adopting viral vectors due

to safety concerns, ignorance, and skewed views. The human body is home to myriads of microorganisms such as viruses, bacteria, fungi, and protozoa who have adopted appropriate mechanisms for mutual benefits to survive and thrive (Roossinck, 2015). For instance, viruses control gut bacterial genes to enhance host digestion and provide defense against bacterial infections and cancer cells. (Roossinck and Bazan, 2017). Although the word ‘virus’ implies parasitism, mortality, and morbidity, viruses are considered nature’s genetic engineers, due to their ability to infect a variety of organisms including bacteria, humans, animals, and plants (Roossinck and Bazan, 2017). Viral vectors can, therefore, infect a wide variety of cells and impart stable and high transgene expression. However, the application of viral vectors in gene delivery does possess serious challenges. When using viral vectors, the general concerns include limited packaging size, immune reactions, off-target effects, inflammation, and insertional mutagenesis, each varying in severity depending on the attributes of the virus. There are high risks for viral vectors as they may elicit inflammatory immune responses, making treatment less efficient and thus, resulting in the creation of antibodies and cell-mediated memory that precludes a second dosage of the same virus. Insertional mutagenesis is another major obstacle in using viral vectors, particularly retroviruses, as there is a risk of the genetic cargo inserting randomly into a tumor suppressor or activating an oncogene via integration into dangerous regulatory locations within the genome.

Researchers have discovered potent viral promoters and enhancer elements that can be utilized to build plasmid vectors for the high-level production of foreign proteins because of advances in science and technology. (Viswanathan et al., 2003). Because they have more genes encoding structural and non-structural proteins, they can infect and spread across host cells than other organisms. (Thomas et al., 2003). Few viruses can transduce, or cells that can stably express a gene together with the host's DNA. (Thomas et al., 2003). They permit

genome modification and gene deletion without affecting their capacity to infect host cells. (Thomas et al., 2003). Due to their adaptability, they are useful biological instruments for introducing foreign genetic elements into eukaryotic cells (Thomas et al., 2003). The development of new viral vectors is being researched constantly, with a particular emphasis on replacing harmful genes with beneficial ones. More than 70% of the global gene therapy clinical studies currently being conducted use viral vectors that are non-pathogenic, replication-defective, and friendly to humans (Yin et al., 2014). The most used viral vectors for gene therapy include adenovirus (AV), Adeno-Associated Virus (AAV), Herpes Simplex Virus (HSV), Retrovirus (RS), Lentivirus (LV), etc.

- 1. Adenovirus (AV):** AVs are DNA viruses that cause benign infections of the upper respiratory tract in humans. AV is a non-enveloped icosahedral viral particle that carries a linear dsDNA genome ranging from 26 to 45 kb. Transcripts undergo alternative splicing to encode genes in both the sense and antisense orientations (Wold and Toth, 2013). AV was the first viral vector developed for gene therapy and was approved for clinical trials in 1990 for the original delivery of adenosine deaminase in the treatment of Severe Combined Immunodeficiency Disease (Goswami et al., 2019). The replication cycle of the virus can be divided into two phases: the early phase which corresponds to events that occur before the replication of the viral DNA, and the late phase which corresponds to the period following the initiation of viral DNA replication (Robbins et al., 1998). The AV genome is flanked by 2 inverted terminal repeats (ITRs) and contains 8 transcription units (Davison et al., 2003). The early regions (E1A, E1B, E2, E3, and E4) are the first to be transcribed and include the protein sequences necessary to activate the transcription of additional viral areas and change the cellular environment to facilitate virus generation (Davison et al., 2003). A transcript that has been alternatively spliced is used to transcribe the late regions (L1-L5) (Khare et al.,

2011). Only one or two viral early genes (E1 and E3) were removed from the first-generation AVs (Ahi et al., 2011). Strong cytotoxic T-cell responses were triggered by the low amounts of other adenoviral genes expressed by cells that had been transduced with these vectors, which swiftly abolished transgene expression (Ghosh et al., 2020). The most significant development to reduce immunogenicity, extend transgene expression, and enhance the prospects of adenovirus vectors for long-term gene therapy has been the development of helper-dependent adenoviruses that are deleted for all viral genes (Schaack, 2005; Alba et al., 2005). This is in contrast to second and third-generation vectors that contain additional deletions in other early genes (E2 and/or E4). The adenoviruses have several features that make them well-suited to be used as a vector for gene therapy applications. First, they have been isolated from numerous species, are widely distributed, and have been reported in more than 100 distinct serotypes (Wold and Toth, 2013). Second, compared to other currently available vectors, AVs can quickly infect a wide variety of human cells and provide significant quantities of gene transfer (Ghosh et al., 2020). Third, AVs may transduce these transgenes in non-proliferating cells despite their relatively substantial DNA capacity (up to 7.5 kilobase pairs) (Ahi et al., 2011). Fourth, through subsequent rounds of viral replication, added foreign genes are typically preserved unchanged because the viral genome does not suffer high rates of rearrangement (Ahi et al., 2011). Finally, utilizing recombinant DNA techniques, AVs are rather simple to alter (Goswami et al. 2019). Although AVs have many advantages, the use of AVs is plagued by challenging limitations. As the packaging capacity of AVs is 8kb for replication-defective vectors and 30kb for helper-dependent vectors, manufacturing vectors with large DNA sizes is difficult (Morsy and Caskey, 1999). Another concern related to their use is that these

viruses can and have elicited potent inflammatory responses that can even be fatal (Lowenstein et al., 2007).

2. **Adeno-Associated Virus (AAV):** AAV is the most exploited virus in the generation of gene therapy vectors. AAV is a non-pathogenic parvovirus that is composed of a linear single-stranded DNA (ssDNA) approximately 4.7 kb in size (Naso et al., 2017, McClements and MacLaren, 2017). It encodes two 145 nucleotide ITRs at the termini that fold in as closed ends in the linear ssDNA genome. The virus lacks genes for polymerase enzyme and is dependent upon cellular polymerase for replication (Chen et al., 2018). The genome of AAV consists of three open reading frames flanked by these ITRs (Yan et al., 2005) that function as the viral origin of replication and the packaging signal (Naso et al., 2017). The rep ORF encodes 4 non-structural proteins that play a vital role in viral replication, transcription regulation, genome integration, and virion assembly (Wang et al., 2018). The cap ORF encodes 3 structural proteins: VP1, VP2, and VP3, that assemble to form a 60-mer (repeat unit) viral capsid (Maurer et al., 2018). The assembly-activating protein, which directs AAV capsid proteins to the nucleolus and participates in capsid assembly, is created by an ORF that exists as an alternative reading frame within the cap gene (Earley et al., 2017). In recombinant forms of AAV, the gene of interest is inserted between the ITRs in place of rep and cap, which are instead encoded in trans, along with helper viral genes, during vector production. The importance of this vector is that it can transduce both dividing and non-dividing cells with stable transgene expression (Van et al., 2008). The transgene expression is much more durable compared to other vectors and can infect various cells such as the liver, muscle, retina, and Central Nervous System (CNS) (Van et al., 2008, Earley et al., 2017). However, the pre-existing immunity to human AAV vectors is comparable to those of AVs, and the integration into the host genome is low but can still on occasion

lead to insertional mutagenesis and accidental activation or inhibition of endogenous gene expression (Colella et al., 2018). But there have been numerous reports of treatment-emergent serious adverse events (TESAE) in GT studies with AAV vector-based products in recent years, including hepatotoxicities, thrombotic microangiopathies (TMA), and brain magnetic resonance imaging (MRI) findings of uncertain significance with some TESAEs that led to the death of study participants (Kuzmin et al., 2021; Wilson et al., 2020; Perez et al., 2020). The modest packaging capacity of AAV vectors, which is typically estimated to be less than 5 kb (Chen et al., 2018), is another significant restriction. These limitations make AAV vector design for gene therapy challenging.

3. **Retrovirus (RV):** Retroviruses are enveloped RNA viruses with a diploid ssRNA genome consisting of at least 4 genes namely, gag, pro, pol, and env (Kurian et al., 2000). The gag gene encodes the primary structural polyprotein required for the assembly of non-infectious and immature viral-like particles (Kootstra and Verma, 2003); the pro gene encodes for the viral protease that is in charge of promoting the maturation of viral particles; the pol gene encodes for reverse transcriptase, RNase H, and integrase; and the env gene is in charge of the viral surface glycoprotein and transmembrane proteins (Gunzburg et al., 1996). The typical feature of retroviral vectors is their ability to integrate randomly into host DNA (Chuah et al., 2003). The viral RNA is reversibly transcribed and integrated into the form of a provirus (Weiss and Tailor, 1995). The RV efficiently utilizes host cellular enzymes for their replication and long-term expression of viral proteins (Vile and Russell, 1995). In the context of gene therapy, clinical applications for monogenic disorders, cancer, and infectious diseases, retroviral vectors have been widely employed to transfer therapeutic genes, providing patients with a stable and effective expression of the transgene (Kurian et al.,

2000, Smith 1995). Retroviruses have numerous advantages over other vectors. The most significant advantage is their ability to transform their ssRNA genome into a dsDNA molecule that stably integrates into the target cell genome which helps the retroviral vectors to modify the host cell nuclear genome permanently (Vile and Russell, 1995). This advantage also is the cause of primary safety risk that arises from their ability to integrate into the host cell chromosome, which raises the possibility of insertional mutagenesis and oncogene activation (Romano, 2012). Also, the target cell range of the vector is a safety concern as the incorporation of a virus envelope that can infect cells from multiple species increases both the risk of generating replication-competent retroviruses and the potential danger of any resulting virus, which could spread from one species to another (Romano et al., 2009).

4. **Lentivirus (LV):** Lentiviruses are a type of retrovirus composed of a single-stranded positive-sense RNA genome (Lundstorm, 2019). Most of the LVs are derived from HIV-1 that retain their capability to integrate into the host genome of infected cells (Liu and Berkhout, 2014). The backbone of LVs is composed of a wild-type HIV genome with all its genes and regulatory elements (Liu and Berkhout, 2014). The first-generation LV vectors consist of the envelope protein and the ‘psi’ packaging signal that is removed and incorporated into a heterologous promoter to reduce recombination potential (Ramezani and Hawley, 2002). The second-generation vectors have the accessory genes (*vif*, *vpr*, *vpu*, and *nef*) removed to reduce the virulence of any potential replication-competent retrovirus (Ramezani and Hawley, 2002). The transactivator gene, *tat*, is deleted in third-generation LVVs, dividing the vector into three plasmids to prevent future recombination possibilities and keeping only the three genes required for transgenic production (*gag*, *pol*, *rev*). Most recently, fourth-generation vectors split the *gag* and *pol* genes onto separate plasmids to reduce

recombination potential which leads to enhancement in transduction efficiency and transgene expression (Smith, 1995, Hall and Agarwal, 2018). Lentiviral vectors have gained immense attention for clinical applications due to their ability to transduce non-proliferating or slowly proliferating cells more efficiently, such as CD34⁺ stem cells (Merten et al., 2016). Gene transfer using LVs into CD34⁺ HSCs has been used to treat several genetic diseases such as β -thalassemia, X-linked adrenoleukodystrophy, metachromatic leukodystrophy, and Wiskott-Aldrich Syndrome (Nienhuis and Persons, 2012, Cartier et al., 2012, Sessa et al., 2012, Singh et al., 2017). The use of LVVs in research is still associated with potential oncogenic risks due to insertional mutagenesis, and the long-term safety of these clinical interventions is still being evaluated. Recently, Bluebird Bio suspended clinical trials for sickle cell anemia that involved the use of LV vectors after participants developed acute myeloid leukemia (AML) or myelodysplastic syndrome (MDS). Despite the engineering of packaging cell lines and deletions of genes necessary for viral replication, the limitations of using LVVs in clinical trials today are primarily brought on by the lack of adequate methods to produce high-titer virus stocks and the safety concerns related to their origin from HIV (Palu et al., 2000).

5. **Herpes Simplex Virus (HSV):** The dsDNA genome of the herpes simplex virus is more than 150 kb long, separated into long and short unique segments (UL and US, respectively), and is flanked by inverted repeating sequences (TRL/IRL and TRS/IRS) (Goins et al., 2014). About 90 genes are encoded by the HSV genome, of which 50 can be removed using recombinant vectors because they are not necessary (Burton et al., 2002). At present gene delivery and gene therapy applications use 3 types of HSV vectors (Burton et al., 2002):

- i. Amplicons, which are gene delivery systems packed through transfection into producer cells of plasmids containing the HSV origin of DNA replication, HSV

cleavage-packaging sequences, and transgene(s) of interest, make up the first system (Goins et al., 2004). To fit huge pieces of foreign DNA, the plasmids create head-to-tail linked concatemers that are packaged into the viral particles (Goins et al., 2004).

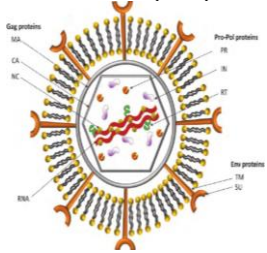
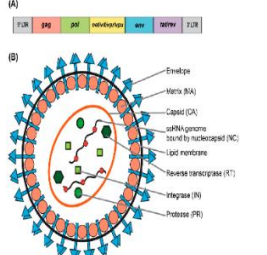
- ii. In a second system, the deletion of genes necessary for the lytic cycle of HSV can create less toxic, replication-defective vectors that elicit a smaller immune response than the wild-type HSV vectors (Efsthathiou and Minson, 1995). Deleting various combinations of the immediate early genes creates several versions of replication-defective HSV vectors (Efsthathiou and Minson, 1995).
- iii. The third system consists of replication-competent HSV vectors that are attenuated versions of HSV and contain genes that are required for replication *in vitro* but lack the genes necessary for replication *in vivo* (Manservigi et al., 2010). The set of deleted genes must produce the proper balance of replicative attenuation for safety and lytic activity for tumor toxicity or vaccination immunogenicity for replication-competent vectors to function as oncolytic therapies and vaccines, respectively (Manservigi et al., 2010).

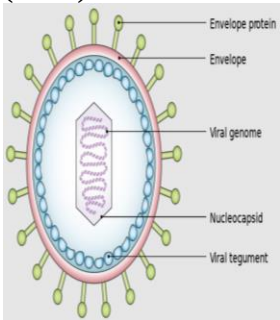
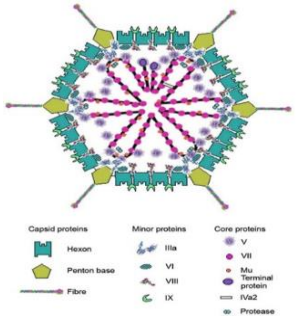
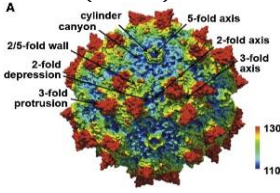
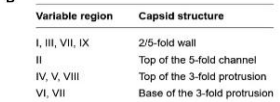
However, because encephalitis is spread by wild-type HSV, it is imperative to create benign replication-defective vectors with early gene deletions. The use of HSV-1 as a vaccination vector for the delivery of foreign antigens is also permitted and supported by recent advancements in the field of HSV-1 technology (Koelle and Corey, 2003). HSV vectors have several benefits for preventing viral infections (Lachmann, 2004). According to Manservigi et al. (2010), HSV is spread through direct touch, mucosal membrane exposure to bodily excretions, and inhalation of respiratory droplets. HSV remains a recognized human disease with inflammatory and immunomodulatory

effects, despite molecular changes to the viral genome drastically reducing virus-associated hazards (Manservigi et al., 2010).

The attributes, advantages, and limitations of these commonly used viral vectors are summarized below (Table 3). Although viral vectors can infect a wide variety of cells and show stable and high transgene expression, the use of viral vectors possess serious challenges. When using viral vectors, the general concerns involved are: i) limited packaging size; ii) immune reactions; iii) off-target effects; iv) inflammation; and v) insertional mutagenesis. There are high risks for viral vectors as they may elicit an immune response making the treatment less efficient and thus, resulting in the creation of antibodies that could preclude a second dosage of the same virus. Additionally, some viral vectors impart a risk of insertional mutagenesis as there is a risk of inserting into a tumor suppressor or upstream and activating a proto-oncogene.

TABLE 3: Major Viral vectors used in gene therapy.

Vector	Genetic Material	Packaging Capacity	Tropism	Inflammatory potential	Vector Genome Forms	Advantages	Limitations
Enveloped							
Retrovirus (RV)	RNA	8kb	Dividing cells only	Low	Integrated	Persistent gene transfer in dividing cells	Only transduces in dividing cells, integration might induce oncogenesis in some application
							
Lentivirus (LV)	RNA	8kb	Broad	Low	Integrated	Persistent gene transfer in most tissues.	Integration might induce oncogenesis in some application
							

Herpes Simplex Virus (HSV) 	dsDNA	40kb (Replication defective vector) 150kb (Amplicon)	Strong for neurons	High	Episomal	Large packaging capacity, strong tropism for neurons.	Inflammatory response, transient transgene expression in cells other than neurons.
Non-Enveloped							
Adenovirus (AV) 	dsDNA	8kb (Replication defective vector) 30kb (Helper Dependent vector)	Broad	High	Episomal	Extremely efficient transduction in most tissues	Capsid mediates a potent inflammatory response.
Adeno-Associated Virus (AAV)  	ssDNA	<5kb	Broad, except for haematopoietic cells.	Low	Episomal (>90%) Integrated (<10%)	Non-inflammatory, Non-pathogenic	Small packaging capacity.

(Adapted From : Tani et al., 2011; Chamcheu et al., 2015)

To overcome the disadvantages associated with the currently used viral vectors, anelloviruses open a new avenue to explore viral vectors that are a part of the human microbial flora but are not associated with any disease. Anelloviruses may represent an ideal vector for gene therapy as they are found in various cells and tissues of the human body and to date are not known to be etiological agents of any diseases. Their capability to travel in exosomes further makes them an ideal candidate for further study.

1.2.2.B. NON-VIRAL GENE DELIVERY

The non-viral vectors are comprised of naked DNA or synthetically produced biological particles bound to or encapsulated to a synthetic compound carrying the therapeutic gene expression cassette and are released upon reaching the targeted site (Chira et al. 2015). Physical methods or Chemical/Synthetics approaches can be used to attain gene therapy using non-viral vectors.

I. Physical methods: These methods involve the use of physical forces that counteract the membrane barrier of the cell thus facilitating the delivery of genetic material.

Different physical methods used for gene therapy are discussed below:

- 1. Microinjection:** Marshall A. Barber (1911) first described the technique of microinjection which involves a micropipette using which the nucleic acid can be directly injected into a living cell at a microscopic level (Barber, 1911; Chi et al., 2020). It is a simple mechanical approach that uses a needle approximately 0.5-5 μ m in diameter which penetrates the cell membrane or the nuclear envelope and finely injects the genetic material. This method is simple, economical, effective, non-toxic, reproducible and is particularly beneficial for delivering large DNA constructs. However, this approach has significant limitations as it requires individual manipulation of each cell and as such demonstrates low overall transfection rates. Microinjection holds high potential to be used in vaccination procedures where low transgene levels are required but may be sufficient to elicit an immune response (Wang et al., 2013).
- 2. Needle Injection:** This is one of the most simple and safe approaches which allows direct local injection of naked DNA into several tissue types such as liver, skin, brain as well as tumors (Al-Dosari and Gao, 2009). In a clinical trial carried out for patients suffering from myocardial ischemia, patients were needle-injected with the vascular endothelial growth factor-2 (VEGF-2) gene and the injection demonstrated some positive therapeutic

results along with improved heart functions (Losordo et al., 2002). Besides naked DNA, this approach could be applied to other nucleotides such as siRNA and DNA/polycations complexes (Wang et al., 2013). This method has attracted a lot of attention to study and evaluates DNA-based vaccination (Al-Dosari and Gao, 2009). Although this approach is simple and safe, the expression level of the therapeutic gene is low especially when the naked DNA is injected intravenously. Along with low gene expression, the DNA is susceptible to nucleases present in the blood (Wang et al., 2013).

3. Jet Injection: The jet injection is a needle-free approach and has been used as a drug delivery method since 1947. In this method, DNA in solution is driven by high-pressurized gases like CO₂ to form a high-speed ultra-fine stream that hits the targeted cells (Wendell et al., 2006). The pressure creates pores on the cell membrane which facilitates intracellular entry of DNA (Barolet and Benohanian, 2018). Jet injection leads to higher gene transfer efficiency compared to other needle injection methods (Barolet and Benohanian, 2018). Also, the pressure of gases could be easily controlled and can be adjusted to the levels of mechanical tolerance of the cells which reduces the chances of tissue damage (Barolet and Benohanian, 2018). The jet injection can be used for many tissues like muscle, skin, fat, and mammary tissues (Boulaiz et al., 2005). Promising results for cancer inhibition have been demonstrated by this approach (Barolet and Benohanian, 2018). To date, no major side effects have been reported for this method apart from minor effects such as hyperemia, edema, and minor bleeding (Lisakowski et al., 2003).

4. Gene Gun: The gene gun is also known as Ballistic DNA injection or DNA-coated particle bombardment and was initially designed for plant transformation but has since been used to transfer genes into cells and tissues like skin, muscle, mucosa, and surgically exposed tissues or tumor cells (Wang et al., 2013). In this approach, an elemental particle

of heavy metals like gold, tungsten, or silver is coated with plasmid DNA. These particles are then accelerated by pressurized gas and fired at the target cells or tissues. The momentum of the particles allows penetration around a few millimeters deep into the tissue and releases the DNA into the cell's path (Gao et al., 2007). Some parameters of the particles like size, speed, and dose play a pivotal role in transgene transfer efficiency and the degree of tissue injury (Wang et al., 2013). One of the biggest advantages of this approach is that it induces more immune responses with lower DNA doses and is very useful for DNA vaccines (Al-Dosari and Gao, 2009). The gene gun is a safe and simple gene transfer approach, but the transient transgene expression limits its use especially in actively dividing cells in which the high proliferation rate rapidly dilutes the transgene (Wang et al., 2013).

5. **Electroporation:** In this method, high-voltage electrical currents are used on the targeted cells which creates transient nanometric pores on the cell membrane which allows the negatively charged DNA to move inside the cells and remain trapped within (Wang et al., 2013). This approach was first used *in vitro* and *in vivo* in 1982 and 1991, respectively (Neumann et al., 1982; Titomirov et al., 1991). *In vivo*, electroporation has been found applicable in most of the cells tested for gene therapy such as muscle, liver, lungs, skin, and in myriads of tumor cells (Nishi et al., 1996; Mir et al., 1999; Gehl et al., 2002; Heller and Heller, 2006). In 2008, the first clinical trial for gene therapy using electroporation was reported (Daud et al., 2008). In this phase I trial, patients suffering from metastatic melanoma received plasmid interleukin, IL-12 into the tumor cells (Daud et al., 2008). The results of this study demonstrated that electroporation-mediated gene therapy is safe, effective, reproducible, and titratable modality (Daud et al., 2008). The efficiency of electroporation for gene therapy depends on various factors such as electric current intensity, the time interval between discharges, and concentration and type of DNA

(Wang et al., 2013). Studies so far have revealed that electroporation can be used to transfect almost all types of cells and is safe, effective, and consistent (Wang et al., 2013). In the case of electroporation, with optimized conditions, the transfection efficiency can even become comparable with those of viral vectors (Andre and Mir, 2004). This method can be used to deliver large DNA fragments (100-150 Kb) and produce long-lasting transgene expression (Canatella et al., 2001; Molnar et al., 2004). Although the approach holds various advantages, its use in in vivo treatment is limited due to some major drawbacks. One of the drawbacks is the limited range between the electrodes which restricts the gene transfer to a large area of tissues (Durieux et al., 2004; Gissel and Clausen, 2001). Surgical procedures are required to place the electrodes in the internal organs and applying high voltage across the tissue or internal organs may lead to tissue or organ damage (Wang et al., 2013). There is a possibility that high voltage may affect the stability of genomic DNA which further raises the safety concern (Wang et al., 2013).

- 6. Sonoporation:** Sonoporation, also known as ultrasound-facilitated gene transfer, is a procedure that uses ultrasonic waves to induce cell membrane permeabilization to allow intracellular gene entry (Wang et al., 2013). However, in every ultrasonic cycle part of the energy of the propagating waves is absorbed by the cells which in turn affects the cell membrane thus, affecting the gene transfer efficiency (Wang et al., 2013). The gene transfer efficiency can be improved by using contrast agents which are air-filled microbubbles that are stabilized by surface-active molecules like polymers or phospholipids (Endoh et al., 2002). On application of ultrasonic waves, the microbubbles start to oscillate rapidly, expand, shrink, and break up releasing local shock waves to transiently permeate the membrane of nearby cells (Endoh et al., 2002). The efficiency of this approach can further be enhanced by treating the cells with lidocaine or heat (Nozaki et al., 2003, Ogawa et al., 2004) and by using DNA/cationic lipid complexes instead of

naked DNA (Koch et al., 2000, Answer et al., 2000). The efficiency of sonoporation depends on ultrasonic frequency and intensity, duration of treatment, amount and the type of DNA used, contrast agents, and the type of tissue (Wang et al., 2013). The advantages of the approach include safety, non-invasiveness, and the ability to transfer genes into the cells of internal organs without involving surgical procedures (Sheyn et al., 2008). Although the gene transfer efficiency through sonoporation is much higher than other conventional naked DNA injection methods, the low gene transfer remains the major drawback of this approach (Wang et al., 2013).

7. **Photoporation:** Photoporation is a physical method that uses a single laser pulse that creates transient pores on the cell membrane allowing the DNA to enter the cell. The efficiency of gene transfer depends mostly on the focal point and pulse frequency of the laser (Ramamoorth and Narvekar, 2015). It has been claimed that the transgene expression of photoporation is like electroporation (Li and Huang, 2007). Recently, it was found that photoporation mediated by vapor nanobubble to have a more gentle and efficient approach for intracellular mRNA delivery in adherent and suspension cells, with promising potential for the future engineering of cells in therapeutic and fundamental research applications (Raes et al., 2020).
8. **Magnetofection:** The principle of this approach is based on magnetically targeted drug delivery (Al-Dosari and Gao, 2009). In this technique, the therapeutic gene is coupled with a magnetic nanoparticle, and the complex is introduced into the cell under the influence of a magnetic field created using rare, earth magnets which are placed beneath cell cultures to increase the sedimentation of complexes and transfection speed (Al-Dosari and Gao, 2009). The combination is delivered intravenously through *in vivo* gene therapy. The complex is attracted to and maintained at the target by the high-gradient external magnet (Dabson, 2006). The cross-linking molecule is subsequently cleaved by

an enzyme, a charge interacts, or the matrix degrades, releasing the genetic material (Dabson, 2006). The approach is mostly used for transfection of primary cells *in vitro* or cells that are difficult to transfect using another approach (Jones et al., 2013).

9. **Hydrodynamic Gene Transfer:** This technique uses high hydrostatic pressure as the driving force to deliver therapeutic genes to the internal organs (Wang et al., 2013). In pre-clinical studies carried out in mice, in which large volumes of DNA solution (more than 8% of body weight) were injected rapidly (around 5 seconds or less) into the tail vein, transgene expression was observed in liver, lung, kidney, and spleen, with liver showing the highest level of gene expression (Liu et al., 1999; Zang et al., 2000). Since hydrodynamic gene transfer is safe, simple, and effective, it can be used to deliver DNA, RNA, proteins, and synthetic compounds into various cells of small animals (Al-Dosari et al., 2005; Romoren et al., 2004; Liang et al., 2004; Bradley et al., 2005). The gene delivery efficiency is determined by the type of organs, injection volume, injection speed, and the concentration of the functional substance being added (Wang et al., 2013). However, the administration of a large volume of DNA solution is the biggest hurdle for its clinical applications in humans and large animals as rapid injection of solution equivalent to 8% of body weight would be intolerable (Wang et al., 2013).

II. CHEMICAL/SYNTHETIC METHODS

1. **Cationic Lipids/ Liposomes:** This approach is the most extensively studied and commonly used non-viral vector for gene therapy. The technique was first used by Felgner et al. (1987) who used a double-chain monovalent quaternary ammonium lipid to condense and transfer DNA into cultured cells which resulted in high expression of the delivered gene. All cationic lipids have a common structure that is composed of three parts: hydrophilic head group, hydrophobic anchor, and linker (Wang et al., 2013). The

hydrophilic head comprises one or more amine groups which include primary, secondary, tertiary amines, and quaternary ammonium salts. Cationic groups such as amidinium, guanidium, and heterocycle have also been utilized as hydrophilic heads (Ruysschaert et al., 1994; Vigneron et al., 1996; Solodin et al., 1995). Based on charges present on the head group, cationic lipids can be classified as monovalent or multivalent and the hydrophobic anchors are the non-polar moieties of the cationic lipids driving the formation of lipid/DNA complexes known as lipoplexes (Ramamoorth and Narvekar, 2015). They can be further categorized based on chemical structure: hydrocarbon-based chain, cholesterol-based chain, and Vitamin D-based chain (Ren et al., 2000). Most of the hydrocarbon anchors are represented by double-chain hydrocarbons that are connected to the linker via ester, ether, carbamoyl, or amide bonds (Wang et al., 2013). The transfection efficiency depends mainly upon the overall geometric shape, the number of charged groups per molecule, the nature of the lipid anchor, and linker bondage (Ramamoorth and Narvekar, 2015). Due to the positive charge on Lipoplexes, they interact with the negatively charged glycoprotein and proteoglycans of the cell membrane which facilitates the cellular uptake of the nucleic acids. The positively charged lipids surrounding the genetic material protect it against intracellular and extracellular nucleases. One of the major drawbacks of lipoplexes is the surface charge which reduces the half-life of lipoplexes in blood circulation thus limiting its utility to vascular endothelial cells (Al-Dosari and Gao, 2009). Neutral polymers like polyethylene glycol (PEG) can be used to surface shield the charges and increase the half-life of lipoplexes (Jones et al., 2013). Although this approach is considered to have low toxicity, it can become cytotoxic beyond the 3:1 ratio of Lipid:DNA (Jones et al., 2013; Ramamoorth and Narvekar, 2015).

2. Cationic Polymer Vectors: In this approach, cationic polymers like polyethyleneimine (PEI), chitosan, polyphosphoesters, etc. are mixed with DNA to form complexes known as polyplexes. Polyplexes are nanosized transfection units and are more stable compared to lipoplexes. The contributions of cationic polymers to transfection have been attributed to several capabilities: i) to enhance the DNA uptake via endocytosis through charge-charge interaction between polyplexes and the anionic sites on the cell surface; ii) to protect DNA from nuclease degradation; and iii) to facilitate DNA escape from endosomes. However, with the usage of different polymers, the transfection activity and toxicity may vary dramatically (Ramamoorth and Narvekar, 2015).

Non-viral methods provide advantages over their viral counterparts, such as relative safety and immunotolerance, the ability to transfer large-size genes, lower toxicity, and easy preparation. Also, non-viral vectors can be modified with ligands for tissue- or cell-specific targeting. However, non-viral methods remain limited by their low transfection efficiency rates and poor transgene expression. Table 4 summarizes the advantages and disadvantages of non-viral vector systems. Due to the poor transfection efficiency and low transgene expression, viral vectors are found to be more efficient vectors and continue to comprise most gene therapy clinical trials.

TABLE 4: Advantages and Disadvantages of non-viral vectors.

METHOD	ADVANTAGES	DISADVANTAGES
Microinjection	Simple, effective, reproducible, non-toxic, able to transfer large size DNA	Not suitable for transfection of large number of cells
Needle Injection	Simple, safe	Low efficiency
Jet Injection	Needle free, easy to control, safe	Low efficiency. Some cases result in local slight tissue damage.
Gene Gun	Safe, effective	Tissue damage
Electroporation	Highly effective, reproducible, potential for localized gene delivery, able to transfer large size DNA	Limited working range restricts gene delivery to large area of tissue. Surgical procedure is required to transfer genes into internal organs. The stability of genomic DNA might be influenced by

		high voltage.
Sonoporation	Safe, non-invasive, able to transfer genes into internal organs without surgical procedure, potential to disrupt blood-brain barrier.	Low efficiency
Photoporation	Safe, non-invasive	Low efficiency
Magnetofection	Flexibility. Low cytotoxicity.	Transient transfection.
Hydrodynamic Gene Transfer	Simple, reproducible, highly effective (especially for liver gene delivery)	Large injection volume is required which restricts its clinical application.
Cationic Lipids	Easy to be prepared, low cost, highly effective in vitro	Toxicity, low efficiency in vivo
Cationic Polymers	Easy to be prepared and chemically modified, low cost, highly effective	Toxicity. Some polymers are non-biodegradable.

(Adapted from: Ramamoorth and Narvekar, 2015; Gao et al., 2007; Wang et al., 2013)

1.2 ssDNA EUCARYOTIC VIRUSES

Recent advancement in metagenomics and bioinformatics tools has radically changed our understanding of viruses. With the help of these tools, various samples have provided detailed information about viral genome organization, regulation of gene expression, replication mechanism, and viral interaction with cellular hosts (Malathi and Renuka Devi, 2019). To determine the viral distribution across various ecosystems, different samples were analyzed from various environments such as marine or freshwater, atmospheric air, soil, hot springs, and habitats like insects, plants, human tissue, etc., revealing that most viral sequences are ssDNA viruses with either circular or linear DNA genomes (Malathi and Renuka Devi, 2019). Currently, 70,000 nucleotide sequences that are available in the NCBI database belong to ssDNA viruses (Malathi and Renuka Devi, 2019). The taxonomy of ssDNA viruses has undergone many revisions since 2015, as more and more sequences are available by metagenomic mining. At present, there are thirteen families established by ICTV, of which 11 families contain circular genomes, and two families (*Parvoviridae* (AAV) and *Bidnaviridae*) have a linear single-stranded DNA genome (Malathi and Renuka Devi, 2019). Out of 11, seven families infect eukaryotes - *Anelloviridae*, *Bacilladnaviridae*, *Circoviridae*, *Geminiviridae*, *Genomoviridae*, *Nanoviridae*, and *Smacoviridae*. The striking

feature of members of *Anelloviridae* is that they possess a negative sense of the ssDNA genome unlike other ssDNA viruses (Miyata et al., 1999). The ssDNA viruses infecting bacteria belong to the families *Microviridae*, *Inoviridae*, and those which are harbored by archaea are grouped under two families, *Pleolipoviridae* and *Spiraviridae*. Table 5 summarizes the ssDNA viruses (Malathi and Renuka Devi, 2019). The basic anchoring feature that unites most of the ssDNA viruses is the replication initiation protein (Rep), which shows a great degree of conservation (Malathi and Renuka Devi, 2019). The coat protein (CP), on the other hand, is incredibly changeable, and in the event of a circular genome, the arrangement of Rep/CP ORFs either divulges or converges towards the replication origin. This similarity between prokaryotic rolling circle plasmids and ssDNA viruses suggests that they have evolved together (Gutierrez et al., 2004). Most of the eukaryotic ssDNA viruses express rep and replicative via rolling circle replication (Gutierrez et al., 2004).

TABLE 5: Properties of ssDNA Viruses.

HOST	VIRUS TAXON	SUB FAMILY	GENERA	VIRUS MORPHOLOGY (SIZE IN nms)	GENOME MORPHOLOGY	GENOM E SIZE (kb)	REPLIC ATION
Bacteria	Microviridae	2	3	Icosahedral 30 nm	Circular	4.4-6.1	RC-Rep
	Inoviridae	-	7	Filamentous/rod shaped, 7 x 700–2000 nm	Circular	4.5-12.4	RC-Rep
Archaea	Pleolipoviridae (haloarchaea)	-	3	Pleomorphic 40 nm	Circular ss or ds	7–10.6	RC-Rep
	Spiraviridae (crenarchaea)	-	1	Coil-shaped Antenna-like projection 28 x 200 nm	Circular	24.9	Unknown
Eukarya	Anelloviridae (human)	-	12	Icosahedral 30–32 nm	Circular	2–4	RC-Rep
	Bidnaviridae (silkworm)	-	1	Icosahedral 25 nm	Linear segmented with terminal inserted repeats	6.0-6.5 per segment	DNA pol B
	Circoviridae	-	2	Icosahedral 20 nm	Circular	1.7–2.3	RC-Rep

(birds and animals)							
Geminiviridae (plants)	-	9	Icosahedral twinned 22 x 38 nm	Circular	2.5–3	RC-Rep	
Nanoviridae (plants)	-	2	Icosahedral 18–19 nm	Circular	0.9–1.1 per segment	RC-Rep	
Smacoviridae (fecal matter of mammals, dragonflies)	-	6	Icosahedral unknown Most probably icosahedral	Circular	2.2	RC-Rep	
Parvoviridae (vertebrates and arthropods)	2	5 and 8	Icosahedral 18–26 nm	Linear TIR	4–6.3	RHR-NS1, NS2	
Genomoviridae (one definite mycovirus several in mammals, birds insert, plants, sewage and sediments)	-	9	Icosahedral 20 nm	Circular	2.2	RC-Rep	
Bacilladnaviridae (diatoms)	-	3	Icosahedral 33–38 nm	Circular, partially ds.	5.8–6	RC-Rep	

KEY: - RC-Rep: Rolling circle Replication, RHR: Rolling hairpin replication, NS1 and NS2:

Non-structural proteins (Malathi and Renuka Devi, 2019).

These viruses have evolutionary relationships with prokaryotic viruses that have been extensively used for biotechnological applications. They hold immense potential to be used for different applications with the aid of gene editing and molecular biology techniques. Among these viruses, anelloviruses are attractive candidates as immunotolerant vectors for gene therapy.

1.3 ANELLOVIRUS

With new species being discovered regularly, metagenomics and bioinformatics advancements have provided numerous novel insights into the microbial world, particularly among viruses. Numerous recently discovered viruses have no known connection to any diseases. Such viruses are referred to as "orphans" in the field of virology because they don't have a "parent," in this case, a disease association. Within the family *Anelloviridae*, the so-called anelloviruses, a sizable number of human viral 'orphans' were discovered. The first human anellovirus, torque teno virus (TTV), was discovered in a Japanese patient with acute post-transfusion hepatitis in 1997 (Nishizawa et al., 1997). Since this initial finding, many different human anelloviruses have been detected and identified in a variety of clinical and environmental samples. These viruses are the most abundant eukaryotic viruses in the human virome indicating a stable co-existence and tolerance (Virgin et al., 2009). Mammalian anelloviruses have also been identified in a wide variety of animal families including *Aotidae*, *Callitrichidae*, *Canidae*, *Cercopithecidae*, *Cricetidae*, *Didelphidae*, *Equidae*, *Felidae*, *Hominidae*, *Indriidae*, *Leporidae*, *Molossidae*, *Muridae*, *Mustelidae*, *Otariidae*, *Phocidae*, *Phyllostomidae*, *Procyonidae*, *Suidae*, *Tupaiidae*, *Ursidae*, and *Viverridae* (Varsani et al., 2021). Apart from mammals, they have also been discovered in the feces of predators, including the South Polar skua (family *Stercorariidae*) and blood-feeding invertebrates of the families *Culicidae* and *Ixodidae* (Varsani et al., 2021). The productive analysis of virus-host interactions was, however, significantly hampered by the dearth of effective instruments. According to the International Committee on Taxonomy of Viruses (ICTV), the classification of the *Anelloviridae* family consists of 14 genera (Biagini, 2009). Recently, 16 new genera of anelloviruses have been added to the list of *Anelloviridae* family (Table 6) (Varsani et al., 2021).

TABLE 6: List of genera in the family *Anelloviridae*.

	GENUS
Previously established	Alphatorquevirus
	Betatorquevirus
	Deltatorquevirus
	Epsilontorquevirus
	Etatorquevirus
	Gammatorquevirus
	Iotatorquevirus
	Kappatorquevirus
	Lambdatorquevirus
	Mutorquevirus
	Nutorquevirus
	Thetatorquevirus
	Zetatorquevirus
	Gyrovirus
Newly established	Chitorquevirus
	Omegatorquevirus
	Omicrontorquevirus
	Pitorquevirus
	Psitorquevirus
	Rhotorquevirus
	Sigmatorquevirus
	Upsilon-torquevirus
	Xitorquevirus
	Aleptorquevirus
	Dalettorquevirus
	Gimeltorquevirus
	Hetorquevirus
	Tettorquevirus
	Wawtorquevirus
	Zayintorquevirus

(Adapted From: Varsani et al., 2021)

Among the known anellovirus genus, alphatorquevirus, betatorquevirus, and gammatorquevirus are the most studied. TTV belongs to alphatorquevirus genera and like TTV two more viruses: torque teno mini virus (TTMV) and torque teno midi virus (TTMDV) belonging to betatorquevirus and gammatorquevirus were discovered in 2000 and 2007 respectively (Takahashi et al., 2000; Ninomiya et al., 2007). All these viruses have similar genome structures but different genome sizes: 3.7-3.8 kb for TTV, 3.2 kb for TTMDV, and 2.8- 2.9 for TTMV (Biagini, 2009; Takahashi et al., 2000; Ninomiya et al., 2007). The most

remarkable feature of anelloviruses is their omnipresence. Since the first discovery of TTV in the blood of a Japanese patient showing symptoms of hepatitis of unknown etiology, since then, anelloviruses have been detected in individuals living on all continents (Nishizawa et al., 1997; Spandole et al., 2015). Anelloviruses have been detected in many biological samples: whole blood, nasal secretions, saliva, bile, feces, tears, semen, breastmilk, and urine (Deng et al., 2000; Goto et al., 2000; Inami et al., 2000; Itoh et al., 2000; Matsubara et al., 2000; Naganuma et al., 2008; Kapusinszky et al., 2012; Furuta et al., 2015). Recently, complete genome sequences of two novel TTVs were discovered in human brain tissue (Kraberger et al., 2020). The fact that anelloviruses are found in a variety of cell types may indicate that these viruses don't have a strong preference for any single cell type or type of tissue. There isn't currently a cell culture system in place to grow human anellovirus, and there isn't an animal model that could offer light on the interactions between the virus and the host. Finding the best animal model to research the human anellovirus is challenging. No *Mus musculus* (house mouse) Anellovirus isolates have been published to date (Nishiyama et al. 2014), hence no mouse model can currently be utilized to research the biology of anelloviruses, even though anelloviruses have been found in several mammals, including rodents.

1.3.1 GENOME REPLICATION AND TRANSCRIPTION OF ANELLOVIRUS

The genome of anelloviruses is composed of a single-stranded negative-sense circular DNA (ssDNA) which ranges from 2.0 to 3.9 kb (Miyata et al., 1999; Jones et al., 2005; Ninomiya et al., 2007). The genome contains overlapping open reading frames (ORFs) and an untranslated region (UTR) (Peng et al. 2002) (Figure 4). TTVs are thought to form non-enveloped icosahedral particles of roughly 30 nm diameter, as evidenced by the electron microscopy photographs of structures that might represent TTV particles in an infected cell

or TTV particles in immune complexes within an infected person (Itoh et al., 2000; Leppik et al., 2007). There is evidence that anelloviruses use the same rolling circle replication method as most circular DNA viruses, both of prokaryotic and eukaryotic origin (Rosario et al., 2012). The genome of anelloviruses possesses an untranslated region that forms hairpins to facilitate rolling circle replication (Peng *et al.*, 2002; de Villiers *et al.*, 2011). The genomes of TTV encode conserved motifs that resemble motifs of rolling-circle replication-associated proteins (Rep proteins) belonging to the *Circoviridae* family (Bendinelli *et al.*, 2001). Additionally, TTV genomes have been found in the human bone marrow and liver as double-stranded circular DNA, a typical replicative intermediate of rolling circle replication (Peng et al., 2002). Anelloviruses cannot produce their own DNA polymerase, hence the machinery of the host cell is completely necessary for anellovirus genome replication (Kakkola et al., 2007). Thus, ssDNA to dsDNA conversion and genome replication takes place in the nucleus.

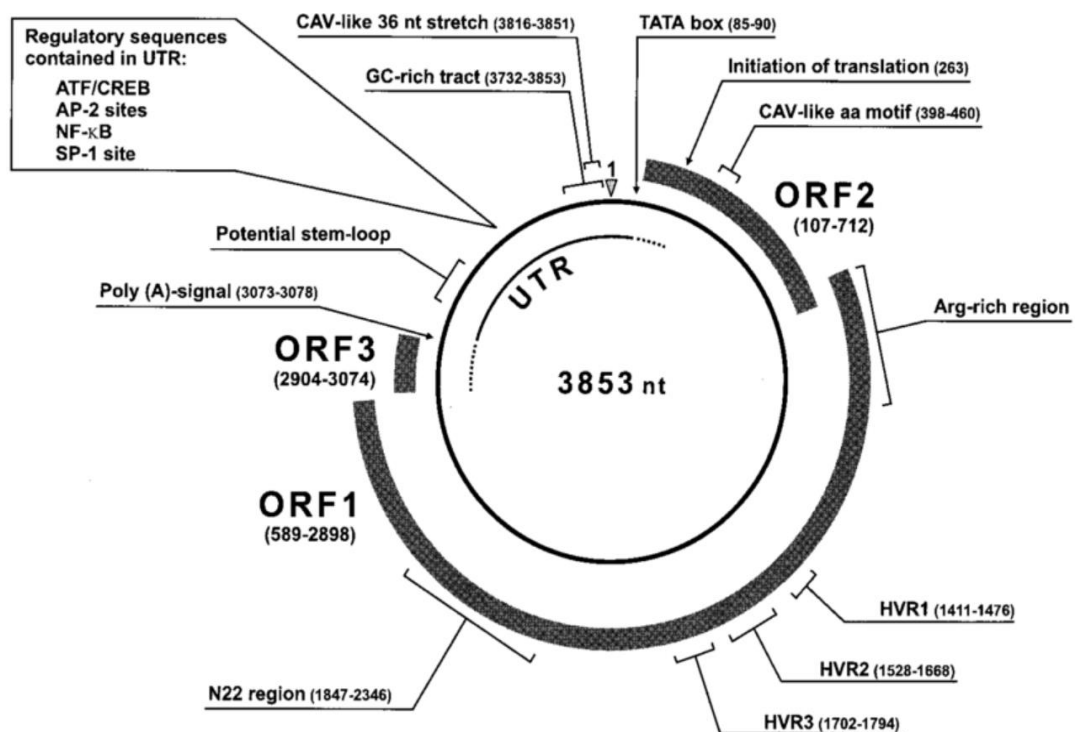


FIGURE 4: Genome organization of anellovirus TTV isolate TA278. (Taken from: Manzin et al., 2015). The TTV genome has a wide range of nucleotide sequences and consists of an untranslated region. Analysis of the DNA coding region revealed that there were at least

three-four partially overlapping open reading frames as well as a hypervariable region (HVR) and an N-terminal arginine fragment. The ORF1 of anelloviruses are involved in capsid production while ORF2 proteins are involved in production of regulatory proteins. The role of other ORFs is still unknown.

The transcriptional profile of anelloviruses has been studied by transfecting TTV DNA into cell cultures and in bone marrow cells derived from TTV-positive individuals (Kamahora et al., 2000; Qiu et al. 2005; Kakkola et al. 2009). During TTV infection, three mRNA species are generated. The longest transcripts are 1.2 kb and 1.0 kb long, while the biggest mRNA can be between 2.8 and 3.0 kb long. Since all mature mRNAs have the same 5' and 3' ends, splicing only happens on one lengthy pre-mRNA. (Qiu et al., 2005). By beginning translation at two separate AUG codons, the longest mRNA is employed to translate both the ORF1 and the ORF2 proteins (Qiu et al., 2005). The smaller mRNAs also undergo translation from two beginning AUG codons, so translation can result in the production of 6 proteins. The ORF2/2 and ORF1/1 proteins are produced by the 1.2 kb mRNA, which is the least abundant of the mRNAs, while the ORF2/3 and ORF1/2 proteins are produced by the 1.0 kb mRNA (Qiu et al. 2005).

The ORF1 protein, which is the largest one encoded by TTV and is found in the cytoplasm of cells transfected with a molecular clone of a TTV strain, suggests that the protein has a structural function (Qiu et al., 2005). The ORF1 protein has a sequence of arginine repeats in its N-terminal that is extremely similar to the ARM motif discovered in the Cap proteins of circoviruses, which is known to have DNA binding properties (Erker et al., 1999; Sarker et al., 2016). This characteristic may suggest that the ORF1 protein, similar to the Cap proteins of circoviruses, is essential for the packing of ssDNA (Sarker et al., 2016). According to Zheng et al. (2007) and Kakkola et al. (2009), the ORF2 protein is most likely a regulatory protein that aids the virus in suppressing the host's immunological responses. While ORF1/1 and ORF1/2 proteins are similarly distributed in the cytoplasm and nucleus of transfected

cells, ORF2/2 and ORF 2/3 are localized in the nucleus (Qiu et al. 2005). Because phosphorylated ORF2/2 proteins have distinctive serine-rich domains at the C-terminus, they likely can bind DNA templates (Asabe et al. 2001). This characteristic points to a function for this protein in the control of gene expression and genome replication (Kakkola et al. 2009).

1.3.2 POTENTIAL INTERACTION OF ANELLOVIRUS WITH THE HOST IMMUNE SYSTEM

TTV is known to interact with many pathogen-associated molecular patterns (PAMP) receptors that stimulate immune and inflammatory responses (Freer et al., 2018). *In vitro*, studies have demonstrated that the ORF 2 protein of TTV suppresses the activity of NFκB (Nuclear factor kappa-light-chain enhancer of activated B-cell) by inhibiting its translocation to the cell nucleus and, thus, its ability to activate transcription of genes, like IL-6, IL-8, and cyclo-oxygenase 2 (Zheng et al., 2007). Moreover, the genome of TTV and its replicative intermediates may trigger toll-like receptors (TLR) in infected cells which leads to the synthesis of pro-inflammatory molecules. This is due to the presence of unmethylated heterodimers of guanosine and cytosine (CpGs) in bacterial and viral DNA but is absent in mammalian DNA and foreign DNA (Freer et al., 2018). One pattern recognition receptor (PRR), toll-like receptor 9 (TLR-9), is specialized to distinguish CpGs (Freer et al., 2018), and depending on the number of nucleotides flanking CpGs, it may stimulate the expression of inflammatory cytokines, such as IFN-α, Interleukin (IL)-6, and IL-12 (Krieg, 2002). Both stimulatory and inhibitory CpGs are present in the DNA of TTV, and most organisms and their relative frequency may contrast impressively, indeed inside strains of the same species, hence likely affecting the way they interact and invigorate TLR-9 (Freer et al., 2018). TTV genogroup 4, which is high in CpG motifs, has been identified at higher levels in pediatric patients with bronchopneumonia compared to those with milder acute respiratory infections

and was found to stimulate TLR-9 in mouse spleen cells in vitro, causing inexhaustible generation of pro-inflammatory cytokines (Zheng et al., 2007).

MicroRNAs (miRNA) are ~22 nucleotides of small, single-stranded, non-coding RNAs delivered by the hosts and pathogens (Freer et al., 2018). They are potent modulators of pathogen recognition and host defense in a vast array of cellular metabolic pathways. For microbe-host interaction, cellular miRNA appears to regulate immune responses and inflammation and to play a direct antiviral role by blocking the translation of viral genes, counteracting persistent replication. It has also been demonstrated in animal models that miRNA can polarize macrophages toward allergic reactions (Zhou et al., 2017). Because miRNAs may both upregulate and downregulate inflammation in a variety of disorders, including asthma, their role in inflammation is likely to be complicated (Dissanayake and Inoue, 2016). Viruses, particularly small ones like TTV, produce their own miRNAs, which work in tandem with viral proteins to control viral gene expression, replication, pathogenicity, immune evasion, and the entire process of virus-related inflammation (Freer et al., 2018). It's worth noting that both cellular and viral miRNAs have been discovered to transport information to distant cells via plasma exosomes (Freer et al., 2018). TTV was also discovered to encode *in vivo* miRNAs that could be implicated in viral immune evasion as well as the control of IFN signaling (Kincaid et al., 2013). As measured by TTV viremia levels, there was no correlation between production and release and virus replication. Notably, TTV miRNA profiles vary based on the health condition of individuals, as did the types of miRNAs produced within sick patients (Vignolini et al., 2016). TTV miRNAs' role and relevance are unknown, and more research is needed. Figure 5 depicts an overview of the processes used by TTV to activate or relax the host's innate and adaptive defenses (Taken from: Freer et al., 2018).

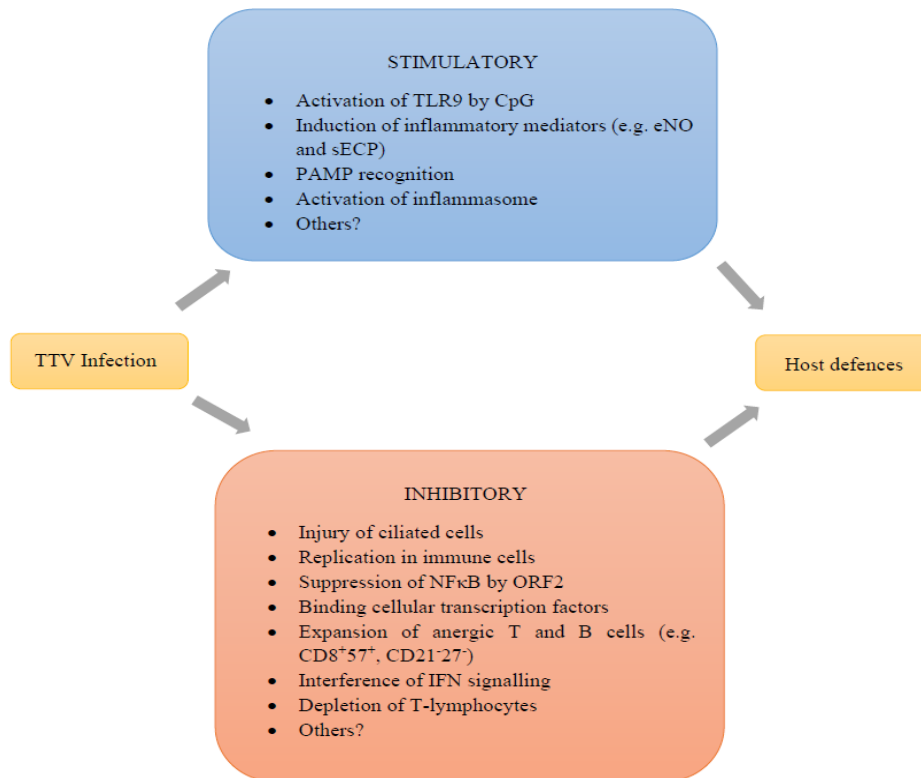


FIGURE 5: Possible mechanisms whereby TTV infection may modulate host defences. TLR, toll-like receptor; CpG, unmethylated heterodimers of guanosine and cytosine; eNO, exhaled nitric oxide; sECP, secreted eosinophil cationic protein; PAMP, pathogen-associated molecular pattern (Adapted from Freer et al., 2018).

1.3.3 ANELLOVIRUS AS A VECTOR FOR GENE THERAPY

Anelloviruses are still poorly understood and extensive research is still required but at present, there is no system or animal model available to grow these viruses *ex vivo*. Attempts have been made to culture anelloviruses and obtain cell-free viruses, but to date, no cell lines have been identified to obtain cell-free virus particles (Kaczorowska and Hoek, 2020). Nonetheless, the limited *in vitro* studies that have been conducted have shown that anelloviruses may be an ideal vector for gene therapy. For instance, the ORF 2 protein likely suppresses the activity of NFκB which is crucial for many genes related to the inflammatory response. ORF2 influences the activity of NFκB by inhibiting its translocation to the nucleus leading to its inability to activate transcription genes such as IL-6, IL-8, and cyclo-oxygenase

(Zheng et al., 2007). ORF 2 also produces the protein called TTV-derived-Apoptosis-Inducing protein (TAIP), which is 105 amino acids long and has been shown to induce apoptosis in cellular carcinoma cells (Kooistra et al., 2004; Bostan et al., 2013). Additionally, anelloviruses circulate in the body within exosomes which may promote infectivity by allowing entry of viruses in non-permissive cells, and viruses inside the exosomes are less susceptible to the neutralizing effects of antibodies (Martelli et al., 2018) and permit potential for redosability. Whether exosome encapsulation is required for effective gene delivery remains to be determined.

1.4 FILAMENTOUS PHAGES

Bacteriophages or phages are viruses that have bacterial hosts and are the most abundant organisms in the biosphere (Smeal et al., 2017). While temperate phages may lay latent as prophages within lysogenic bacteria for several generations before initiating a lytic cycle, lytic phages must destroy their hosts after infection (Clokier et al., 2011). Filamentous phages that belong to the family *Inoviridae* however, go beyond the standard dichotomy of viral lysis and lysogeny. The majority of characterized inoviruses continue to grow from their hosts constantly and maintain active infections throughout several bacterial generations without killing their hosts. All filamentous bacteriophages have certain features that distinguish this group. They are usually between 1-2 μm in length and 6-7 nm in diameter, and all contain a ssDNA genome (Clokier et al., 2011). The virion length of the phage is determined by the genome size, thus the larger the genome, the longer the phage becomes. The phage coat is composed of five proteins – two sets of two different minor coat proteins which are located at either end of the bacteriophage (Clokier et al., 2011) and a dominant major coat protein in the structure that forms the tube-like structure (shingle-like helix) that encloses the ssDNA genome (Clokier et al., 2011). Figure 6 shows the basic structure of filamentous bacteriophage (Yacoby and Benhar, 2008). The best studied of these are the Ff phages that infect male (F+)

strains of *Escherichia coli* (Yacoby and Benhar, 2008). Members of the Ff group include filamentous phage f1, fd, and M13, which are all largely similar (Yacoby and Benhar, 2008). Filamentous bacteriophages found in this group infect F⁺ *E. coli* by recognition and attachment to the tip of the F conjugative pilus. As with all other filamentous phages, infection with Ff bacteriophage does not lyse cells (Yacoby and Benhar, 2008) but rather extrudes through the cell membranes. The nanoscale size, simple life cycles, and speed and ease with which they can be prepared, manipulated, and characterized make phages attractive and versatile biotechnological systems (Smeal et al., 2017). The life cycle of filamentous phages is almost identical (Smeal et al., 2017). M13 phage, which is the most studied and exploited Ff is our phage of primary interest and will be discussed here in greater detail.

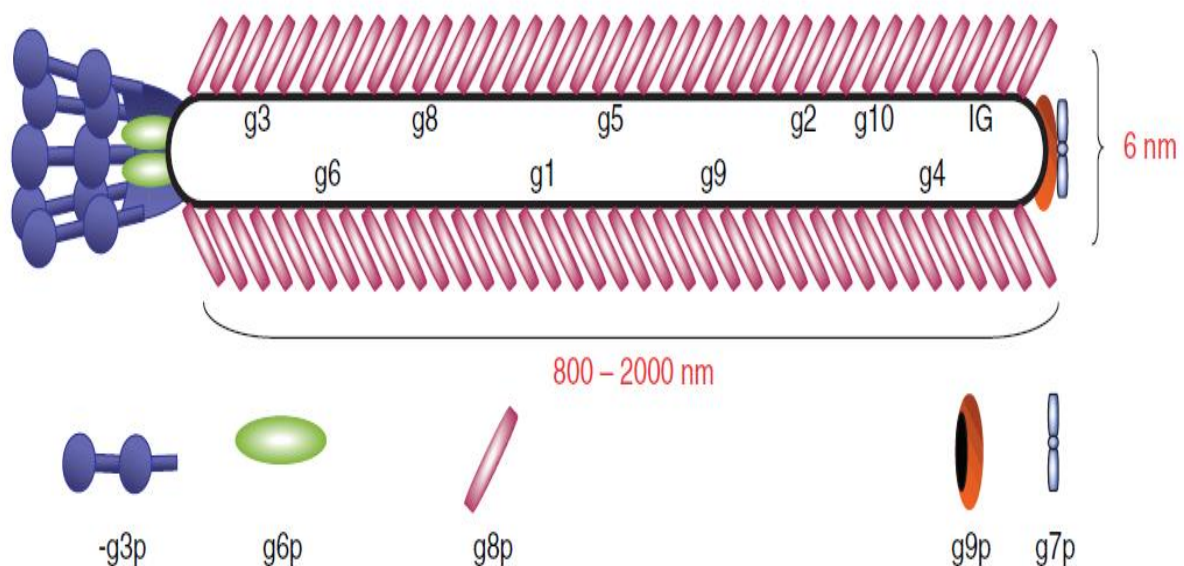


FIGURE 6: Structure of the filamentous bacteriophage. The diagram shows a single-stranded DNA core (thick black line) surrounded by a proteinaceous coat. The positions of the 10 phage genes and the intergenic region that contains the phage origin of replication are marked.

1.4.1 LIFE CYCLE OF M-13 FILAMENTOUS (Ff) PHAGE

M13 is a filamentous phage that infects *E. coli* carrying the F-episome. Active infection by M13 does not kill the host cell but rather the phage particles are continuously produced and extruded through the cell membrane as the infected cell continues to grow and divide (Young et al., 2000). The M13 phage particle consists of a single-stranded DNA (ssDNA) genome surrounded by approximately 2700 copies of a major coat protein, p8, and 5 copies each of 4 minor coat proteins: p3 and p6 on one end and p7 and p9 on the other (Figure 7A) (Smeal et al., 2017). The M13 genes and corresponding proteins are named 1–11 and are referred to as g1-g11 and p1–p11 and the genome is organized in two separate transcriptional units (Figure 6B) (Vanwezenbeek et al., 1980, Smeal et al., 2017). There are five promoters, and three terminators that produce a cascade of mRNA species that contribute to the regulation of phage protein concentrations (Edens et al., 1978; Moses et al., 1980). An extended non-coding region in the M13 genome includes a promoter, a terminator, the origins of positive and negative strand DNA synthesis, as well as the signal that is required for packaging of the ssDNA genome into a phage particle (Figure 7B) (Dotto et al., 1981, Smeal et al., 2017).

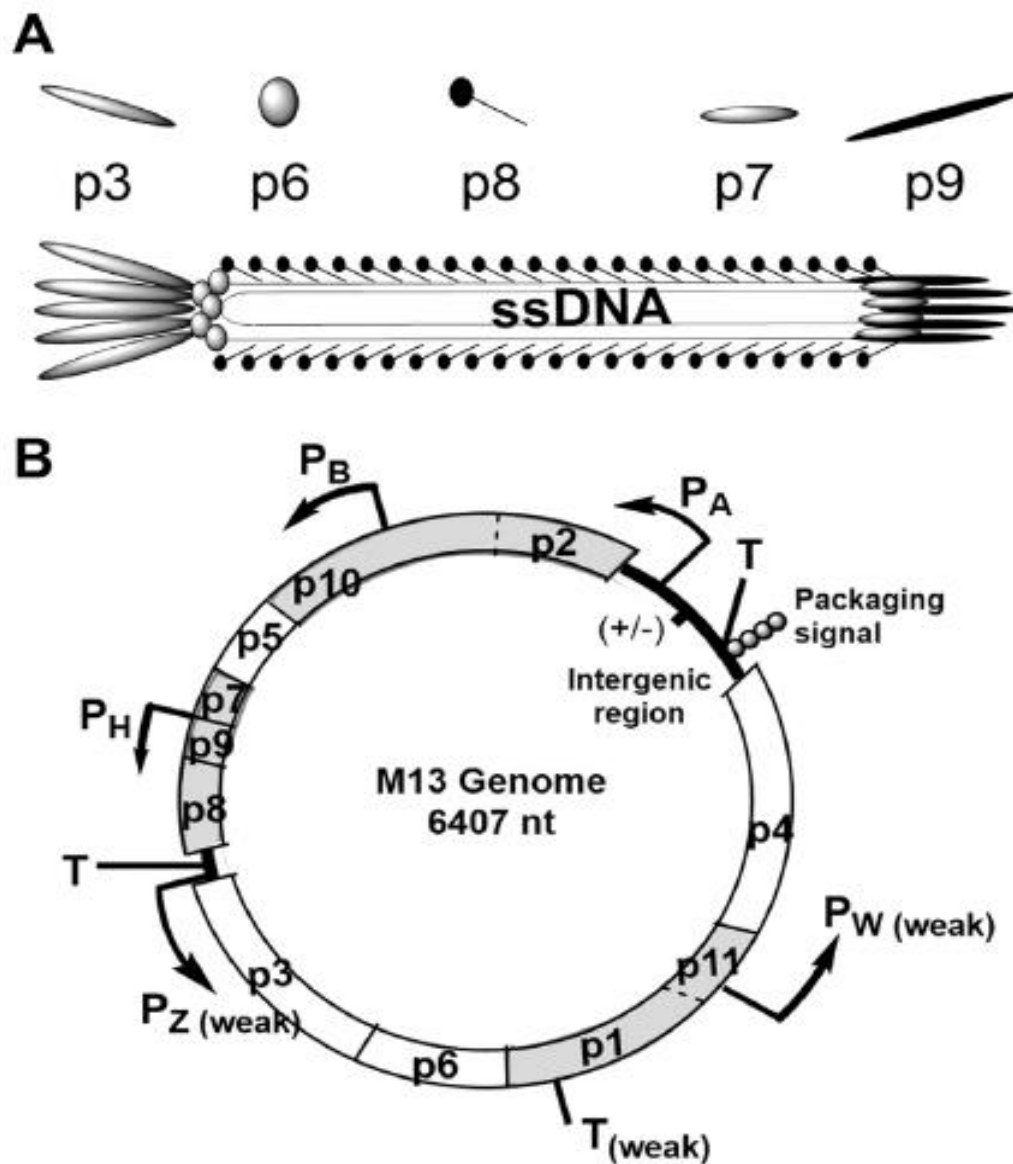


FIGURE 7: A) Schematic representation of M13 phage particle depicting the organization of coat proteins around the ssDNA genome. B) Map of the M13 phage genome identifying all genes, the intergenic region, promoters, and terminators. Coding regions for each M13 protein are labelled “p#”. Shaded regions represent portions of the genome in which the genes overlap (e.g., the gene for p10 is completely contained within the coding region for p2). P_A, P_B, P_H are strongly active promoters. Two additional weak promoters are labelled P_Z and P_W. Two terminators (T) are strong, rho-independent terminators; the third, weaker terminator (T_(weak)) is rho-dependent (Smeal et al., 2017).

The M13 life cycle begins with the passage of the phage genome into a host cell, which is mediated by the coat protein, p3. The exact process of phage genome entry into the host is

poorly understood (Bennett et al., 2011), but upon entry, the ssDNA genome is rapidly converted into a double-stranded form by host replication enzymes. The double-stranded replicative form DNA (RF DNA) immediately initiates mRNA transcription from the 5 constitutive promoters, producing mRNAs and directing the synthesis of all 11 phage-encoded proteins. Additional single-stranded copies of the M13 genome are made through rolling circle replication using the RF DNA as a template. While concentrations of phage proteins are low, the single-stranded copies are converted into additional RF DNA. As the concentrations of phage proteins, particularly p5 increase the rate of synthesis of RF DNA from ssDNA decreases. The p5 protein then binds to the ssDNA copies which prevents the conversion into RF DNA thus, sequestering the ssDNA genomes for packaging into progeny phages. The p5 sequestered ssDNA is recognized by the membrane-spanning phage assembly complex made up of phage proteins: p1, an inner membrane-spanning protein, p4, an outer membrane-spanning protein, and p11, an inner membrane-anchored periplasmic protein (Haigh and Webster, 1999; Marciano et al., 1999). Through a poorly understood mechanism, the assembly complex first attaches the minor coat proteins p7 and p9 to the p5 sequestered ssDNA and then proceeds to pass the ssDNA through the cell membranes, concomitantly removing p5 and assembling p8 around the ssDNA genome. The process is completed by the addition of the other minor coat proteins p3 and p6, and the particle is released from the cell (Bennett et al., 2011, Smeal et al., 2017) (Figure 8).

translation through variations in the strength of ribosome binding sites. Several nested feedback mechanisms are used in the steps of protein synthesis and DNA replication to regulate the number of M13 genomes in the infected cell (Smeal et al., 2017). Phage proteins p2, p5, and p10 play a key role in directing the M13 life cycle by controlling levels of phage DNA replication, deciding the fate of newly synthesized DNA, and controlling the expression of phage proteins (Smeal et al., 2017). DNA replication is initiated by the phage protein p2, which is inhibited by protein p10 (Smeal et al., 2017). A replicating ssDNA genome may either be converted into double-stranded forms to serve as a template for transcription and DNA replication, or it may be sequestered by p5, directed into phage assembly machinery, and packaged into progeny M13 particles. Besides modulating the translation of itself, the ssDNA binding protein also affects the translation of p1, p2, p3, and p10 (Smeal et al., 2017). The levels of phage DNA and proteins produced adjust as the infected cell grows and divides.

1.4.2 USE OF FILAMENTOUS PHAGES FOR GENE THERAPY

In recent years the evolution of phage technology has favored bacteriophage applications in drug and gene delivery studies. This is largely driven by the robust ability of Ff to display peptides, proteins, or antibodies on their surface through the phage display technique (Yacoby and Benhar, 2008). Phage display was first mentioned in 1985 (Smith, 1985) as a molecular method for creating polypeptides with distinctive properties. The target protein is expressed on the surface of the phage particle by translationally fusing the DNA encoding the polypeptide with the genes for the phage coat protein (N or C terminal) (Smith, 1985; Sidhu, 2000). Other phages including Lambda, T4, and T7 are also utilized as phage display systems, in addition to the filamentous phage M13 of *E. coli* (Benhar 2001). It is possible to screen and isolate highly specific peptides with an affinity for the target proteins using phage display libraries. Phage display has opened the great potential for phages to be used as

therapeutic gene delivery vehicles (Suwan et al., 2019). The theory behind utilizing phages as DNA vaccines, in which the phage coat shields the DNA within from destruction after it has been injected into the mammalian host and as it circulates to its target, is similar to the theory behind using phages for targeted gene delivery. The two strategies, however, differ philosophically. Phages can target particular cell types, which is necessary for tissue-specific gene delivery since they can show foreign proteins on their surface (Clark and March, 2006). Targeting and processing compounds are displayed on the surfaces of phages via phage display and artificial covalent conjugation (Larocca et al., 1999). Peptide sequences like fibroblast growth factor have been utilized to target cells expressing the right receptors for the targeted delivery of phages (Hart et al., 1994; Sperinde et al., 2001). Protein sequences like the adenovirus penton base, which facilitates entry, attachment, and endosomal release, have also been employed with some degree of effectiveness to improve the uptake and endosomal release of phages (Piersanti et al., 2004). Additionally, the simian virus 40 (SV40) T- antigen nuclear localization signal and the human immunodeficiency virus (HIV) tat protein's protein transduction domain have been employed to improve the absorption and nuclear targeting of modified phages like lambda (Nakanishi et al., 2003). The Tat consists of a protein transduction domain that allows Tat to enter the cell by crossing the cell membrane (Nakanishi et al., 2003). The NLS within the domain further mediates the translocation into the nucleus. On the other hand, SV40 T-antigen interacts with importin α to be imported into the nucleus (Bakhshinejad and Sadeghizadeh, 2014). Additionally, integrin-binding peptides that improve binding and uptake and DNase II inhibitors that reduce DNA degradation are exhibited peptides that can help in gene transfer by phages (Hart et al., 1994; Sperinde et al., 2001). To screen the ability of phages to target specific cells and tissues, phage display libraries have been used in several studies, and in all cases, phages were found in specific tissues (Rajotte et al., 1998). For instance, after isolating phages that target the liver, Clark

and March (2006) inoculated mice with phage display libraries, and concentrated phages were isolated in the organ after extracting the livers of the animals. Similarly, in vitro strategies have been used for the isolation of phage-displayed peptides that enhanced cytoplasmic uptake into mammalian cells (Ivanenkov and Menon, 2000).

1.4.3 ADVANTAGES AND LIMITATIONS OF FILAMENTOUS PHAGES APPLICATION TO GENE THERAPY

Filamentous phages like M13 are evolving as relevant vectors for the safe and targeted systemic administration of transgenes (Hajitou et al., 2008, Pranjol and Hajitou, 2015). They have several advantages over traditional viral and non-viral vectors. First, their protein coat consists of a repeating protein unit arranged in an alpha-helical array which allows the phage to package (in theory) unlimited DNA capacity, as the capsid coat merely needs to elongate to accommodate the transgene. Hence, the phage is efficient at condensing and packing the DNA. Second, a large variety of peptide ligands can be added with ease due to the protein coat's high tolerance for translational fusions, resulting in ligand-directed transduction of the target cell type. They are also considered safe because they have been used for many years to treat bacterial illnesses in people and because their use in food preparation has been authorized by the Food and Drug Administration (FDA-USA) (Lang, 2006). Third, phage vectors are scalable at high titers and are produced at relatively low marginal cost, which is desirable for industrial processing. Despite significant advantages for phages to be used as a vector for gene therapy application, the system is, however, posed with several obstacles. Phages have low stability in blood circulation as they are immunogenic in nature and are rapidly cleared by the reticuloendothelial system (RES) (Azam et al., 2021). Compared to other nonviral and viral delivery vector systems, phages remain poor delivery vectors as they lack intrinsic strategies for mammalian cell transduction (Donnelly et al., 2015). Therefore, new strategies need to be explored to allow efficient phage binding, entry into mammalian

cells, and subsequent increase of their gene transferability, which should subsequently develop their use in a wide range of gene transfer applications. Perhaps a safer more effective approach is to use eucaryotic packaging of ssDNA via viruses such as TTV that are a part of the human microbiome to deliver therapeutic transgene(s) of interest to a diverse range of cell types.

2. RATIONALE, HYPOTHESIS, AND OBJECTIVES

2.1 RATIONALE

Gene therapy (GT) holds immense potential for the treatment and potential cure of many acquired and inherited diseases. However, the success of GT depends greatly on the effective delivery of the genetic material to the relevant tissue(s). The efficiency of vector transport and unloading of genetic material into target cells, perseverance, activity, and immunological responses are the primary technological limits or critical considerations in achieving successful gene therapy. In addition to the technological limitations, cost-effectiveness, and ethical and social considerations also must be considered in vector development. Also, the high cost of manufacturing viral vectors leads to another bottleneck in production and utilization. For instance, Glybera, used to treat Lipoprotein Lipase Deficiency (LPLD), was withdrawn from the market in 2017 due to high production costs that compromised cost-effectiveness and hence, demand. Although both viral and non-viral vectors demonstrate some of the ideal characteristics for gene therapy, no single vector is a complete package without limitations, which presents a significant hurdle for genetic medicine. Novel approaches to vector generation seek to offer high transfection efficiency, transgene expression, and to be safer and more efficiently produced. Among the various unconventional options available, one is the use of prokaryotic filamentous phages or phagemids. While phages are capable of cell-targeting by displaying peptides, proteins, or antibodies on the surface through phage display techniques (Yacoby and Benhar, 2008), this approach generally results in low transfection efficiency. Anelloviruses, prolific human eukaryotic viruses, on the other hand, that share the same genomic topology of ss circular DNA as filamentous phages may prove to be highly suitable candidates for GT as they are not associated with any disease and can infect various cell types and tissues without stimulating an immune response. Their natural presence as part of the human virome indicates that they

are well-tolerated and as such, may be redosable and titratable in their application to genetic medicine.

In this project, the M13 filamentous phage system in *E. coli* will be used to replicate the transgene circuit and another *E. coli* strain will be used to produce anellovirus capsid protein. The project will further focus on the packaging of ssDNA by anellovirus capsid protein in vitro. If production and packaging can be achieved, this will overcome one of the largest hurdles in the growth and development of this virus as a potent GT platform. Anelloviruses can circulate in the body within exosomes which may promote infectivity by allowing entry of viruses in non-permissive cells, and viruses inside the exosomes are less susceptible to the neutralizing effects of antibodies (Martelli et al., 2018) and permit potential for redosability. Whether infectivity requires exosome transmission is currently unknown and it is acknowledged that this may be a limitation of this production method.

2.2 HYPOTHESIS

Minimal circular ssDNA molecules encoding a transgene cistron of interest and carrying a packaging signal for anellovirus can be produced and purified as M13 miniphagemids from host *E. coli* cells. In tandem, His-tagged anelloviral capsid proteins can be expressed and functionally processed in *E. coli* to yield self-assembling homo-multimeric anellovirus capsids. M13-processed ss circular DNA vector in combination with *E. coli*-produced capsids will self-assemble into circular ssDNA containing anellovirus capsid gene delivery vectors as a novel synthetic method for the assembly and production of anelloviruses.

2.3 OBJECTIVES

The main goal of this project is to produce ssDNA using the replication system of filamentous phage M13 that will replicate a minimal transgene circuit containing the anellovirus DNA binding signal which in turn will be recognized by anellovirus capsid protein produced in *E. coli* in vitro. The specific objectives are:

1. Identification and evaluation of anellovirus capsid protein and DNA-binding domain sequences. Selection of a suitable candidate for further production and purification.
2. Construction of relevant plasmids:
 - a. a specialized packaging plasmid containing the gene of interest and DNA binding domain for transgene production,
 - b. a tagged capsid protein plasmid containing the gene for the candidate anellovirus capsid protein designed for protein capsid protein purification.
3. Production and purification of the vector components in *E. coli*:
 - a. the ssDNA phagemid DNA packaging vector containing anellovirus DNA binding signal
 - b. the candidate anellovirus His-capsid protein.
4. Assessment of the ability to successfully package the minimal ssDNA vector carrying the packaging signal for anellovirus capsid in a cell-free *in vitro* packaging system.

3. MATERIALS AND METHODS

3.1. Strains, vectors, and primers.

E. coli BL21 strains were used in the production of anellovirus capsid proteins. *E. coli* JM109 was used for DNA vector plasmid amplification and production of M13 phagemids that packages the ssDNA containing anellovirus packaging signal. All the bacterial strains, plasmids, and primers used in the study are listed in Tables 7, 8, and 9 respectively.

TABLE 7: Bacterial strains used in the study

Strain	Genotype	Source
JM109	F' traD36 proAB+ lacIq lacZ_M15/_ (lac-proAB) endA1 glnV44 thi-1 e14 ⁻ (McrA ⁻) recA1 gyrA96 relA1 hsdR17	NEB
BL21	F ⁻ ompT gal dcm lon hsdS _B (r _B ⁻ m _B ⁻) [malB ⁺] _{K-12} (λ ^S)	NEB

NEB: New England Biolabs

TABLE 8: Plasmids used in the study

Plasmid	Genotype	Source
pET-28_TTV19Cap	T7 promoter, lac operator, 6X His-tag, TEV cleavage site, TTV19 ORF1	This Study
pTTV19 IR_CMV_gfp_M13	M13 origin, TTV19 intergenic region, CMV promoter, <i>gfp</i> gene inserted between M13 initiation and termination sequence, Amp ^R	Wong, 2022
M13SW8	M13KO7, Packaging signal removed, Kn ^R	Wong, 2022

Table 9: Primers used in this study.

Primer	Sequence
gfp-F Forward	5'CAAGATGAAGAGCACCAAAGG3'
gfp-R Reverse	5'CGAAGTGGTAGAAGCCGTAG3'

3.2. Plasmid design and construction.

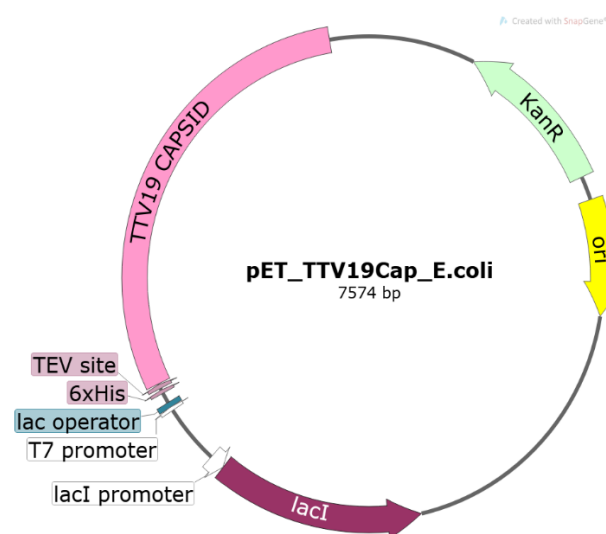
TTV19 was selected as a suitable candidate for the capsid protein expression in *E. coli*. The ORF1 of TTV19 is involved in the production and assembly of capsid protein. The ORF1 sequence of TTV19 was obtained from NCBI (GenBank: AB025946.2) and was codon-

optimized for protein production in bacteria. The codon was optimized using Gensmart™ codon optimizing tool software. Commercially, no anellovirus antibodies are available, so a 6X His-Tag (encoded by pET vector) was added to the N-terminal of TTV19 ORF1 along with the TEV cleavage site (Fig. 9A) to be used for both purification as well as identification of capsid protein and additionally, the removal of the His-tag is possible, if interference with capsid assembly occurs. The sequence was synthesized and added to the pET-28 vector (Genscript. Piscaway, USA).

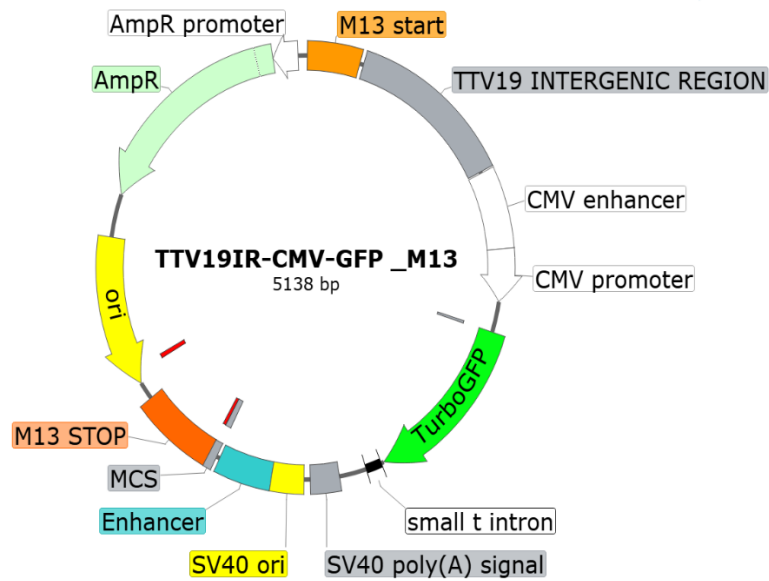
The TTV19 intergenic region containing a DNA binding signal was identified from the NCBI database (GenBank: AB025946.2). The intergenic sequence was synthesized (Genscript. Piscaway, USA) and introduced into the pCMV_gfp plasmid (Wong, 2022), and the construct was named pTTV19 IR_CMV_gfp_M13 (Fig. 9B).

The helper plasmid pM13SW8 (Wong, 2022) was used to produce and package ssDNA containing TTV19 intergenic region from the DNA vector plasmid. The helper plasmid consists of all the necessary genes for M13 replication, transcription, and packaging, as shown in Fig. 9C.

A



B



C

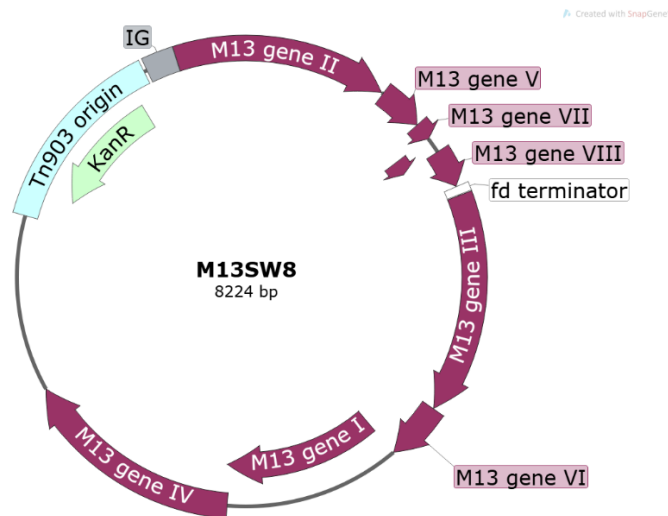


FIGURE 9. Plasmid maps. A) **pET-28_TTV19Cap Vector.** The pET-28_TTV19Cap vector consists of a T7 promoter, a lac operator, RBS, 6X His-Tag, TEV cleavage site, ORF1 of TTV19, and kanamycin resistance gene. B) **DNA vector plasmid pTTV19 IR_CMV_gfp_M13.** The DNA vector plasmid consists of the M13 start sequence, TTV19 intergenic region, CMV enhancer, CMV promoter, *gfp*, SV40 poly (A) signal, and M13 stop sequence. C) **Helper plasmid pM13SW8.** The helper plasmid pM13SW8 consists of all the necessary genes from M13KO7 for M13 replication and transcription, M13KO7 packaging signal is removed, and kanamycin resistance.

3.3. Transformation of *E. coli* cells with plasmids.

Calcium-competent *E. coli* cells were transformed by vectors and plated on LB plates + selective antibiotic overnight and plates were observed for growth the next day. The transformed colonies were maintained on LB plates + antibiotic. Subsequently, a plasmid extraction was carried out to confirm the transformation of the cells using the Monarch Plasmid Miniprep Kit (New England Biolabs, Whitby, CA) and digested with BamHI HF (New England Biolabs, Whitby, CA). Verification of successful transformation of the plasmids and correct plasmid band sizes after restriction digestion was confirmed by AGE analysis.

3.4. Protein production and extraction.

The transformed *E. coli* cells were grown in LB + Kn broth until the culture reached an optical density of $A_{600} = 0.6$. Once the culture reached an optical density of $A_{600} = 0.6$, the cells were induced with 1mM IPTG and were further grown for 4-5 h for protein production. The cells were pelleted by centrifuging at 10,000 x g for 10 min. The pellet was washed with cold PBS twice. The collected cell pellet was extracted using a cell extraction buffer from Invitrogen. 1 ml of extraction buffer was added per 10^8 cells. The cells were lysed in cell extraction buffer for 30 min on ice, with vortexing at 10 min intervals. The lysed cells were centrifuged at 10,000 x g for 10 min. at 4° C. The supernatant was collected for further analysis.

3.5. SDS-PAGE and Western Blot.

The protein detection component of this study was completed by conducting SDS-PAGE, which comprised running the protein samples through resolving and stacking gels. The resolving gel of 12% concentration was made using distilled H₂O, 1M Tris-HCl pH 8.8, 30% Acrylamide/Bis, 10% SDS, 10% APS, and TEMED. The stacking gel had a concentration of

4% and was made using distilled H₂O, 1M Tris-HCl pH 6.8, 30% Acrylamide/Bis, 10% SDS, 10% APS, and TEMED. The resolving and stacking gels were added to SDS gel plates, and 1mm gels were prepared. Approximately 10 ug of protein was mixed with 10 ul of 4X laemmli buffer. The sample was heated at 95° C for 5 min. using a dry blocker heater. 20 µl of samples and 3.5 µl of BLUEye prestained protein ladder (Frogga Bio Scientific Solution, Concord, CA) were loaded into the wells. The gels were run at 90 V for the first 15 min and then at 120 V for approximately 1 h in 1X running buffer.

Protein from the gel was transferred onto 0.45 µm nitrocellulose membrane (Bio-Rad, Mississauga, CA) using a Trans-Blot SD Semi-Dry Transfer cell (Bio-Rad, Mississauga, CA) at a voltage of 25 V for 30 min. To transfer the proteins from the gel to the membrane, a sandwich was prepared using extra thick blot paper (Bio-Rad, Mississauga, CA), BIO-RAD 0.45 µm nitrocellulose membrane, gel, and BIO-RAD extra thick blot paper. Before the sandwich was prepared, the extra thick blot paper and nitrocellulose membrane were placed in 1X transfer buffer. After the proteins were transferred onto the nitrocellulose membrane, the membrane was stained with ponceau stain for 5 min to verify the efficient run of samples through the gel and transfer onto the nitrocellulose membrane. The membrane was subsequently washed with MilliQ water to remove the ponceau stain. The membrane was then blocked using 5% skimmed milk in 1X Tris-HCl buffered saline/ Tween-20 (TBS-T) for a minimum of 1 h while shaking on ROCKER 35 (Labnet International Inc., New Jersey, USA). After blocking, the membrane was washed thrice with TBS-T for 10 min each while shaking. The membrane was then treated with primary antibody, 6X-His Tag monoclonal antibody (3D5) HRP (Catalog No. R931-25, Thermo Fisher Scientific, Mississauga, CA) at a concentration of 1 µg/ml diluted in blocking buffer (5% skimmed milk in TBS-T) overnight while shaking. The membrane was then rewashed 3X with TBS-T for 10 min each while shaking. After primary antibody treatment, the membrane was treated with secondary

antibody Goat anti-Mouse IgG (H+L) HRP from (Catalog No. 31444, Invitrogen, Burlington, CA) diluted with blocking buffer (5% skimmed milk in TBS-T) at a ratio of 1:10000 for 1 h while shaking. The membrane was then washed again three times with TBS-T for 10 min while shaking. To visualize the presence of protein, approximately 10 ml of PierceTM 1-StepTM Ultra TMB blotting solution (Thermo Fisher Scientific, Mississauga, CA) was added to the membrane. The membrane was then covered in foil and incubated for 30 min at RT while shaking. The expected full TTV19 capsid protein was expected to present as a band at ~75 kDa.

3.6. Protein Purification.

The over-expressed protein of interest was purified using His60 Ni superflow resin and gravity column kit (Clontech laboratories Inc., Mountain View, CA). The His60 Ni Gravity column (1 ml resin) and all the buffers were equilibrated to the working temperature of 4° C. The top and the bottom stopper of the column were removed, and the column was washed with 5-10 column volumes of His60 Ni equilibration buffer, after which the bottom stopper on the outlet column was replaced. The extracted protein sample was added to the column. The target protein was allowed to bind the resin by slowly inverting the column and incubating the column at 4° C for 1 h. After incubation, the column was installed in a vertical position, and the resin was allowed to settle. Once the resin was settled, 1 ml fractions were collected. The column was then washed with 10 column volumes of His60 Ni equilibration buffer followed by 10 column volumes of His60 Ni wash buffer. The targeted protein was then eluted with approximately 10 column volumes of elution buffer and, 1 ml fractions were collected. The eluted samples were subjected to SDS-PAGE and western blot analysis, to detect the presence of the protein of interest in the eluted samples.

3.7. Protein quantification.

The Pierce BCA Protein Assay Kit (Thermo Fischer Scientific, Waltham, USA) was used to determine the amount of protein present in the purified fractions. Following their protocol, the albumin standards were prepared at concentrations ranging from 0 to 2000 µg/ml. Next, 10 µL of the standards and samples were plated on a non-treated Nunc 96 flat-well plate (Thermo Fischer Scientific, Waltham, USA). Two hundred µL of working reagent (Reagent A and Reagent B at 50:1 ratio) was then added to each well, mixed with the samples, and incubated at 37° C for 30 min. The absorbances were subsequently read at 562 nm using the plate reader (Biotek, Winooski, USA). The protein concentration within the purified samples was calculated using the line of best-fit equation determined from the albumin standards.

3.8. Phagemid production.

A fresh colony of *E. coli* JM109 containing plasmids M13SW8 and TTV19 IR_CMV_gfp_M13 was grown in LB + Kn broth overnight. The next day, the culture was centrifuged (8,000 X g, 10 min) to separate the bacterial pellet from the supernatant containing the phage lysate.

3.9. Purification of phagemid.

The M13 phagemid lysate was then filtered through a 60 ml syringe using a 0.22 µm filter into a fresh 50 ml centrifuge tube. The lysate was then subjected to polyethylene glycol (PEG) precipitation. The filtered lysate was concentrated by adding 1/5th the volume in PEG/NaCl and mixed by inversion. The mixture was incubated at 4° C for 2 h and centrifuged at 12,000 X g for 15 min at 4° C. The white pellet was re-suspended in cold TN buffer and a second PEG precipitation was carried out to further concentrate the phagemid. The PEG-precipitated phagemid was then treated with DNaseI (Promega, Madison, USA) for

1h at 37° C to remove any contaminating bacterial DNA. The concentrated phagemid was stored at 4° C.

3.10. Quantification of phagemid through plaque assay.

A plaque assay was carried out to quantify viable M13 phagemid. *E. coli* strain JM109 was used for carrying out plaque assay as it is an F⁺ strain suitable to produce filamentous phages. A series of M13 phagemid dilutions were prepared in TN buffer to obtain countable (30-300) plaques on a spot plate. 200 µL of early log-phase of *E. coli* strain JM109 was mixed with 3 ml of 0.8% top agar supplemented with 5 mM MgSO₄, which was then poured on a pre-warmed LB agar plate. A 5 µL sample of each serially diluted M13 phagemid was then spotted on the top agar. The plates were then incubated at 37° C for 4-8 h. The lowest dilution that generated distinct plaques on the bacterial lawn was selected to generate full plates for plaque assay. The full plate plaque assay was carried out using the same procedure, instead of spotting the phagemids on the top agar, the phagemids were mixed directly into the bacterial culture before adding to the top agar and pouring it on the LB plates. The titre was determined from the mean of three counts and expressed as plaque-forming units (PFU) per millilitre (ml).

3.11. Phenol-chloroform extraction.

The PEG-concentrated phagemid lysate was mixed with phenol (1:1, v/v) and subsequently vortexed. The mixture was centrifuged at 4° C for 5 min, and the top aqueous layer was extracted. This layer was extracted again with an equivalent volume of phenol: chloroform twice and once with chloroform. The phagemid ssDNA was precipitated overnight in 95% ethanol at -80° C. The precipitated ssDNA was washed twice with 70% ethanol, dried, and re-suspended in DNase and RNase- free water. The ssDNA was quantified using Nanodrop

analysis. The ssDNA was then subjected to gel electrophoresis to confirm the size of the ssDNA.

3.12. *In vitro* assembly of viruses.

To a final concentration of 7.5 mg/ml, the capsid proteins were suspended in 5X assembly buffer (pH 7.2, 250 mM Tris-HCl buffer containing 250 mM NaCl, 50 mM KCl, and 25 mM MgCl₂). A 1X assembly buffer solution containing 50 mM Tris-HCl, 50 mM NaCl, 10 mM KCl, and 5 mM MgCl₂ with a pH of 7.2 was created by combining capsid protein and ssDNA in a 6:1 mass ratio (w/w) of capsid protein to ssDNA and a 4:1 (v/v) ratio of Milli-Q water to buffer. This was then incubated for ssDNA at 4° C for at least overnight before analysis and purification. The capsid proteins were suspended in an assembly buffer without ssDNA that was used as the negative control for qPCR and functional assembly. Similarly, ssDNA in the absence of protein served as an unencapsidated control for DNase digestion assays.

3.13. DNase enzymatic digestion assay.

Anellovirus capsid-assembled ssDNA and protein and/or ssDNA only control was subjected to DNase treatment. 1 µL of DNase enzyme was added to the samples. The samples were then incubated at 37° C for 1 h. After the treatment of samples with DNase, the samples were treated with 1µL of DNase stop solution and incubated at 65° C for 15 min. To visualize ssDNA digestion AGE was performed. The samples were stored at 4° C for further analysis.

3.14. Iodixanol ultracentrifugation.

The crude protein extract and assembled ssDNA with the crude capsid protein were subjected to iodixanol ultracentrifugation. The iodixanol gradient (10-60%) was prepared using OptiPrepTM and 1X NTC solution. To prepare a step gradient, a 1.5 ml layer of each concentration starting from the highest to the lowest (60% → 10%) was added to thin wall

tubes for the SW41 rotor (polyallomer centrifuge tubes, 14 X 89 mm; Beckman Coulter, 331372). The gradient tube was then stored at 4° C. Alternately, the samples were prepared by pelleting through ultracentrifugation in an SW41 T1 rotor at 4° C for 90 min at 26,000 rpm. The pellet was re-suspended in chilled 1X NTC solution. The samples were then layered onto the iodixanol density gradient, followed by ultracentrifugation in an SW41 T1 rotor at 4° C for 150 min at 35,000 rpm. After ultracentrifugation, 12 fractions (with a volume of ~ 1 mL each) were collected for each sample by puncturing a hole within the centrifuge tubes from the top to the bottom of the tube.

3.15. Transmission Electron Microscopy (TEM).

The assembled capsid protein and ssDNA and capsid proteins were visualized using TEM. The samples were poured in a volume of 20 µL onto copper grids covered with carbon made by Electron Microscopy Sciences (ThermoFisher Scientific, Mississauga, CA) and incubated for 1-2 min. The sample solutions were removed from the grid using Whatman Cellulose Filter Paper. Each grid was then washed by adding 40 µL of MilliQ water and incubating the grids for 1 min. The MilliQ water was then removed from the grid using filter paper. Each grid was stained with 20 µL of 2% phosphotungstic acid for 1 min. The stain was removed using filter paper, and the grids were washed 5X with MilliQ water. The grids were then air-dried overnight and subsequently imaged using the Phillips CM10 transmission electron microscope at a magnification of 64 000X.

3.16. Polymerase chain reaction (PCR).

PCR was carried out for extracted ssDNA from capsid proteins using primer sequences indicated in Table 9. The C1000 Thermal Cycler was used for PCR amplification (Bio-Rad, Hercules, USA). The PCR reaction contained 10 µL of Phusion Flash High Fidelity PCR Master Mix (Thermo Fisher Scientific, Waltham, USA), 5 µL of sample DNA to be

amplified, 0.5 μ M of forward and reverse primers, and sterile water within a final reaction volume of 20 μ L. Table 10 outlines the settings applied for the PCR reaction. The PCR amplified products were then run on a 2% gel for approximately 1.5 h at 88 V to verify for amplification of the $\sim$ 100 bp *gfp* gene region.

Table 10: PCR amplification protocol.

Cycle Step	Temperature	Time	Number of Cycles
Initial Denaturation	98° C	10 sec	1 cycle
Denaturation	98° C	5 sec	30 cycles
Annealing	62° C	5 sec	
Extension	72° C	2 sec	
Final Extension	72° C	1 min	1 cycle
Incubation	4° C	Incubate	N/A

3.17. Quantification of virion particles.

Virus particles were quantified using quantitative PCR (qPCR). A calibration curve for quantifying virion particles was constructed using the pTTV19 IR_CMV_gfp_M13 plasmid that encodes the *gfp* gene. The *gfp* gene region was used for the amplification of the DNA samples, where 100-fold serial dilutions (gc/ μ l) of the plasmid were made (10^0 - 10^8) to generate the calibration curve.

For each qPCR amplification, the PCR reaction was prepared using 5 μ L of PowerUp SYBR Green Mix (Thermo Fisher Scientific, Waltham, USA), 1 μ L each of 500 nM primer (forward and reverse), 2 μ L of the DNA sample, and 1 μ L of dH₂O for a final reaction volume of 10 μ L. PCR cycling conditions were as follows: 50° C for 2 min, 95° C for 2 min, followed by 40 cycles at 95° C for 15 s and 60° C for 1 min. Next, the melt curve was set for 1 cycle at

95° C for 15 s, 60° C for 1 min, and 95° C for 15 s. PCR reactions were run in triplicate on the StepOne Plus Real-Time PCR system (Applied Biosystems, Waltham, USA). The quantification cycle or threshold cycle number (C_t) for each reaction was recorded and used in subsequent analysis.

The following equation converts the mass of dsDNA to the quantity of genome copies:

$$gc = \frac{\text{Concentration of template DNA} \times 6.022 \times 10^{23}}{\text{Length of the DNA (bp)} \times 1 \times 10^9 \times 660}$$

where gc is the concentration of virion genome copies (genome copies/ μ L).

Virus concentrations were estimated from their respective calibration curve. First, the C_t values from the control were plotted against the log of the known concentrations for each concentration in the dilution series.

Linear regression produces a familiar equation:

$$C_t = mx + b$$

where C_t is the measured threshold cycle number, m is the slope, x is the base-10 log of the concentration in gc/ μ L and b is the x-intercept.

PCR amplification efficiency is given by $E = 10^{-1/m} - 1$.

The virion concentration, V can be estimated simply by the equation:

$$V = 10^{(C_t - b)/m} \times 2$$

where multiplication by 2 adjusts for the estimation of single-stranded products (gc/ μ L) from dsDNA standards. Only C_t measurements within the bounds of the calibration curve were used to calculate virion concentration.

4. RESULTS

4.1. Selection of a suitable anellovirus candidate for capsid protein production and *in vitro* assembly of ssDNA.

TTV19 was used as the reference viral species to identify other human anellovirus strains that could prove to be promising candidates for this study. Fourteen other strains similar to TTV19 were found with maximum percent identity (Table 11). The ORF1 of anellovirus is involved in capsid protein production and the intergenic region contains the DNA binding signal. Based on virus distribution, stability, and DNA packaging efficiency, TTV19 was selected in the end as the most promising viral candidate to be used. While other species demonstrated high virus distribution among different cells and comparable stability, they also offered lower DNA packaging efficiency compared to TTV19, which was considered a criterion of imperative importance.

TABLE 11: Genomic characteristics and locations of transcription elements of the Anellovirus species.

Anellovirus Species	LOCUS	Genomic length (nt)	ORF-1 (Capsid Protein)	Intergenic Region
TTV19	AB025946	3808	595-2832	3697-3808
TTV san IR-1031	AB038619	3795	595-2835	3685-3795
TTV san-S039	AB038620	3792	595-2832	3681-3792
TTV-TJN02	AB028669	3794	599-2839	3677-3781
TTVMO2	MN116509	3265	511-2673	3191-3265
Anelloviridae sp. isolate ctjd005	MH649042	3718	483-2762	3577-3718
TTV-Hebei-1	KJ082064	3649	620-2824	3508-3649
Anelloviridae sp. isolate P12-c1	MW455448	3741	594-2813	3377-3741
Anelloviridae sp. isolate P3-c6	MW455379	3309	649-2889	3131-3309
Anellovirus cyc/2019	MT501653	3651	620-2824	3492-3651
Anellovirus cyb/2019	MT501649	3650	620-2824	3461-3601
TTV10 isolate BIS8-17	GU797360	3767	612-2846	3634-3767
TTV isolate TCHN-G2	AF345529	3214	495-2732	3135-3214
TTV isolate TCHN-D2	AF345525	3196	479-1759	3120-3196
TTV10 isolate JT34F	AB064607	3770	612-2849	3447-3770

As the selected candidate, The ORF1 gene of TTV19 was codon optimized and inserted into the pET vector containing a 6X His-Tag and TEV cleavage site.

4.2. Production of TTV19 capsid protein for *in vitro* virus assembly.

4.2.1. Production of TTV19 capsid protein.

The possibility of TTV19 capsid protein production in a bacterial system was investigated by first transforming *E. coli* BL21 cells with the constructed pET-28_TTV19Cap vector that expresses TTV19ORF1 in combination with an N-terminal 6X His tag. To visualize successful transformants, the extracted pET-28_TTV19Cap vector plasmid was linearized with Bam HI and visualized via AGE (Fig. 10). The extracted pET-28_TTV19Cap vector plasmid resolved to its respective genome size i.e., ~7.5kb.

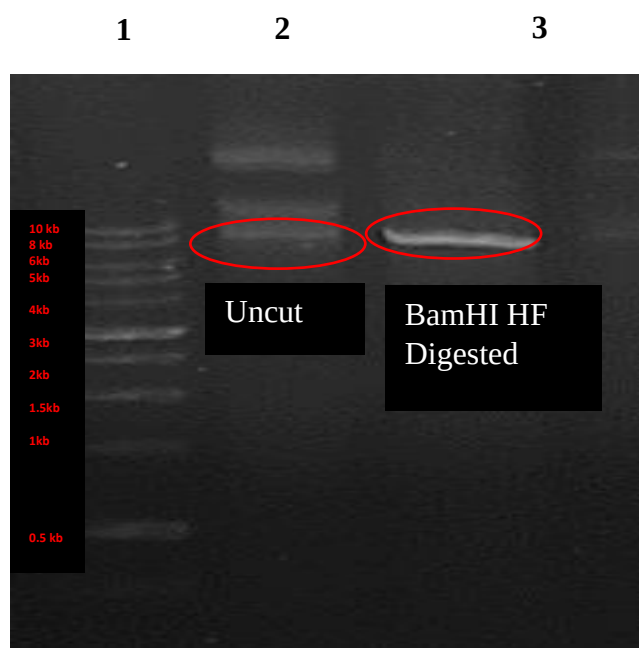


FIGURE 10. Agarose gel of pET-28_TTV19Cap and its restriction digestion. ORF1 of TTV19 was inserted into the pET-28 vector for production of TTV19 capsid protein. A) Lane 1 shows 1kb DNA ladder. B) pET-28_TTV19Cap plasmid, which is 7.5 kbp extracted from transformed *E.coli* strain BL21. C) Lane 2 shows extracted pET-28_TTV19Cap plasmid restriction digested with BamHI HF.

Next, the production of TTV19 capsid protein in *E. coli* BL21 was evaluated using SDS-PAGE and western blot analysis (Fig. 11). TTV19 capsid protein expression was confirmed by the presence of a band at ~75 kDa which was the expected size. However, along with the

expected capsid protein size, a band of lower molecular weight (~50 kDa) was observed. While the reason for this band is unconfirmed, it is likely due to proteolytic cleavage and has been observed by other investigators (See Discussion).

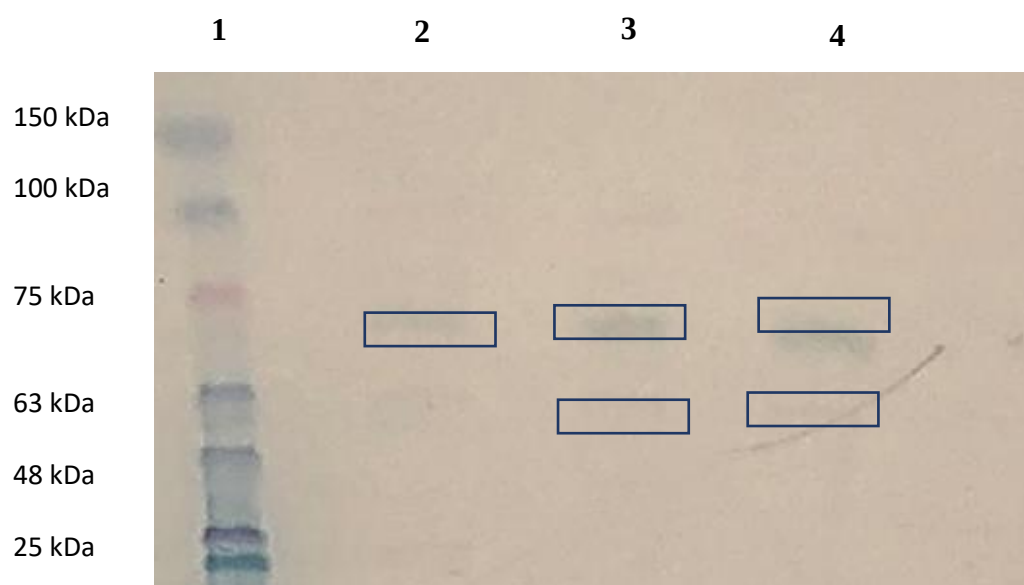


FIGURE 11. Production of TTV19 capsid protein. TTV19 capsid protein was produced via *E. coli* BL21 cells. The expressed proteins in the cell lysate were detected using western blot. A) Lane 1 is protein molecular weight ladder in kDa. B) Lanes 2, 3, and 4 shows protein bands at ~75kDa and ~50kDa.

4.2.2. Optimization of TTV19 Capsid protein production.

Once confirming that *E. coli* BL21 cells were able to produce the TTV19 capsid protein, experiments were conducted to obtain maximal capsid production within this system. The TTV19 capsid protein sequence is under the control of the *lac* operator, regulating the T7 promoter, and thus requires IPTG induction to result in high-level expression. Therefore, TTV19 capsid protein expression was evaluated following IPTG induction at various optical densities. The expression of the protein was confirmed using SDS-PAGE and western blot analysis (Fig. 12). From the analysis it was observed that capsid proteins were expressed at all the tested optical densities by the presence of a band at ~75 kDa. However, the maximal

expression was observed at optical density $A_{600} = 0.6$. As such, for this study, TTV19 capsid proteins were expressed after inducing the cells with IPTG at $A_{600} = 0.6$. Along with proteins expressed at ~75 kDa, protein bands of lower molecular weight were observed that were ~25 kDa. These bands were smaller than those observed in Fig 11. From this, it can be inferred that the proteins potentially get degraded in a sequential manner to produce a mature virus capsid head, or there is a proteolytic cleavage occurring in the bacterial system.

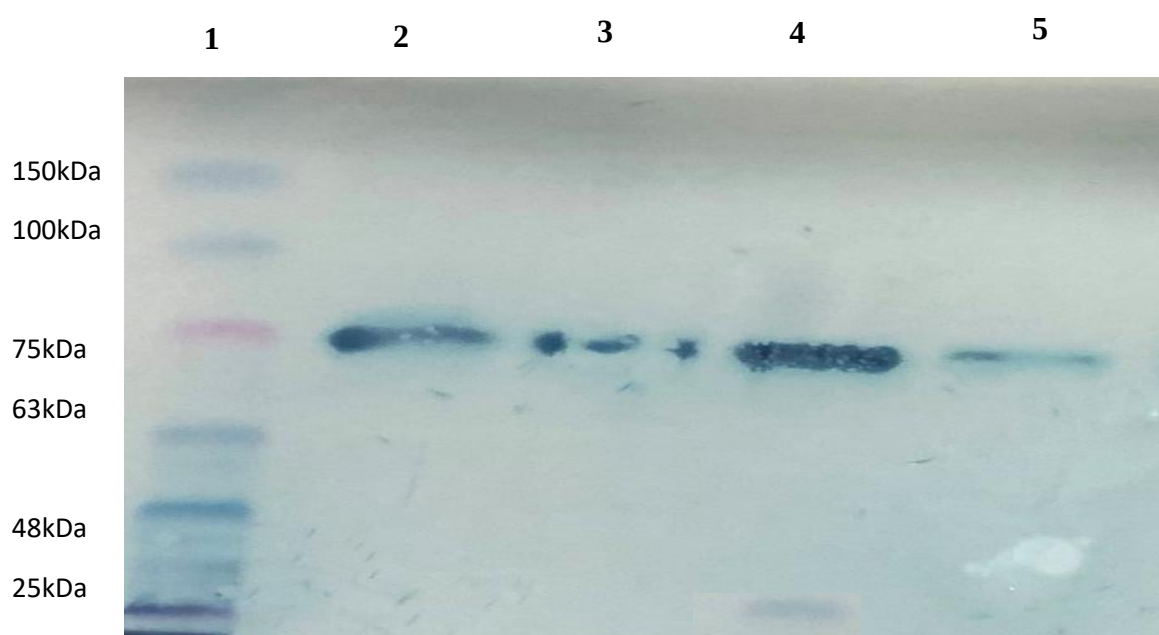


FIGURE 12: Production of TTV19 Capsid protein at different O.D. The expressed proteins in the cell lysate were detected using western blot. A) Lane 1 is protein molecular weight ladder in kDa. B) Lane 2 is IPTG induction at $A_{600} = 0$ which showed the presence of protein band at 75kDa. C) Lane 3 is IPTG induction at $A_{600} = 0.4$ which showed the presence of protein band at 75kDa. D) Lane 4 is IPTG induction at $A_{600} = 0.6$ which showed the presence of two protein bands at 75kDa and 25kDa. E) Lane 5 is IPTG induction at $A_{600} = 0.8$ which showed the presence of two protein bands at 75kDa and 25kDa.

4.2.3. Purification of expressed TTV19 capsid protein.

Once, the TTV19 capsid proteins were optimally expressed in *E. coli* BL21, the expressed proteins were purified for *in vitro* assembly of virus particles using the Ni-NTA purification column, as the capsid protein sequence has a 6X His-Tag at the N-terminal end of the capsid

protein. After purification, the eluted protein samples were subjected to SDS-PAGE and western blot analysis (anti-His Ab) to detect the presence of purified capsid proteins, as shown in Fig 13. From Fig.13, it was observed that the proteins were successfully purified, although, somewhat unexpectedly, the molecular weight of the purified protein was ~25 kDa, which is smaller compared to the expected protein molecular weight (~75 kDa). From Fig. 13A, it was observed that the expressed proteins in crude extracts produced lighter protein bands at ~75 kDa, along with another protein band at ~50 kDa. However, once the proteins were purified, proteins that were ~25 kDa in size were observed in elution fractions 1, 2, and 3 following purification. While the origin of the smaller His⁺ protein is unknown, it is quite likely that the mature 75 kDa ORF1 protein has been further processed to generate a 25 kDa product. Whether this processing is specific to capsid maturation and resulted in a mature virus capsid head, or proteolytic cleavage occurred in the bacterial system was as yet unknown. Therefore, the purified protein and the crude extract containing the capsid proteins were both used for further *in vitro* assembly of virus particles.

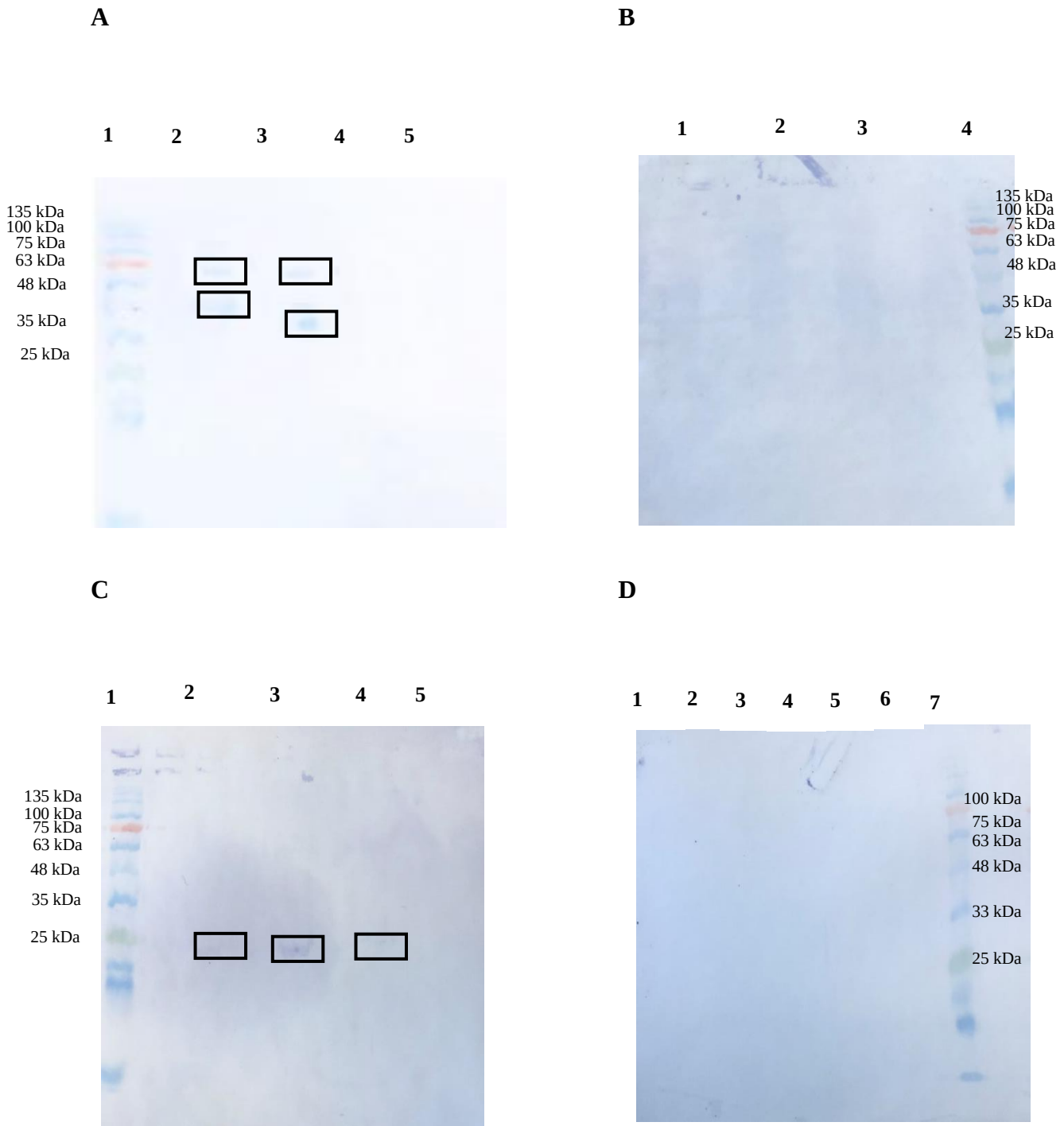


FIGURE 13: Purification of His-Tagged TTV19 capsid protein. A) Lane 1 is protein molecular weight ladder, Lane 2& 3 is crude extract containing TTV19 capsid protein, Lane 4&5 is BW1 and BW2. B) Lane 1,2, & 3 are BW5, BW4, and BW3 respectively, Lane 4 is protein molecular weight ladder. C) Lane 1 is protein molecular weight ladder, Lane 2,3,4,&5 is protein samples after elution, E1, E2, E3, and E4 respectively, D) Lane 1,2,3,4,5,&6 are eluted proteins samples E10, E9, E8, E7, E6, and E5 respectively, and Lane 7 is protein molecular weight ladder. Key: CE = Crude protein extract containing TTV19 capsid protein, BW = Samples collected by washing the columns before eluting the sample, and E = Sample elution from the column.

Next, the concentration of purified protein samples was determined for *in vitro* virus assembly. A standard curve as shown in Fig. 14 was obtained for albumin concentrations. From the standard graph, a linear equation was obtained which is $y = 0.0006x + 0.0211$ and R^2 value of 0.9986. Using this formula, the unknown concentration of purified TTV19 capsid proteins was determined. From all the purification steps carried out during the study, the average concentration of the protein sample was found to be $39.4985 \pm 9.148 \mu\text{g/ml}$. The purified TTV19 capsid protein samples were used for *in vitro* assembly of ssDNA.

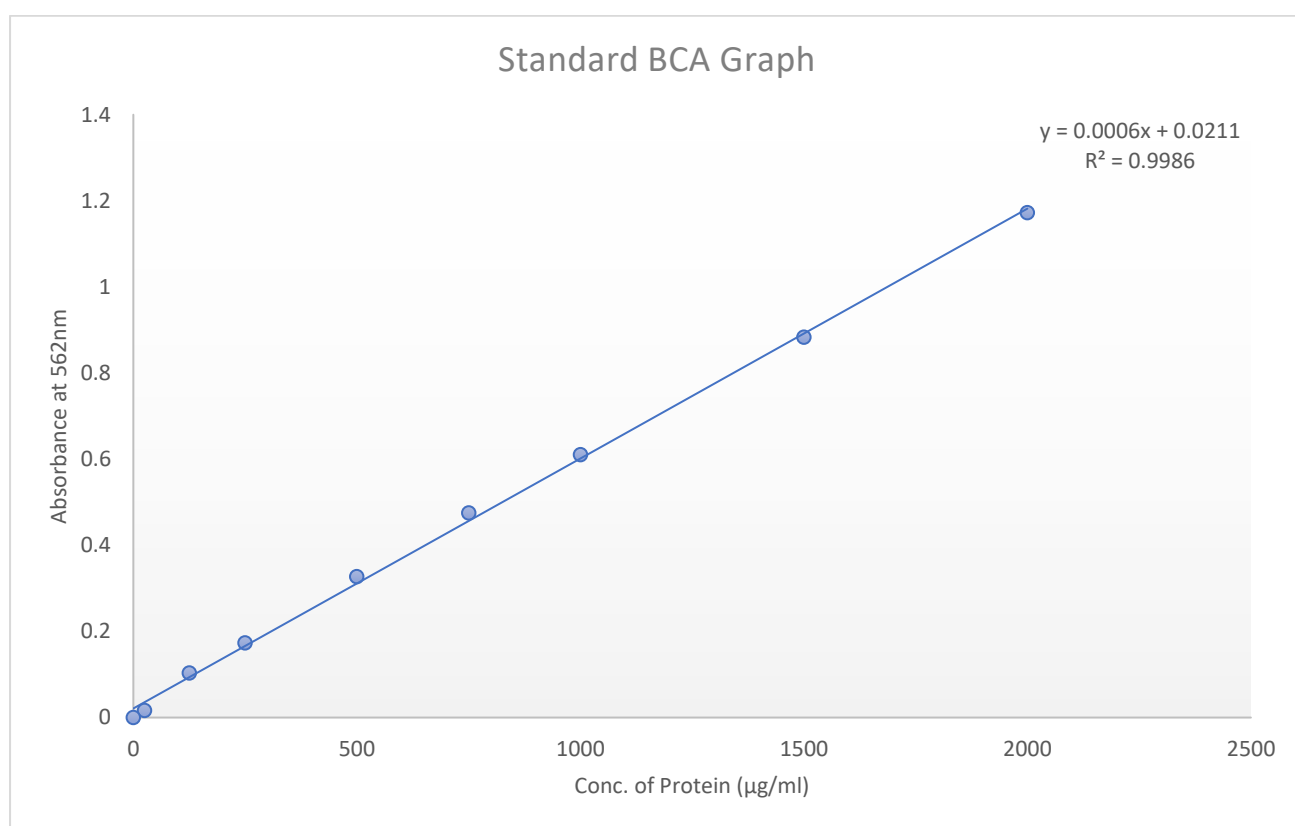


FIGURE 14: Standard BCA graph for protein concentration determination. The standard BCA graph was obtained using albumin using concentrations ranging from 0 µg/ml to 2000 µg/ml.

4.3. Production of ssDNA containing the TTV19 DNA binding signal using M13 replication system.

The second component of the viral construct focused on the production of ssDNA for *in vitro* assembly of virus particles. This was achieved by transforming *E. coli* JM109 cells by pTTV19 IR_CMV_gfp_M13 (for packaging of ssDNA) and the helper plasmid (pM13SW8) for M13 phagemid production.

4.3.1. Production, purification, and quantification of M13 phagemid containing the TTV19 DNA binding signal.

A plaque assay was conducted to determine phage concentration, but before performing a full plate plaque assay, spot plating was performed to determine the lowest dilution which generates distinct plaques on bacterial lawns. However, no plaques were obtained on the plates, as shown in Fig. 15. This was observed since M13SW8 lacks a packaging signal and cannot extrude from the cell and conversely, extruding phagemids cannot propagate upon reinfection of a target cell as they lack complementary genes.

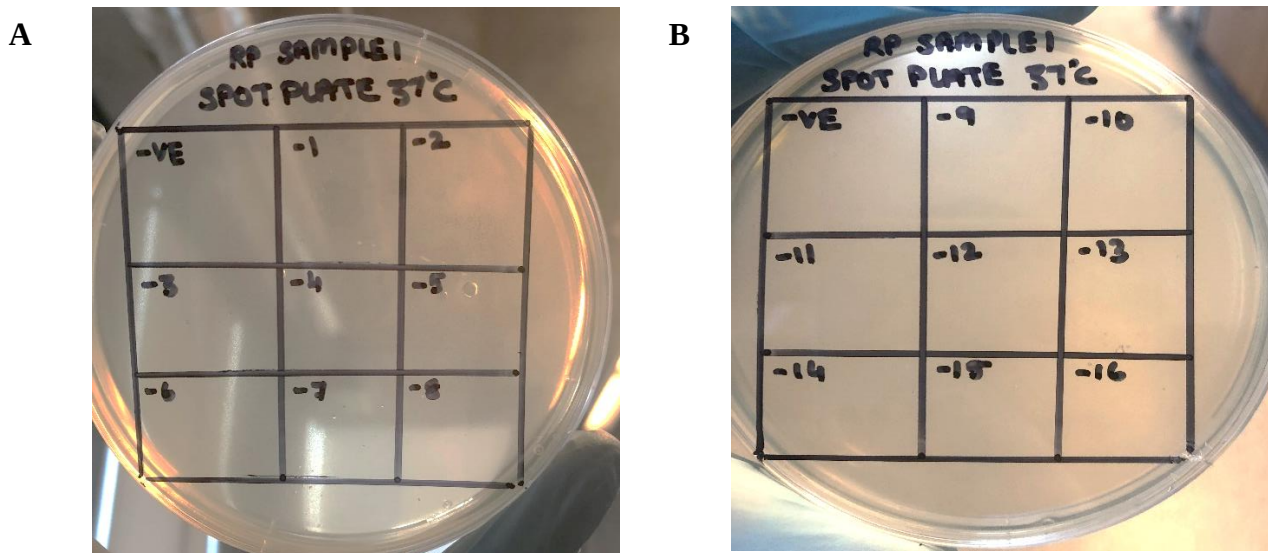


FIGURE 15: Phagemid Plaque assay using spot plate method. A) Spot plaque assay for dilutions 10^{-1} to 10^{-8} . B) Spot plaque assay for dilutions 10^{-9} to 10^{-16} .

Hence, to obtain the phagemid titre, qPCR was performed where a segment of the *gfp* gene region was amplified. The average phage concentration was found to be $1.53 \times 10^9 \pm 1.66 \times 10^8$ phage/ml.

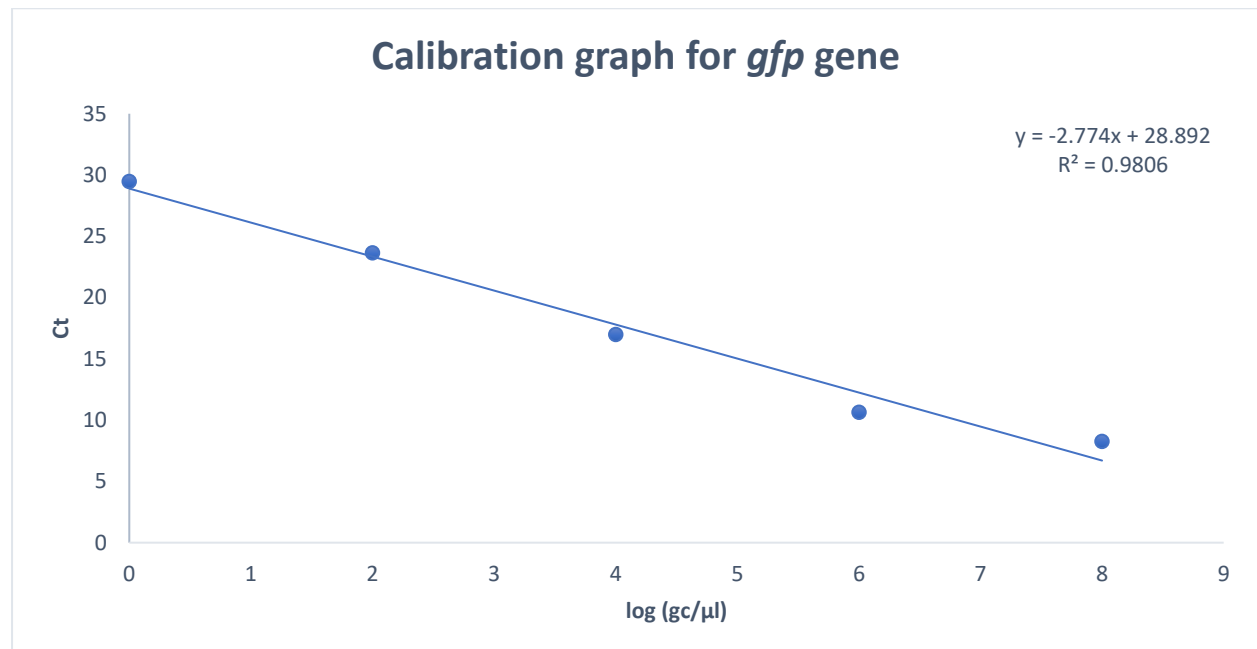


FIGURE 16: Calibration graph to estimate phage concentration. Representative calibration curve for *gfp* gene region measured from qPCR of pTTV19 IR_CMV_gfp_M13 plasmid.

4.3.2. Extraction of ssDNA from M13 phagemid for *in vitro* assembly.

After M13 phagemids were produced and purified, ssDNA was extracted from M13 phagemids. The extracted ssDNA was visualized using AGE and appeared on the gel at a size of ~ 0.8 kb (Fig. 17). The ssDNA is ~ 1.6 kb in size but is expected to appear on the gel as half the size ~ 0.8 kb since the DNA sample is single-stranded and will run faster.

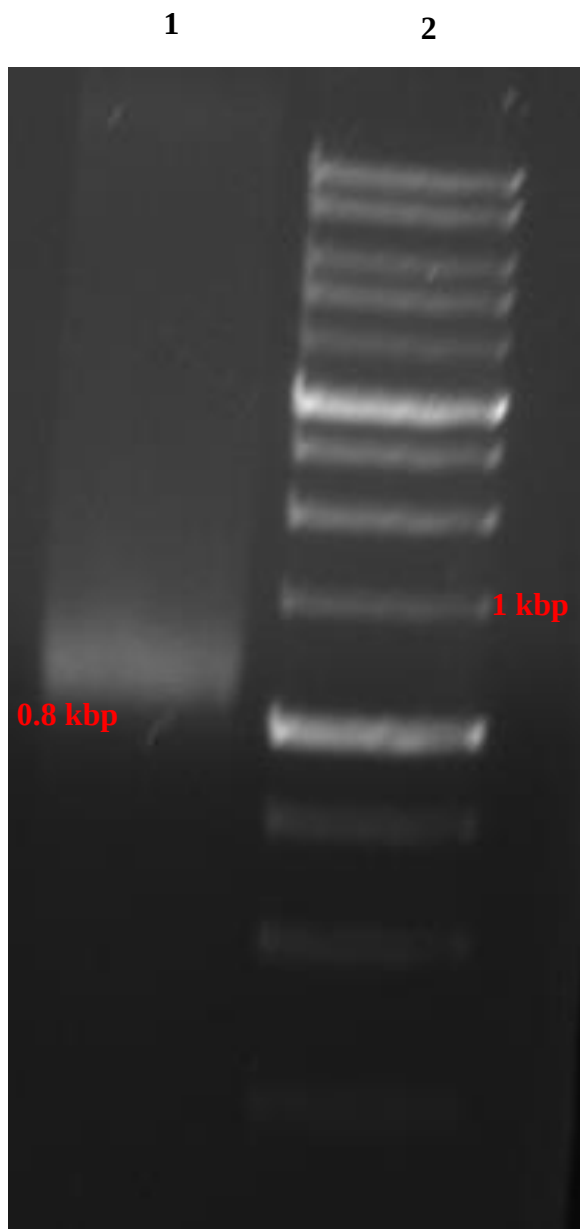


FIGURE 17: ssDNA obtained from M13 phagemid carrying the gene cassette. A) Lane 1 shows ssDNA obtained from M13 phagemid which is ~1.6kb. Since the DNA is ss it will run faster, and the DNA should be ~0.8kb as observed above. B) Lane 2 is DNA ladder.

4.4. *In vitro* assembly of ssDNA containing the gene of interest and TTV19 intergenic region, and TTV19 capsid protein.

The next step of the study involved the *in vitro* assembly of capsid proteins and the extracted ssDNA. This was achieved by assembling His-Tag purified TTV19 capsid protein and ssDNA. Along with His-Tag purified capsid proteins, crude extract containing TTV19 capsid protein and ssDNA were also assembled *in vitro*. The assembly of virus particles and packaging of ssDNA was confirmed using the experiments described below.

4.4.1. Iodixanol ultracentrifugation of crude extract containing TTV capsid protein and assembled crude extract containing TTV19 capsid protein and ssDNA.

Once the crude extract containing TTV19 capsid proteins assembled ssDNA, the assembled products were subjected to iodixanol ultracentrifugation for purification of the assembled virus particles. Along with the assembled capsid protein sample, crude extract containing only TTV19 capsid was also subjected to iodixanol ultracentrifugation. The purified samples were then subjected to SDS-PAGE and western blot to detect the presence of purified capsid protein in the fraction. The purified proteins for both samples were obtained in fraction 10 at a band size of ~ 25 kDa, as shown in Fig. 18 and Fig. 19. Hence for further study, fraction 10 of the iodixanol centrifuged samples of crude extract containing TTV19 capsid protein, and fraction 10 of the iodixanol centrifuged samples of assembled ssDNA and crude extraction containing TTV19 capsid protein were used.

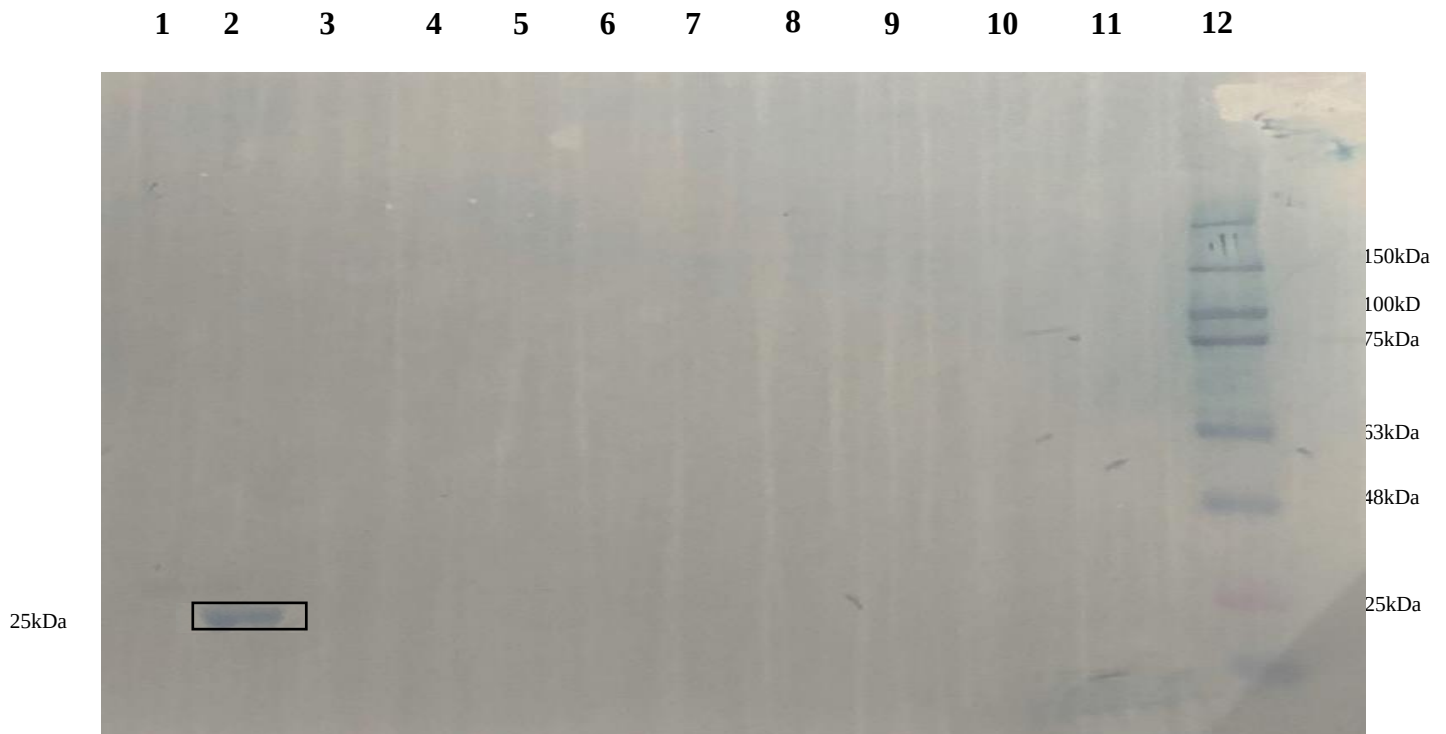


FIGURE 18: Iodixanol ultracentrifugation of crude extract containing TTV19 capsid protein. Lane 1 – 10 is F11 to F1 respectively, and Lane 12 is protein molecular weight ladder Key: F = Fraction.

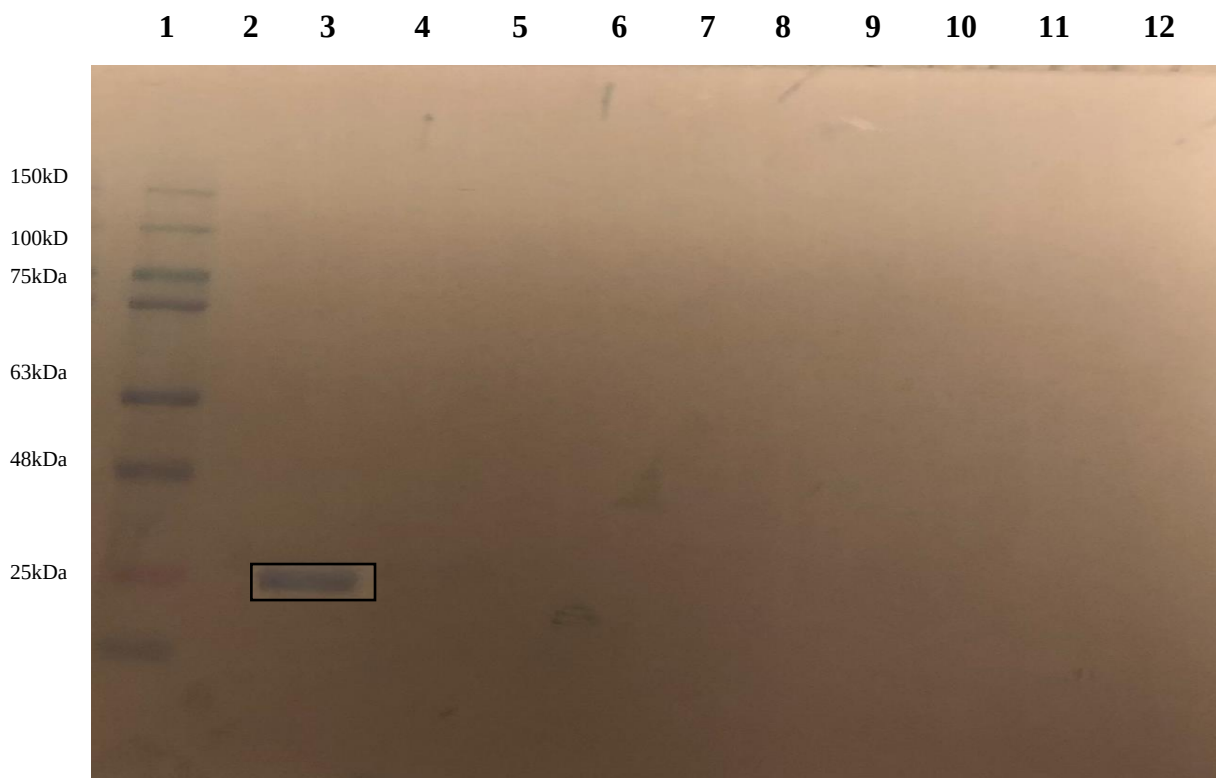


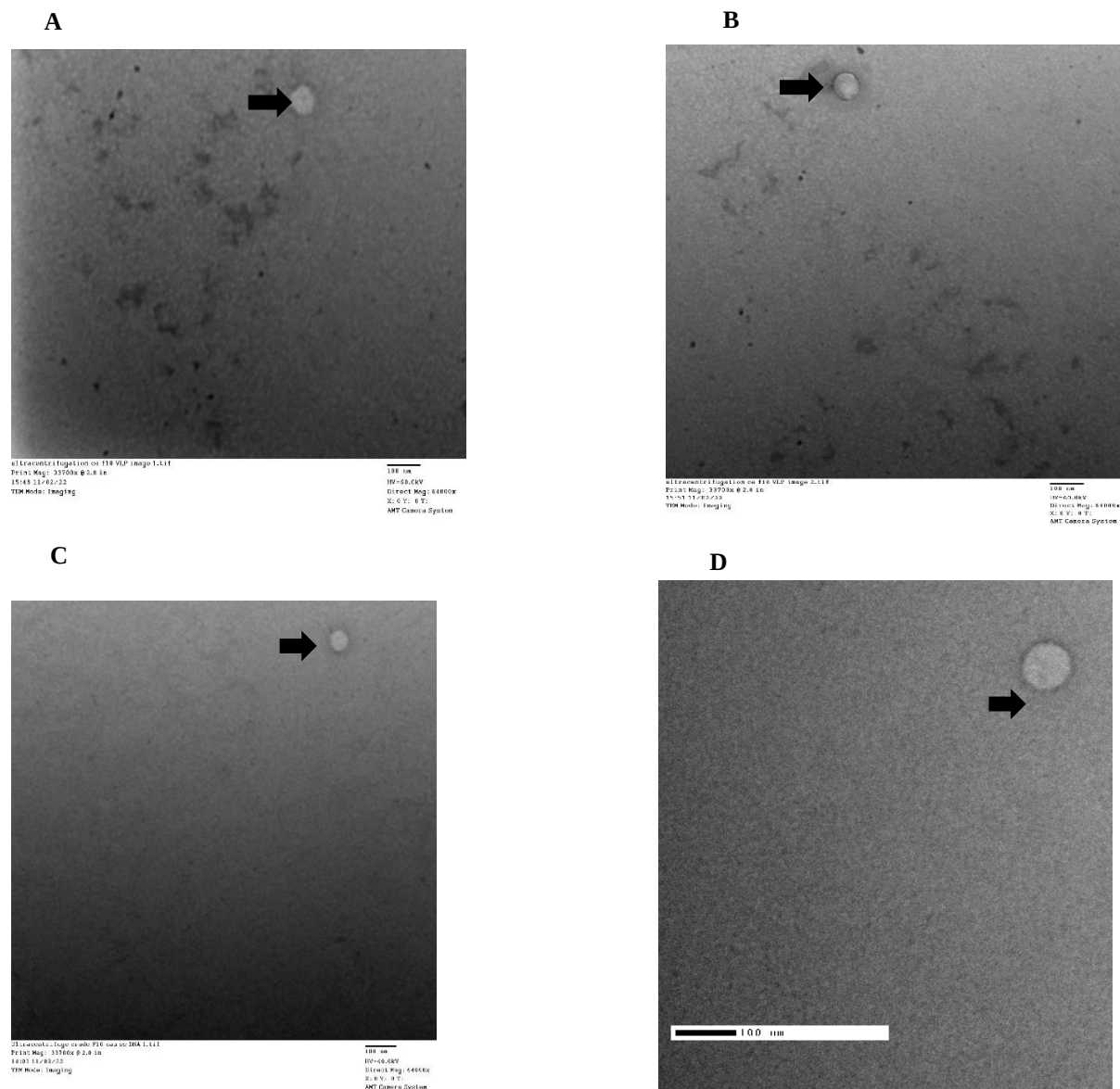
FIGURE 19: Iodixanol ultracentrifugation of *in vitro* assembled ssDNA and crude extract containing TTV19 capsid protein. Lane 1 is protein molecular weight ladder, and Lanes 2 – 12 is F11 to F1 respectively. Key: F = Fraction.

4.4.2. Transmission electron microscopy (TEM) for *in vitro* assembled TTV19 capsid protein and ssDNA containing the gene of interest.

After assembling capsid proteins and ssDNA *in vitro*, the samples were examined by TEM to try to identify any present assembled virus particles. TTV19 virus is ~50nm and assembles as an icosahedral head. TEM images for crude extract containing TTV19 capsid protein showed the presence of what seemed to be icosahedral heads approximately ~50 nm in size (Fig. 20 A & B). While not particularly conclusive this finding supports the hypothesis that the capsid proteins are assembling into virus-like particles (VLP) in the bacterial system.

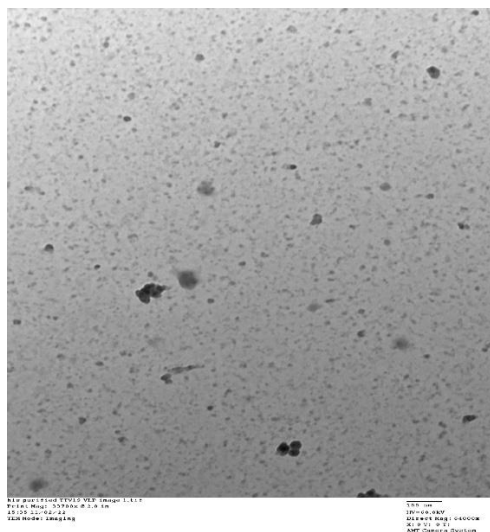
However, while both assembled crude extract containing TTV19 capsid protein and ssDNA, and assembled capsid protein and ssDNA treated with DNase indicated the presence of icosahedral heads of the correct size (Fig. 20 C, D, E, & F) they were very few in number.

The negative controls Proteinase K treated assembled crude extract containing TTV19 capsid protein and ssDNA treated with DNase enzyme did show the presence of an icosahedral heads which demonstrates that the virus particles are assembling (Fig. 20 G & F). To confirm the assembly of virus particles further testing is required using more conclusive methods.

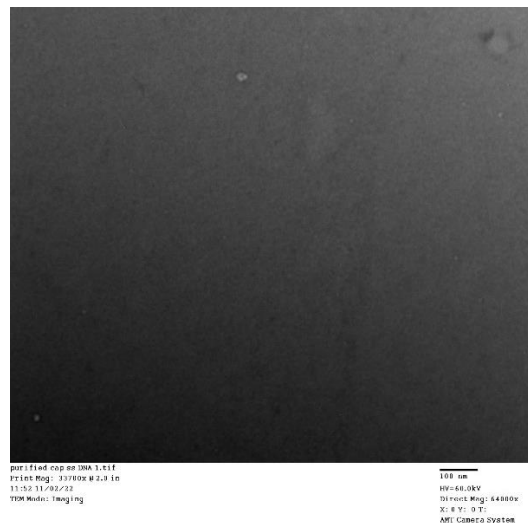


don't assemble on their own. Also, this could be observed due to low concentrations of viral capsid proteins. Next, we subjected assembled His-Tag purified capsid protein and ssDNA (Fig. 21B). It was found that very few assembled into a viral particle. The assembled His-Tag purified TTV19 capsid protein treated with DNase was also subjected to TEM analysis, which showed the presence of icosahedral heads of proper size nm (Fig. 21C & D). The absence of an icosahedral head for proteinase K-treated samples and ssDNA digested with DNase enzyme (Fig. 21 E & F) validates that the viral particles might be assembling to form icosahedral particles. However, the results show that the viruses might be assembling into icosahedral heads, but whether the TTV19 capsid proteins are packaging ssDNA could not be evaluated by this method alone, and as such, further testing was required to confirm DNA encapsidation in TTV19 capsids.

A



B



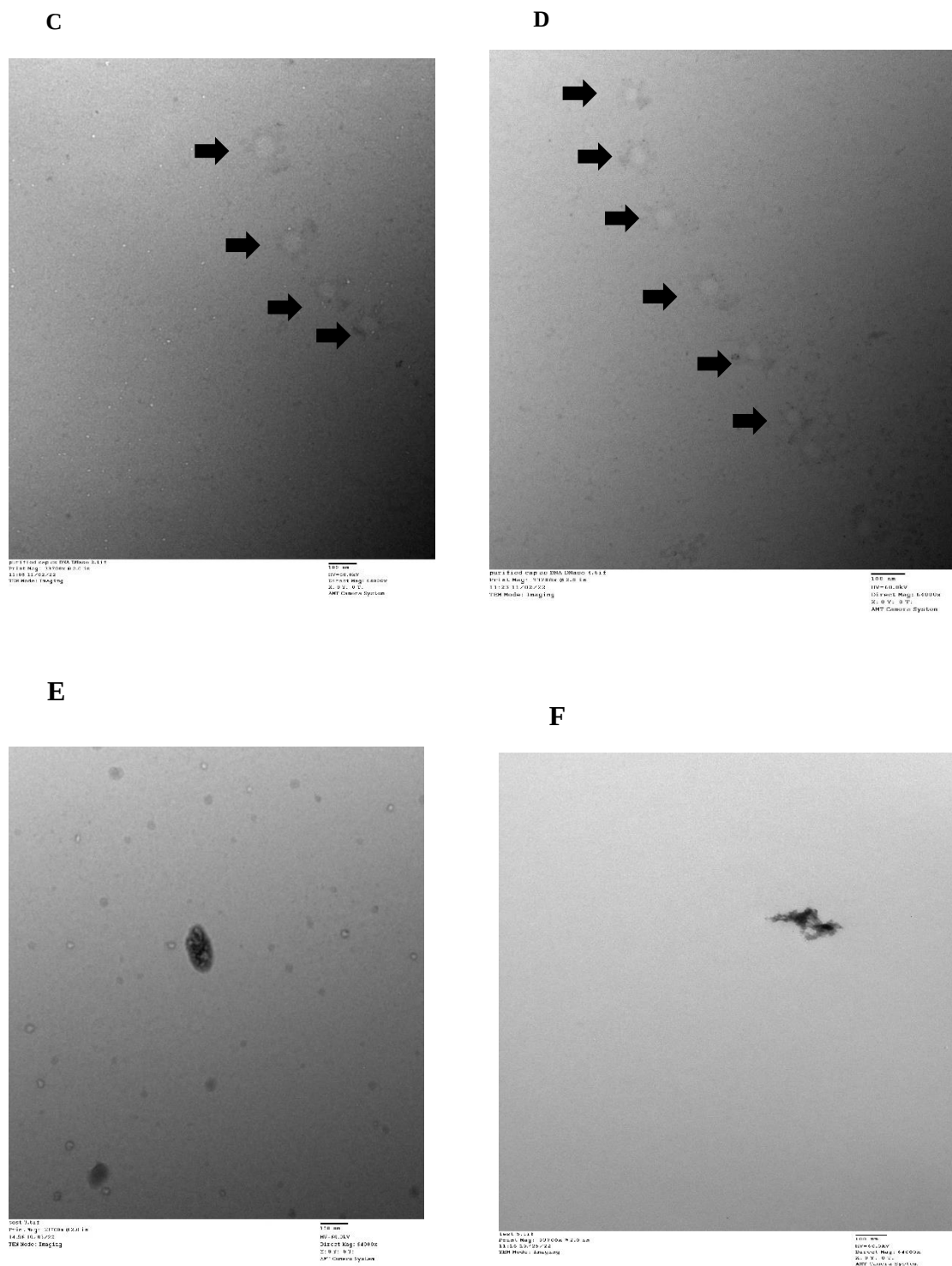


FIGURE 21: TEM images His-Tag purified TTV19 extract samples. A: His-Tag purified TTV19 capsid proteins; B: assembled His-Tag Purified TTV19 capsid protein and ssDNA; C&D: DNase treated assembled His-Tag purified TTV19 capsid protein and ssDNA. E: Proteinase K treated assembled His-Tag Purified TTV19 capsid protein and ssDNA; and F: ssDNA treated with DNase.

4.4.3. Detection of packaging of ssDNA by TTV19 capsid protein *in vitro* by PCR.

Next, to confirm that the capsid proteins are assembling ssDNA *in vitro*, PCR was performed to amplify a segment *gfp* gene region of ~100 bp carried by the DNA vector. Following PCR amplification, the amplified DNA fragment (~100 bp) was visualized by AGE (Fig. 22). The ssDNA extracted from an assembled crude extract containing TTV19 capsid protein and ssDNA exhibited *gfp* gene amplification (~100 bp). This indicates that the capsid proteins present in crude extract in assembling ssDNA *in vitro* (Fig. 22A). On the other hand, ssDNA extracted from assembled His-Tag purified capsid protein and ssDNA did not show any amplification of the *gfp* gene (Fig. 22B) which might indicate that the capsid proteins are not packaging ssDNA but assembling to form VLP. To further prove that more sensitive and accurate testing like DNA quantitative method should be performed.

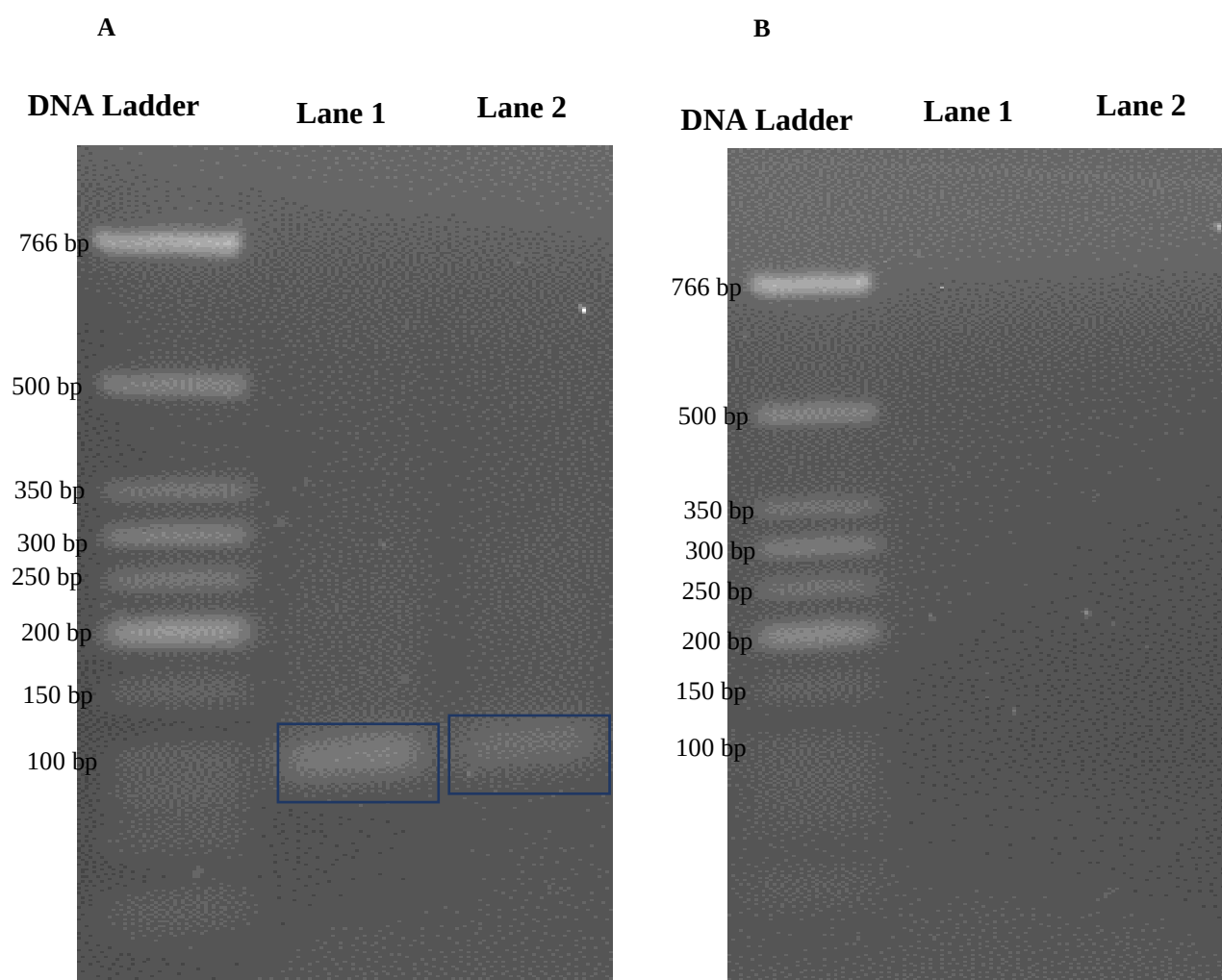


FIGURE 22: 1.8% gel showing PCR amplification of *gfp* gene region. A: PCR amplification of the *gfp* gene from assembled crude extract containing TTV19 capsid protein and ssDNA. B: PCR amplification of the *gfp* gene from assembled His-Tag purified TTV19 capsid protein and ssDNA.

4.4.4. Confirmation of packaging of ssDNA by TTV19 capsid protein and determination of concentration of TTV19 virion particles.

Capsid proteins from crude extract were able to assemble ssDNA *in vitro*, while His-Tag purified capsid proteins were not able to assemble ssDNA *in vitro*. To further validate the results, the DNA quantification approach of quantitative PCR (qPCR) was carried out to confirm ssDNA packaging by capsid proteins *in vitro*. This approach can also be used to determine the concentration of assembled virus particles *in vitro*. To estimate virus

concentration, the standard calibration curve was generated using the plasmid pTTV19 IR_CMV_gfp_M13 (Fig. 23).

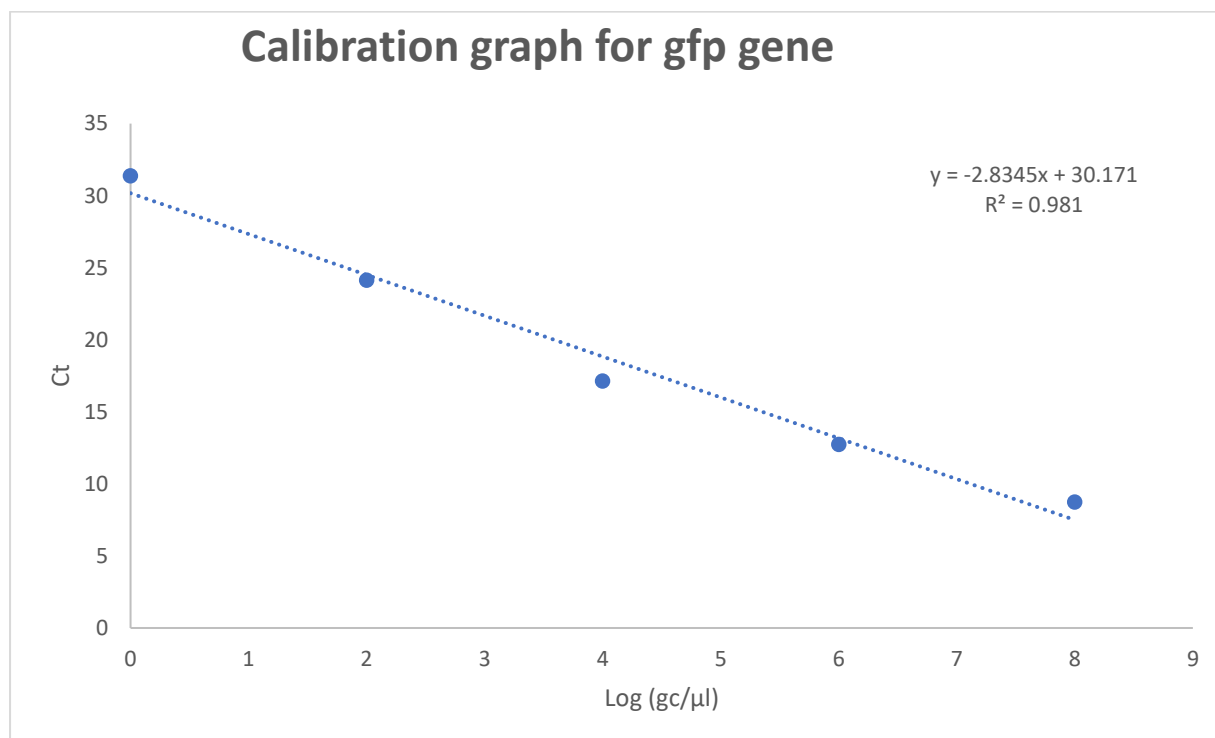


Figure 23: Calibration graph to estimate TTV19 virion concentration. Representative calibration curve for *gfp* gene region measured from qPCR of pTTV19 IR_CMV_gfp_M13 plasmid.

Once, qPCR was run on the samples, it was observed that the ssDNA that was generated by M13 processing was successfully packaged by crude extract containing TTV19 capsid proteins as well as His-Tag purified TTV19 capsid proteins based on the amplification curves obtained (Appendix C) as well as threshold cycle number (Table 12). Single-stranded DNA extracted using heat treatment as well as by phenol-chloroform from assembled capsid protein from the crude extract was successfully amplified, indicating that the DNA was likely packaged by crude extract containing capsid protein. As expected, ssDNA treated with DNase enzyme did not show any amplification. Using the threshold cycle number, the

concentration of virus particles was determined as shown in table 12. Next, we extracted ssDNA using heat treatment as well as phenol-chloroform from assembled His-Tag-purified capsid protein. The product DNA was successfully amplified indicating that the ssDNA was likely packaged by capsid protein within the crude extract. In contrast, extracted ssDNA treated with DNase enzyme did not show any amplification as expected. This result was similar to the one obtained for crude extract containing capsid proteins and similarly, the virus concentration was determined (Table 12).

TABLE 12: C_t Values and virion concentration for the qPCR amplified samples

Test Sample	C_t Value	Virion concentration (No. of virus particles/μl)
ssDNA + Crude extract containing TTV19 capsid protein → DNase Treatment → Heat → DNase	30.43 ± 0.97	NA
ssDNA + Crude extract containing TTV19 capsid protein → DNase Treatment → Heat	20.80 ± 2.31	1.2 x 10 ⁴ ± 1.82 x 10 ²
ssDNA + Crude extract containing TTV19 capsid protein → DNase Treatment → Phenol-chloroform extraction → DNase	32.85 ± 2.90	NA
ssDNA + Crude extract containing TTV19 capsid protein → DNase Treatment → Phenol-chloroform extraction	23.68 ± 0.89	5.73 x 10 ⁴ ± 5.47 x 10 ²
ssDNA + His-Tag Purified TTV19 capsid protein → DNase Treatment → Heat → DNase	31.49 ± 0.47	NA
ssDNA + His-Tag purified TTV19 capsid protein → DNase Treatment → Heat	18.41 ± 0.58	3.01 x 10 ⁴ ± 1.23 x 10 ³
ssDNA + His-Tag purified TTV19 capsid protein → DNase Treatment → Phenol-chloroform extraction → DNase	31.93 ± 0.17	NA
ssDNA + His-Tag Purified TTV19 capsid protein → DNase Treatment → Phenol-chloroform extraction	17.94 ± 0.25	4.19 x 10 ⁴ ± 8.48 x 10 ³
ssDNA + DNase	31.89 ± 0.57	NA

5. DISCUSSION

Anelloviruses are non-enveloped, circular ssDNA negative sense viruses, ranging from 2.9 to 3.9 kb in genomic length (Kaczorowska and Hoek, 2020). They are referred to as ‘orphan viruses’ as they lack disease association and are omnipresent, making themselves a commensal part of the human microbiome (Kaczorowska and Hoek, 2020). The *Anelloviridae* family consists of 30 genera out of which Alphatorquevirus, betatorquevirus, and gammatorquevirus have been the most studied to date (Biagini, 2009; Varsani et al., 2021). These viruses are very similar to each other in nature but differ slightly in genome size (Kaczorowska and Hoek, 2020). These viruses are known to produce 6 different types of proteins: ORF1, the primary capsid precursor protein; ORF2, involved with regulatory functions and immune suppression (Qiu et al., 2005; Sarkar et al., 2016); and four additional of unknown functions (Qiu et al., 2005). Due to the lack of a cell culture system and animal model, this virus is still poorly understood based on limited *in vitro* studies, despite its obvious potential benefits as a redosable gene delivery vehicle. With the help of bioinformatics, anelloviruses may indeed represent a safe and redosble vector for gene therapy and other genetic medicine applications. Anelloviruses circulate in the body often within exosomes which may further promote infectivity by allowing entry of viruses in non-permissive cells, and tolerability, as viruses inside exosomes are less susceptible to the neutralizing effects of antibodies (Martelli et al., 2018). Based on these characteristics, the present study was designed to produce anellovirus capsid proteins using a bacterial expression system. The produced capsid proteins were then used for *in vitro* anellovirus vector assembly.

To produce anellovirus capsid protein a suitable candidate was identified. TTV19, belonging to alphatorquevirus was used as a reference sequence as all the gene sequences for this virus have been identified and it is one of the most abundant anellovirus species found in the

human body (Valenca et al., 2021). Using NCBI BLASTn, other potential candidates were identified, as shown in Table 11. For this study, TTV19 was finally selected as the ideal candidate as it possessed the maximum DNA packaging capacity (3.8 kb) and the gene sequences for capsid proteins and DNA binding signal had been identified. Since no animal models or cell culture systems are available for the propagation of anelloviruses, it is important to investigate other systems for anellovirus production.

Viral vectors are complex, and their production and processing are rife with challenges, such as scalability issues, difficulties maintaining productivity, and an overall lack of robustness (Tariq *et al.*, 2021). In recent years, the use of bacterial systems has become more popular for virus-like particles (VLP) production as this system is a) cost-effective, b) easy to scale up, c) easily manipulated, d) offers high-level expression, e) offers a faster growth rate, f) is genetically stable, and g) and confers simplicity of expression (Kim *et al.*, 2014; Masavuli *et al.*, 2017). VLPs like AAV, HPV, norovirus (NoV), Hepatitis E virus (HEV), etc. have now all been successfully produced in bacterial systems (Chen *et al.*, 2016; Muller *et al.*, 2022; Huo *et al.*, 2018; Gupta *et al.*, 2020). Also, functional viral particles like AAV are known to assemble in the bacterial system (Muller *et al.*, 2022). In addition to bacterial systems, yeast cells, insect cells, mammalian cells, and plants have all been explored as viral and/or VLP production systems. VLPs can be generated in yeast cells by cost-effectively expressing recombinant proteins leading to high expression of vectors that can be effectively scaled to permit commercial-level production (Tariq *et al.*, 2021). Yeast-derived VLP vaccines such as Gardasil® and Gardasil9® against HPV and Mosquirix™, against *Plasmodium falciparum*, have been approved by FDA (Tariq *et al.*, 2021). Baculovirus-based protein expression in insect cell lines (Tariq *et al.*, 2021) represents another effective production system for capsid protein expression and VLP formation that has further been applied as a GT production system for AAV (Kotin *et al.*, 2023). Functional AAV and HSV vectors have been

successfully generated in the Baculoviral system (Moreno *et al.*, 2022; Grady *et al.*, 2014). This expression system is fast and versatile, offers high protein expression, and can perform any necessary eukaryotic post-translational modification (PTM), which may be crucial for the self-assembly and/or release of some virus particle (Chaves *et al.*, 2018). Such PTMs cannot be carried out in bacterial systems, thus limiting the use of bacterial systems for propagation. Although this system has many advantages, it is posed with significant challenges like protein contamination by enveloped baculoviruses, difficulty in scaling, and simpler N-glycosylation PTMs than those generated in mammalian cells (Tariq *et al.*, 2021). Mammalian cells have been used for decades to study protein expression and VLP formation and assembly (Tariq *et al.*, 2021). This system is better compared to other systems due to their capability to properly fold, assemble, and post-translational modification (PTM) to proteins (Dumont *et al.*, 2016). However, such system requires largescale production facilities, like fermentation bioreactors, which are very expensive to construct. Apart from high cost, this system requires lengthy-expression time, has low protein yield, and is susceptible to infections with mammalian pathogens (Tariq *et al.*, 2021). Plant and plant cells are other attractive systems for capsid protein expression and VLP formation (Rybicki, 2020). Plant systems are highly scalable, cost-effective, capable of carrying out PTM, and have a high expression (Rybicki, 2020). However, this system is not very suitable for VLP formation due to low yields of VLPs compared to other systems (Tariq *et al.*, 2021).

This study focused on expressing anellovirus capsid proteins in a bacterial system to explore if this was a viable approach to production. The ORF1 of anellovirus, which encodes its capsid precursor protein, was codon optimized for an *E. coli* bacterial system and amplified via the pET His-tagged system to simplify purification. Apart from His-tag, other tags like FLAG-tag, NE-tag, Strep-tag, Spot-tag, T7-tag, TrpE-tag etc. (Kimple *et al.*, 2013). Using these tags, the expressed proteins can be detected and purified using affinity chromatography.

Interestingly, we did not find that the His-tag posed any issues to TTV19 capsid assembly, suggesting that this may be a scalable and simplistic approach to capsid production. Recently, anellovirus production was successfully demonstrated in insect cells where they were able to form VLPs (Liou et al., 2022). One of the most popular platforms, the *E. coli* cell system, has been successfully used to assemble a variety of VLPs. AAV capsid proteins were previously successfully expressed using a bacterial system. It was found that AAV capsid proteins that were expressed in *E. coli* assembled into VLP's intracellularly that were readily purified (Muller *et al.*, 2022). HPV VLPs are one of the most well-known examples of VLP formation in *E. coli*. The capsid proteins are expressed in *E.coli*, which then self-assemble into VLPs (Chen *et al.*, 2016). Based on this previous study, this approach was similarly applied to this study. The TTV19 capsid expressed in *E. coli* should be ~75 kDa in weight if unprocessed, but while it was observed that initially, the proteins expressed in *E. coli* were ~75 kDa, the protein was further processed to a lower molecular weight and upon purification, yielded proteins of a molecular weight of ~25 kDa. Similar results were obtained by a group of researchers while studying the structure of a virus-like particle (VLP) derived from a Betatorquevirus isolate, called LY1, expressed in SF-9 insect cells (Liou et al., 2022). They found that proteolysis occurs at the C-terminus of the capsid protein and does not play an important role in VLP formation. The enzyme or pathway involved in proteolysis is yet unidentified, but it was suggested that proteolysis of the ORF1 C-terminus may be a natural part of anellovirus formation (Liou et al., 2022). In this study, the reasons for such proteolysis were investigated using the online tool Expasy. It was found that ORF1 of TTV19 had several enzymatic cleavage sites (Appendix D) although however, none of the identified enzymes would have generated the cleavage pattern observed. Interestingly, TTV19 capsid protein was produced in *E. coli* BL21 that are generally protease-deficient, so it is less likely that an *E. coli* protease was responsible for the cleavage. Further, some of the enzymes that

could potentially lead to protein cleavage patterns as observed are not naturally produced in *E. coli* BL21. Other synthetic chemicals cannot be the reason for protein cleavage as they were not provided during the study. While the reason for the cleavage remains unknown, it is very interesting that the final 25 kDa protein does appear to form capsids, which may be a natural process that might be occurring intrinsically within the protein.

As a result of these observations *in vitro* assembly of ssDNA by anellovirus capsid protein was conducted by both His-Tag purified TTV19 capsid protein as well as crude extract containing TTV19 capsid protein to maximize chances of successful packaging. The purpose for using the crude extract containing TTV19 capsid protein for *in vitro* assembly was due to proteolysis a major portion of protein was lost that might be important for capsid assembly.

The second part of the study involved the production of a synthetically originated ssDNA that could be packaged by TTV19 capsid protein *in vitro*. This was achieved using the M13 phage replication system using the helper plasmid, M13SW8, and DNA vector plasmid, pTTV19 IR_CMV_gfp_M13, and produced *in vivo* in *E. coli* host cells. Helper plasmid encodes for proteins necessary for M13 phage replication, transcription, and assembly of M13 phagemids intracellularly in bacterial systems (Wong, 2022). The M13 genome consists of an origin region that has nucleotide sequences for initiation and termination of replication. While the ssDNA to be packaged by M13 phagemids was produced using a DNA vector plasmid, the M13 origin of replication was separated flanked by a mammalian transgene expression cassette, isolating it from the bacterial backbone (Wong, 2022). This permits the separation of the M13 initiation sequence from the termination sequence and the ability to clone a gene cassette of interest between the initiation and termination gene sequences without any intervening prokaryotic sequences or helper phage contamination (Wong, 2022). Using this revolutionary new miniphagemid approach M13 phagemids were produced that packaged ssDNA but also contained the TTV19 DNA binding signal. The extracted ssDNA from M13

phagemids could then be used for *in vitro* assembly. It is interesting to note that this is evidence of cross application of M13 phagemid DNA toward synthetic production of anellovirus. This system should be similarly applicable to reconstitute the original genome sequence of TTV19 in its entirety via this miniphagemid ssDNA production system to synthetically reconstitute a viable anellovirus using *in vivo* generated components from *E. coli*. A similar approach has been used for creating hybrid viruses by combining components of two different viruses that can be used as vectors for various biomedical platforms. One such example of hybrid viruses is the AAV/M13 hybrid vector which has the potential to be used for a variety of applications such as gene therapy, drug delivery, and vaccine development (Suwan *et al.*, 2019). They offer the ability to deliver genes to target cells with high efficiency and specificity, while also displaying proteins or peptides on the surface of the virus for binding and screening purposes (Yoo *et al.*, 2015).

The last step of the study involved *in vitro* assembly of ssDNA and TTV19 capsid protein which was successfully carried out. For *in vitro* assembly, His-Tag purified TTV19 capsid proteins, and crude extract containing TTV19 capsid proteins were used. The purpose of using crude extract containing capsid proteins was the loss of protein fragments during purification. The fragment that was lost during purification might have an important role in the formation of capsid proteins and the assembly of the virus particle. Both crude extract containing TTV19 capsid protein and His-Tag purified capsid protein seemed to successfully assemble ssDNA. Since the ssDNA was observed to be assembled by His-Tag purified proteins that were 25 kDa in size, this potentially suggests that processing of the capsid protein from ~75 kDa to 25 kDa is necessary to produce mature virus particle formation. Additionally, the larger fragment which was lost during purification is not necessary for DNA assembly or virus formation. However, in the case of some viruses like HIV, the precursor

capsid protein (gp 160) is cleaved to produce two glycoproteins, gp 120 and gp 41. Both proteins are important to produce mature virus particles (Martins *et al.*, 2022).

The concentration of virus particles produced in the study was also similar for both ssDNA assembled by crude extract containing TTV19 capsid protein and His-Tag purified TTV19 capsid proteins. This further demonstrates that the contents of the bacterial system do not hinder virus assembly and may actually fully support it and could perhaps even fully support the production and complete synthetic assembly of anelloviruses *in vivo* in an *E. coli* system. Based on this work, the assembly of M13 miniphagemid-derived ssDNA by *E. coli*-produced capsid proteins can be achieved solely using a bacterial system.

6. SUMMARY

In this study, we hypothesized that ssDNA molecules encoding a transgene cistron of interest and carrying a packaging signal for anellovirus can be produced and purified as M13 miniphagemids from host *E. coli* cells. In tandem, anelloviral capsid proteins can be expressed and functionally processed in *E. coli* to yield self-assembling homo- multimeric anellovirus capsids. M13-processed ss circular DNA vector in combination with *E. coli*-produced capsids can self-assemble into circular ssDNA containing anellovirus capsid gene delivery vectors. From the results, it was observed that TTV19 capsid protein can be expressed in *E. coli*. These proteins seem to undergo additional processing and can functionally self-assemble into capsids of the expected size. In combination, ssDNA containing TTV19 DNA binding signal was produced in tandem using the M13 replication system and was obtained from M13 phagemids using phenol-chloroform extraction. TTV19 capsid proteins were able to successfully assemble ssDNA in vitro, thus, supporting the hypothesis of the study. The assembled TTV19 virion particle can be further tested for its transfection efficiency and transgene expression in mammalian cells to understand whether the nucleoprotein composition of this virus is suitable as a GT vector or perhaps further requires exosomal incorporation for effective transduction. Thus, the vector produced in this study is a promising candidate for the gene therapy platform. This approach opens new avenues for the production of novel promising vectors for gene therapy platforms. The application can be further extended to overcome the production limitations of producing cost-effective and safer viral vectors.

7. FUTURE RESEARCH

The expression of anellovirus capsid proteins in a bacterial system and assembly of TTV19 capsid protein and ssDNA *in vitro* has opened new opportunities for biomedical applications. The assembled TTV19 virion particles can be further used for gene therapy platforms by analyzing their transfection efficiency and transgene expression. In this study, components of *in vitro* assembly of TTV19 virions were produced in a bacterial system. This approach can be further extended by investigating the ability of anellovirus capsid proteins to assemble *in vivo* in a bacterial system. This can be achieved by replacing M13 major coat protein, gpVIII with anellovirus ORF1 in helper plasmid M13SW8. This will help the expression of capsid proteins as well as the assembly of ssDNA intracellularly. The assembled virus particles can be tested for their infectivity and gene delivery efficiency. Production of viral vectors *in vivo* using a bacterial system will overcome a lot of limitations surrounding viral vectors for gene therapy platforms. This approach can be extended to other viral vectors like AAV and HSV. However, anellovirus capsid proteins expressed in the bacterial system might not be able to transfect mammalian cells as the system lacks a post-translational modification system found in eukaryotes. This will need to be tested on human cell lines for efficient transduction. Post-translational modification may be necessary for anellovirus capsid proteins to carry out effective transduction, in which a eukaryotic system like insect cell lines may be necessary to express anellovirus capsid proteins.

Apart from being a promising candidate for gene therapy platforms, anelloviruses could play an important role in cancer treatment. ORF 2 of anelloviruses can produce the protein called TTV-derived-Apoptosis-Inducing protein (TAIP), which is 105 amino acids long and has been shown to induce apoptosis in cellular carcinoma cells (Bostan *et al.*, 2013). This ORF2 protein can be used as a display peptide on VLPs for cancer treatment. In summary, anellovirus may be a versatile new therapeutic platform in which case developing efficient

methods for cost-effective production will become essential. In this work, we have substantiated the rationale to further explore the production of anelloviruses by an artificial method based on the packaging of M13 ssDNA miniphagemids.

REFERENCES

- Ahi Y. S., Bangari D. S., and Mittal S. K. (2011). Adenoviral vector immunity: its implications and circumvention strategies. *Curr Gene Ther.*, 11(4):307-320.
- Alba R. B., and Chillon M. (2005). Gutless adenovirus: last-generation adenovirus for gene therapy. *Gene Ther.*, 12: S18-S27.
- Al-Dosari M. S., and Gao X. (2009). Nonviral gene delivery: principle, limitations, and recent progress. *AAPS J.*, 11(4), 671-681.
- Aleti V., Reddy G. B., Parikh K., Arun P., and Chilukuri N. (2013). Persistent and high-level expression of human liver prolidase in vivo in mice using adenovirus. *Chem Biol Interact.*, 203:191–5. doi: 10.1016/j.cbi.2012.08.021
- Andre F., and Mir L. M. (2004). DNA electrotransfer: its principles and an updated review of its therapeutic applications. *Gene Ther.*, 11, S33-S42.
- Antosova Z., Mackova M., Kral V., and Macek T. (2009). Therapeutic application of peptides and proteins: parenteral forever. *Trends Biotechnol.*; 27(11):628–35. Epub 2009/09/22. <https://doi.org/10.1016/j.tibtech.2009.07.009> PMID: 19766335.
- Anwer K., Kao G., Proctor B., Anscombe I., Florack V., Earls R., Wilson E., McCreery T., Unger E., Rolland A., and Sullivan S. M. (2000). Ultrasound enhancement of cationic lipid-mediated gene transfer to primary tumours following systemic administration. *Gene Ther.*, 7(21), 1833-1839
- Asabe S., Nishizawa T., Iwanari H. (2001). Phosphorylation of serine-rich protein encoded by open reading frame 3 of the TT virus genome. *Biochem. Biophys. Res. Commun.*, 286:298–304.
- Barber M. A. (1911). A technic for the inoculation of bacteria and other substances into living cells. *J. Infect. Dis.*, 8(3), 348-360.

- Barolet D., and Benohanian A. (2018). Current trends in needle-free jet injection: an update. *Clinical, Cosmetic and Investigational Dermatology*, 11 231–238
- Barry M. A., Dower W. J., and Johnston S. A. (1996). Toward cell-targeting gene therapy vectors: selection of cell-binding peptides from random peptide presenting phage libraries. *Nat. Med.*, 2(3):299-305.
- Bendinelli M., Pistello M., Maggi F. (2001). Molecular properties, biology, and clinical implications of TT virus, a recently identified widespread infectious agent of humans. *Clin. Microbiol. Rev.*, 14: 98–113.
- Benhar I. (2001). Biotechnological applications of phage and cell display. *Biotechnol. Adv.*, 19(1):1-33.
- Bennett N. J., Gagic D., Sutherland-Smith A. J., and Rakonjac J. (2011). Characterization of a dual-function domain that mediates membrane insertion and excision of Ff filamentous bacteriophage. *J. Mol. Biol.*, 411(5), 972–985.
- Biagini P. (2009). Classification of TTV and related viruses (anelloviruses). *Curr. Top. Microbiol. Immunol.*, 331:21–33.
- Biagini P., Bendinelli M. S. H., King A. M. Q., Adams M. J., Carstens E. B., and Lefkowitz E. J. (2011). *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*. Elsevier Scientific Publ. Co., pp. 331-341.
- Blaese R. M., Culver K. W., Miller A. D., Carter C. S., Fleisher T., Clerici M., Shearer G., Chang L., Chiang Y., Tolstoshev P., Greenblatt J. J., Rosenberg S. A., Klein H., Berger M., Mullen C. A., Ramsey W. J., Muul L., Morgan R. A., and Anderson W. F. (1995). T-Lymphocyte-Directed Gene Therapy for ADA-SCID: Initial Trial Results After 4 Years. *Science*, 270, 475–480.

- Boohaker R. J., Lee M. W., Vishnubhotla P., Perez J. M., and Khaled A. R. (2019). The use of therapeutic peptides to target and to kill cancer cells. *Curr. Med. Chem.*; 19(22):3794–804. Epub 2012/06/26. PMID: 22725698; PubMed Central PMCID: PMC4537071.
- Bostan N., Amen N., and Bokhari H. (2013). Current and future prospects of Torque teno virus. *J. Vaccines Vaccin.*, S:1. DOI: 10.4172/2157-7560.S1-004
- Boulaiz H., Marchal J. A., Prados J., Melguizo C., and Aranega A. (2005). Non-viral and viral vectors for gene therapy. *Cell Mol. Biol.*, 51(1), 3-22.
- Bruno B. J., Miller G. D., and Lim C. S. (2013) Basics and recent advances in peptide and protein drug delivery. *Ther Deliv.*; 4(11):1443–67. Epub 2013/11/16. <https://doi.org/10.4155/tde.13.104> PMID: 24228993; PubMed Central PMCID: PMC3956587.
- Burton E. A., Fink D. J., and Glorioso J. C. (2002). Gene delivery using herpes simplex virus vectors. *DNA Cell Biol.*, 21(12):915-936
- Canatella P. J., and Prausnitz M. R. (2001). Prediction and optimization of gene transfection and drug delivery by electroporation. *Gene Ther.*, 8(19), 1464-1469.
- Cartier N., Hacein-Bey-Abina S., Bartholomae C. C. (2012). Lentiviral hematopoietic cell gene therapy for X-linked adrenoleukodystrophy. *Methods Enzymol.*, 507:187-198.
- Chang E., Jin S., Nam M., Kim S. (2019). Human Induced Pluripotent Stem Cells: Clinical Significance and Applications in Neurologic Diseases. *J Korean Neurosurg Soc* 62 (5) : 493-501.
- Chaves L. C., Ribeiro B. M., and Blissard G. W. (2018). Production of GP64 free virus-like particles from baculovirus-infected insect cells. *J. Gen. Virol.* 99, 265–274.

- Chen H. W., Huang C. Y., Lin S. Y., Fang Z. S., Hsu C. H., and Lin J. C. (2016). Synthetic virus-like particles prepared via protein corona formation enable effective vaccination in anavian model of corona virus infection. *Biomaterials* 106,111–118.
- Chen Y. H., Keiser M. S., and Davidson B. L. (2018). Viral vectors for gene transfer. *Curr. Protoc. Mouse Biol.*, 8(4):e58.
- Chi Z., Xu Q., and Zhu L. (2020). A Review of Recent Advances in Robotic Cell Microinjection. *IEEE access* Volume 8.
- Chilukuri N., Duysen E. G., Parikh K., Sun W., Doctor B. P., Lockridge O., and Saxena A. (2008). Adenovirus-mediated gene transfer of human butyrylcholinesterase results in persistent high-level transgene expression in vivo. *Chem Biol Interact.*, 175:327–31. doi: 10.1016/j.cbi.2008.04.009
- Chira S., Jackson C. S., Oprea I., Ozturk F., Pepper M.S., Diaconu I., Braicu C., Raduly L., Calin G. A., and Berindan-Neagoe I. (2015). Progresses towards safe and efficient gene therapy vectors. *Oncotarget* Vol. 6, No. 31.
- Chuah M. K., Collen D., and Vanden Driessche T. (2003). Biosafety of adenoviral vectors. *Curr. Gene Ther.*, 3(6):527-543.
- Clark J. R., and March J. B. (2006). Bacteriophages and biotechnology: vaccines, gene therapy and antibacterials. *Trends Biotechnol.*, 24(5):212-218.
- Clokier M. R., Millard A. D., Letarov A. V., and Heaphy S. (2011). Phages in nature. *Bacteriophage* 1(1):31-45. doi:10.4161/bact.1.1.14942.
- Cohen I. G., and Adashi E. Y. (2016). The FDA is prohibited from going germline. *Science.*, 353:545–6. doi: 10.1126/science. aag2960

- Colella P., Ronzitti G., and Mingozzi F. (2018). Emerging issues in AAV mediated in vivo gene therapy. *Mol Ther. Methods Clin. Dev.*, 8:87-104.
- Daud A. I., DeConti R. C., Andrews S., Urbas P., Riker A. I., Sondak V. K., Munster P. N., Sullivan D. M., Ugen K. E., Messina J. L., and Heller R. (2008). Phase I trial of interleukin-12 plasmid electroporation in patients with metastatic melanoma. *J. Clin. Oncol.*, 26(36), 5896-5903.
- Davison A. J., Benko M., and Harrach B. (2003). Genetic content and evolution of adenoviruses. *J. Gen. Virol.*, 84:2895-2908
- de Villiers E. M., Borkosky S. S., Kimmel R. (2011). The diversity of torque teno viruses: in vitro replication leads to the formation of additional replication-competent subviral molecules. *J. Virol.*, 85:7284–95.
- Deev R., Plaksa I., Bozo I., and Isaev A. (2017). Results of an International post-marketing surveillance study of pl-VEGF165 Safety and Efficacy in 210 Patients with Peripheral Arterial Disease. *Am. J. Cardiovasc. Drugs.*, 17:235–42. doi: 10.1007/s40256–016-0210–3
- Deng X., Terunuma H., Handema R. (2000). Higher prevalence and viral load of TT virus in saliva than in the corresponding serum: another possible transmission route and replication site of TT virus. *J. Med. Virol.*, 62:531–7.
- Dissanayake E., and Inoue Y. (2016). MicroRNAs in allergic disease. *Curr. Allergy Asthma Rep.* 16:67. doi: 10.1007/s11882-016-0648-z
- Donnelly A., Teerapong Y, Kaoutar B., Keittisak S., and Amin H. (2015). Bacteriophage Mediates Efficient Gene Transfer in Combination with Conventional Transfection Reagents. *Viruses*: 7, 6476–6489; doi:10.3390/v7122951

- Duan L., Xu L., Xu X., Qin Z., Zhou X., Xiao Y., Liang Y., and Xia J. (2021). Exosome-mediated delivery of gene vectors for gene therapy. *Nanoscale*, 13,1387–1397
- Dumont J., Euwart D., Mei B., Estes S., and Kshirsagar R. (2016). Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. *Crit. Rev. Biotechnol.* 36, 1110–1122.
- Dunbar C. E., High K. A., Joung J. K., Kohn D. B., Ozawa K., and Sadelain M. (2018). Gene therapy comes of age. *Science.*, 359: ean4672. doi: 10.1126/science. aan4672
- Dunn I. S. (1996). Mammalian cell binding and transfection mediated by surfacemodified bacteriophage lambda. *Biochimie.*, 78(10):856-861.
- Durieux A. C., Bonnefoy R., Busso T., and Freyssen D. (2004). In vivo gene electrotransfer into skeletal muscle: effects of plasmid DNA on the occurrence and extent of muscle damage. *J. Gene Med.*, 6(7), 809-816.
- Earley L. F., Powers J. M., Adachi K. (2017). Adeno-associated virus (AAV) assembly-activating protein is not an essential requirement for capsid assembly of AAV serotypes 4, 5, and 11. *J. Virol.*, 91(3):e01980-16.
- Edens L., Konings R. N. H., and Schoenmakers, J. G. G. (1978). Cascade Mechanism of transcription in bacteriophage M13-DNA. *Virology* 86(2),354–367.
- Efstathiou S., and Minson A. C. (1995). Herpes virus-based vectors. *Br Med Bull.*, 51(1):45-55.
- Eibach D., Hogan B., Sarpong N., Winter D., Struck N. S., Adu-Sarkodie Y., Owusu-Dabo E., Schmidt-Chanasit J., May J., and Cadar D. (2019). Viral metagenomics revealed novel betatorquevirus species in pediatric inpatients with encephalitis/ meningoencephalitis from Ghana. *Scientific Reports*, 9:2360.

- Endoh M., Koibuchi N., Sato M., Morishita R., Kanzaki T., Murata Y., and Kaneda Y. (2002). Fetal gene transfer by intrauterine injection with microbubble-enhanced ultrasound. *Mol. Ther.*, 5(5), 501-508.
- Erker J. C., Leary T. P., Desai S. M. (1999). Analyses of TT virus full-length genomic sequences. *J. Gen. Virol.*, 80:1743–50.
- Fosgerau K., and Hoffmann T. (2015). Peptide therapeutics: current status and future directions. *Drug Discov. Today*; 20(1):122–8. Epub 2014/12/03. <https://doi.org/10.1016/j.drudis.2014.10.003> PMID: 25450771.
- Freer G., Maggi F., Pifferi M., Di Cicco M. E., Peroni D. G., Pistello M. (2018). The Virome and Its Major Component, Anellovirus, a Convolved System Molding Human Immune Defenses and Possibly Affecting the Development of Asthma and Respiratory Diseases in Childhood. *Frontiers in Microbiology* 9: 7
- Friedmann T. (1992). A brief History of Gene therapy. *Nature Genetics* Volume 2.
- Friedmann T., and Roblin R. (1972). Gene Therapy for Human Genetic Disease. *Science*, 175, 949–955. 13
- Furuta R. A., Sakamoto H., Kuroishi A. (2015). Metagenomic profiling of the viromes of plasma collected from blood donors with elevated serum alanine aminotransferase levels. *Transfusion*, 55:1889–99.
- Gao X., Kim K. S., and Liu D. X. (2007). Nonviral gene delivery: What we know and what is next. *AAPS. J.*, 9(1), E92-E104.
- Gehl J., Skovsgaard T., and Mir L. M. (2002). Vascular reactions to in vivo electroporation: characterization and consequences for drug and gene delivery. *Bba-Gen Subjects*, 1569 (1-3), 51-58.

- Ghosh S., Brown A. M., Jenkins C., and Campbell K. (2020). Viral Vector Systems for Gene Therapy: A Comprehensive Literature Review of Progress and Biosafety Challenges. *Applied Biosafety: Journal of ABSA International*, Vol. 25(1) 7-18.
- Gissel H., and Clausen T. (2001). Excitation-induced Ca²⁺ influx and skeletal muscle cell damage. *Acta Physiol. Scand.*, 171(3), 327-334.
- Goins W. F., Huang S., Cohen J. B., and Glorioso J. C. (2014). Engineering HSV-1 vectors for gene therapy. *Methods Mol Biol.*, 1144:63-79
- Goins W. F., Wolfe D., Krisky D. M. (2004). Delivery using herpes simplex virus: an overview. *Methods Mol Biol.*, 246: 257-299.
- Goswami R., Subramanian G., Silayeva L., Newkirk I., Doctor D., Chawla K., Chattopadhyay S., Chandra D., Chilukuri N., and Betapudi V. (2019). Gene Therapy Leaves a Vicious Cycle. *Front. Oncol.*, 24.
- Goto K., Sugiyama K., Ando T. (2000). Detection rates of TT virus DNA in serum of umbilical cord blood, breastmilk and saliva. *Tohoku. J. Exp. Med.*, 191:203–7.
- Guan X., Guo Z., Wang T., Lin L., Chen J., Tian H. (2017). A pH- Responsive Detachable PEG shielding strategy for gene delivery system in cancer therapy. *Biomacromolecules*, 18:1342–9. doi: 10.1021/acs.biomac.7b00080
- Gunzburg W. H., Fleuchaus A., Saller R., and Salmons B. (1996). Retroviral vector targeting for gene therapy. *Cytokines Mol. Ther.*, 2(3):177-184.
- Gutierrez C., Ramirez-Parra E., Castellano M. M., Sanz-Burgos A. P., Luque A, Missich R. (2004). Geminivirus DNA replication and cell cycle interactions. *Vet. Microbiol.*, 98:111–9.
- Haigh N. G., Webster R. E. (1999). The pI and pXI assembly proteins serve separate and essential roles in filamentous phage assembly. *J.Mol.Biol.*, 293(5), 1017–1027.

- Hajitou A., Lev D. C., Hannay J. A., Korchin B., Staquicini F. I., Soghomonyan S., Alauddin M. M., Benjamin R. S., Pollock R. E., Gelovani J. G. (2008). A preclinical model for predicting drug response in soft-tissue sarcoma with targeted aavp molecular imaging. *Proc. Natl. Acad. Sci. USA.*, 105:4471–4476. doi: 10.1073/pnas.0712184105.
- Hall S., and Agrawal D. K. (2018). Delivery of viral vectors for gene therapy in intimal hyperplasia and restenosis in atherosclerotic swine. *Drug Deliv Transl Res.*, 8(4):918-927.
- Hanna E., Remuzat C., Auquier P., and Toumi M. (2017). Gene therapies development: slow progress and promising prospect. *J. Mark Access Health Policy.*, 5:1265293. doi: 10.1080/20016689.2017.1265293
- Hart S. L., Knight A. M., Harbottle R. P., Mistry A., Hunger H. D., Cutler D. F. (1994). Cell binding and internalization by filamentous phage; displaying a cyclic Arg-Gly-Asp-containing peptide. *J. Biol. Chem.*, 269(17):12468-12474.
- Heller L. C., and Heller R. (2006). In vivo electroporation for gene therapy. *Hum. Gene Ther.*, 17(9), 890-897
- Herweijer H., and Wolff J. A. (2003). Progress and prospects: naked DNA gene transfer and therapy. *Gene Ther.*, 10 6), 453-458.
- Ibraheem D., Elaissari A., and Fessi H. (2014). Gene therapy and DNA delivery systems. *Int J Pharm.*, 459:70–83. doi: 10.1016/j.ijpharm.2013.11.041
- Inami T., Konomi N., Arakawa Y. (2000). High prevalence of TT virus DNA in human saliva and semen. *J. Clin. Microbiol.*, 38:2407–8.
- Itoh Y., Takahashi M., Fukuda M. (2000). Visualization of TT virus particles recovered from the sera and feces of infected humans. *Biochem. Biophys. Res. Commun.*, 279:718–24.

- Itoh Y., Takahashi M., Fukuda M. (2000). Visualization of TT virus particles recovered from the sera and feces of infected humans. *Biochem. Biophys. Res. Commun.*, 279:718–24.
- Ivanenkov V. V., and Menon A. G. (2000). Peptide-mediated transcytosis of phage display vectors in MDCK cells. *Biochem. Biophys. Res. Commun.*, 276(1):251-257.
- Jones M. S., Kapoor A., Lukashov V. V. (2005). New DNA viruses identified in patients with acute viral infection syndrome. *J. Virol.*, 79:8230–6.
- Kaczorowska J., and van der Hoek L. (2020). Human anelloviruses: diverse, omnipresent and commensal members of the virome. *FEMS Microbiology Reviews*, fuaa007, 44, 305–313.
- Kakkola L., Hedman K., Qiu J. (2009). Replication of and protein synthesis by TT viruses. *TT Viruses*. Springer, 53–64.
- Kakkola L., Tommiska J., Boele L. C. (2007). Construction and biological activity of a full-length molecular clone of human Torque teno virus (TTV) genotype 6. *FEBS J.*, 274:4719–30.
- Kamahora T., Hino S., and Miyata H. (2000). Three spliced mRNAs of TT virus transcribed from a plasmid containing the entire genome in COS1 cells. *J. Virol.*, 74:9980–6.
- Kapusinszky B., Minor P., Delwart E. (2012). Nearly constant shedding of diverse enteric viruses by two healthy infants. *J. Clin. Microbiol.*, 50:3427–34.
- Khare R., Chen C. Y., Weaver E. A., and Barry M. A. (2011). Advances and future challenges in adenoviral vector pharmacology and targeting. *Curr. Gene. Ther.*, 11(4):241-258.
- Kimble M. E., Brill A. L., and Pasker L. (2013). Overview of affinity tags for protein purification. *Curr. Protoc. Protein Sci.*, 73.
- Klein T. M., Wolf E. D., Wu R., and Sanford, J. C. (1987). High-velocity microprojectiles for delivering nucleic-acids into living cells. *Nature*, 327(6117), 70-73.

- Kleiner M., Hooper L. V., and Duerkop B. A. (2015). Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes. *BMC genomics*, 16:7.
- Koch S., Pohl P., Cobet U., and Rainov N. G. (2000). Ultrasound enhancement of liposome-mediated cell transfection is caused by cavitation effects. *Ultrasound Med. Biol.*, 26(5), 897-903.
- Koelle D. M., and Corey L. (2003). Recent progress in herpes simplex virus immunobiology and vaccine research. *Clin Microbiol Rev.*, 16(1):96-113.
- Kootstra N. A., and Verma I. M. (2003). Gene therapy with viral vectors. *Annu. Rev. Pharmacol. Toxicol.*, 43:413-439.
- Korzh V., and Strähle U. (2002). Marshall Barber and the century of microinjection: from cloning of bacteria to cloning of everything. *Differentiation*, 70:221–226.
- Kotin R. M., Destro F., Srinivasan P., Kanter J. M., Neufeld C., Wolfrum J. M., Barone P. W., Springs S. L., Sinskey A. J., Cecchini S., and Braatz R. D. (2023). Mechanistic Modeling Explains the Production Dynamics of Recombinant Adeno-Associated Virus with the Baculovirus Expression Vector System. Cold Spring Harbor Laboratory, bioRxiv.
- Krieg A. M. (2002). CpG motifs in bacterial DNA and their immune effects. *Annu. Rev. Immunol.* 20, 709–760. doi: 10.1146/annurev.immunol.20.100301.064842
- Kurian K. M., Watson C. J., and Wyllie A. H. (2000). Retroviral vectors. *Mol Pathol.*, 53(4):173-176.
- Kuzmin D. A., Shutova M. V., Johnston N. R., Smith O. P., Fedorin V. V., Kukushkin Y. S., van der Loo J. C. M., and Johnstone E. C. (2021). The clinical landscape for AAV gene therapies. *Nat. Rev. Drug Discov.*, 20:173-174.
- Lachmann R. (2004). Herpes simplex virus-based vectors. *Int. J. Exp. Pathol.*, 85(4):177-190.

- Lagasse H. A., Alexaki A., Simhadri V. L., Katagiri N. H., Jankowski W., Sauna Z. E., and Kimchi-Safarty C. (2017). Recent advances in (therapeutic protein) drug development. *F1000Res.*, 6:113. doi: 10.12688/f1000research.9970.1
- Lang L. H. (2006). FDA approves use of bacteriophages to be added to meat and poultry products. *Gastroenterology*, 131:1370.
- Larocca D., Kassner P. D., Witte A., Ladner R. C., Pierce G. F., and Baird A. (1999). Gene transfer to mammalian cells using genetically targeted filamentous bacteriophage. *FASEB J.*, 13(6):727-734.
- Lau J. L., and Dunn M. K. (2018). Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorganic & Medicinal Chemistry* 26, 700–2707.
- Leppik L., Gunst K., Lehtinen M. (2007). In vivo and in vitro intragenomic rearrangement of TT viruses. *J. Virol.*, 81:9346–56.
- Liou S., Cohen N., Zhang Y., Acharekar N. M., Rodgers H., Islam S., Zeheb L., Pitts J., Arze C., Swaminathan H., Yozwiak N., Ong T., Hajjar R. J., Chang Y., Swanson K. A., and Delagrave S. (2022). Cold Spring Harbor Laboratory. doi: <https://doi.org/10.1101/2022.07.01.498313>
- Liu Y. P., and Berkhout B. (2014). HIV-1-based lentiviral vectors. *Methods Mol Biol.*, 1087:273-284.
- Losordo D. W., Vale P. R., Hendel R. C., Milliken C. E., Fortuin F. D., Cummings N., Schatz R. A., Asahara T., Isner J. M., and Kuntz R. E. (2002). Phase 1/2 placebo-controlled, double-blind, dose escalating trial of myocardial vascular endothelial growth factor 2 gene transfer by catheter delivery in patients with chronic myocardial ischemia. *Circulation*, 105(17), 2012-2018.

- Lowenstein P. R., Mandel R. J., Xiong W. D., Kroeger K., and Castro M. G. (2007). Immune responses to adenovirus and adeno-associated vectors used for gene therapy of brain diseases: the role of immunological synapses in understanding the cell biology of neuroimmune interactions. *Curr. Gene Ther.*, 7(5):347-360.
- Lui V. W. Y., Falo L. D., and Huang L. (2001). Systemic production of IL-12 by naked DNA mediated gene transfer: toxicity and attenuation of transgene expression in vivo. *J. Gene Med.*, 3(4), 384-393.
- Lukashev a. n., and Zamyatnin A. A. (2016). Viral Vectors for Gene Therapy: Current State and Clinical Perspectives. *ISSN 0006_2979, Biochemistry (Moscow), Vol. 81, No. 7, pp. 700_708.*
- Lundstorm K. (2003). Latest development in viral vectors for gene therapy. *TRENDS in Biotechnology*, Vol.21 No.3.
- Lundstrom K. (2019). RNA viruses as tools in gene therapy and vaccine development. *Genes (Basel)*, 10(3):E189.
- Lysakowski C., Dumont L., Tramer M. R., and Tassonyi E. (2003). A needle-free jet-injection system with lidocaine for peripheral intravenous cannula insertion: a randomized controlled trial with cost effectiveness analysis. *Anesth. Analg.*, 96(1), 215-219.
- Malathi V. G., and Renuka Devi 9. (2019). ssDNA viruses: key players in global virome. *Virus Dis.*, 30(1):3–12
- Manservigi R., Argnani R., and Marconi P. (2010). HSV recombinant vectors for gene therapy. *Open Virol J.*, 4:123-156.
- Manunta M. D., Tagalakakis A. D., Attwood M., Aldossary A. M., Barnes J. L., Munye M. M. (2017). Delivery of siRNA to epithelial cells mediated by a targeted nanocomplex: a therapeutic strategy for cystic fibrosis. *Sci. Rep.*, 7:700. doi: 10.1038/s41598-017-00662-2

- Manzin A., Mallus F., Macera L., Maggi F., and Blois S. (2015). Global impact of Torque teno virus infection in wild and domesticated animals. *Journal of Infection in Developing Countries* 9(6): 562-570
- Marciano D. K., Russel M., and Simon S. M. (1999). Anaquous channel for filamentous phage export. *Science*, 284(5419), 1516–1519.
- Martins S. A., Santos J., Silva R. D. M., Rosa C., Verde S. C., Correia J. D. G., and Melo S. (2022). How promising are HIV-1-based virus-like particles for medical applications. *Front. Cell. Infect. Microbiol.* Vol. 12
- Matsubara H., Michitaka K., Horiike N. (2000). Existence of TT virus DNA in extracellular body fluids from normal healthy Japanese subjects. *Intervirology*, 43:16–19.
- Maurer A. C., Pacouret S., Cepeda D. A. K, Blake J., AndresMateos E., and Vandenberghe L. H. (2018). The assembly-activating protein promotes stability and interactions between AAV's viral proteins to nucleate capsid assembly. *Cell Rep.*, 23(6):1817-1830
- McCarthy M. (2015). Scientists call for moratorium on clinical use of human germline editing. *BMJ.*, 351:h6603. doi: 10.1136/bmj.h6603.
- McClements M. E., and MacLaren R. E. (2017). Adeno-associated virus (AAV) dual vector strategies for gene therapy encoding large transgenes. *Yale J. Biol. Med.*, 90(4):611-623.
- Meissner A., Wernig M., and Jaenisch R. (2007). Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nat. Biotechnol.*, 25: 1177-1181.
- Merten O. W., Hebben M., and Bovolenta C. (2016). Production of lentiviral vectors. *Mol. Ther. Methods Clin Dev.*, 3:16017.

- Mir L. M., Bureau M. F., Gehl J., Rangara R., Rouy D., Caillaud J. M., Delaere P., Branellec D., Schwartz B., Scherman D. (1999). High-efficiency gene transfer into skeletal muscle mediated by electric pulses. *Proc. Natl. Acad. Sci. USA*, 96(8), 4262- 4267.
- Miyata H., Tsunoda H., Kazi A. (1999). Identification of a novel GC- rich 113-nucleotide region to complete the circular, singlestranded DNA genome of TT virus, the first human circovirus. *J. Virol.*, 73:3582–6.
- Molnar M. J., Gilbert R., Lu Y. F., Liu A. B., Guo A., Larochelle N., Orlopp K., Lochmuller H., Petrof B. J., Nalbantoglu J., and Karpati G. (2004). Factors influencing the efficacy, longevity, and safety of electroporation-assisted plasmid-based gene transfer into mouse muscles. *Mol. Ther.*, 10(3), 447-455.
- Morrow T. (2017). Novartis’s kymriah: harnessing immune system comes with worry about reining in costs. *Manag. Care.*, 26:28–30.
- Morsy M. A., and Caskey C. T. (1999). Expanded-capacity adenoviral vectors— the helper-dependent vectors. *Mol. Med. Today*, 5(1):18-24.
- Moses P. B., Boeke J. D., Horiuchi K., and Zinder N. D. (1980). Restructuring the bacteriophage F1genome-expression of gene8 in the intergenic space. *Virology*104 (2), 267–278.
- Muller K. M., Le D. T., Radukic M. T., Teschner K., and Becker L. (2022). Synthesis and Concomitant Assembly of Adeno-Associated Virus like Particles in *Escherichia coli*. *ACS Synth. Biol.* 11, 3601–3607
- Naganuma M., Tominaga N., Miyamura T. (2008). TT virus prevalence, viral loads and genotypic variability in saliva from healthy Japanese children. *Acta. Paediatr.*, 97:1686–90.

- Nakanishi M., Eguchi A., Akuta T., Nagoshi E., Fujita S., Okabe J. (2003). Basic peptides as functional components of non-viral gene transfer vehicles. *Curr. Protein Pept. Sci.*, 4(2):141-150.
- Naldini L. (2015). Gene therapy returns to centre stage. *Nature.*, 526:351–60. doi: 10.1038/nature15818
- Nance M. E., Hakim C. H., Yang N. N., and Duan D. (2017). Nanotherapy for duchenne muscular dystrophy. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.*, 10: e1472. doi: 10.1002/wnan.1472
- Naso M. F., Tomkowicz B., Perry W. L., and Strohl W. R. (2017). Adeno-associated virus (AAV) as a vector for gene therapy. *BioDrugs.*, 31(4):317-334.
- Neumann E., Schaefferidder M., Wang Y., and Hofschneider P. H. (1982). Gene-transfer into mouse lyoma cells by electroporation in high electric-fields. *Embo. J.*, 1(7), 841-845.
- Nienhuis A. W., and Persons D. A. (2012). Development of gene therapy for thalassemia. *Cold Spring Harb Perspect Med.*, 2(11).
- Ninomiya M., Nishizawa T., Takahashi M. (2007). Identification and genomic characterization of a novel human torque teno virus of 3.2 kb. *J. Gen. Virol.*, 88:1939–44.
- Ninomiya M., Nishizawa T., Takahashi M., Lorenzo F. R., Shimosegawa T., and Okamoto H. (2007). Identification and genomic characterization of a novel human torque teno virus of 3.2 kb. *J. Gen. Virol.*, 88:1939–1944.
- Nishi T., Yoshizato K., Yamashiro S., Takeshima H., Sato K., Hamada K., Kitamura I., Yoshimura T., Saya H., Kuratsu J., and Ushio, Y. (1996). High-efficiency in vivo gene transfer using intraarterial plasmid DNA injection following in vivo electroporation. *Cancer Res.*, 56(5), 1050-1055.

- Nishiyama S., Dutia B. M., Stewart J. P. (2014). Identification of novel anelloviruses with broad diversity in UK rodents. *J. Gen. Virol.*, 95:1544–53.
- Nishiyama S., Dutia B. M., Stewart J. P. (2014). Identification of novel anelloviruses with broad diversity in UK rodents. *J. Gen. Virol.*, 95:1544–53.
- Nishizawa T., Okamoto H., Konishi K., Yoshizawa H., Miyakawa Y., and Mayumi M. (1997). A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. *Biochem. Biophys. Res. Commun.*, 241:92–97.
- Nozaki T., Ogawa R., Feril L. B., Kagiya G., Fuse H., and Kondo T. (2003). Enhancement of ultrasound-mediated gene transfection by membrane modification. *J. Gene Med.*, 5(12), 1046-1055.
- Ogawa R., Kagiya G., Feril L. B., Nakaya N., Nozaki T., Fuse H., and Kondo T. (2004). Ultrasound mediated intravesical transfection enhanced by treatment with lidocaine or heat. *J. Urology*, 172(4), 1469-1473.
- Palu G., Parolin C., Takeuchi Y., and Pizzato M. (2000). Progress with retro-viral gene vectors. *Rev Med Virol.*, 10(3):185-202.
- Peng Y., Nishizawa T., Takahashi M. (2002). Analysis of the entire genomes of thirteen TT virus variants classifiable into the fourth and fifth genetic groups, isolated from viremic infants. *Arch Virol.*, 147:21–41.
- Perez B. A., Shutterly A., Chan Y. K. Byrne B. J., and Corti M. (2020). Management of Neuroinflammatory Responses to AAV-Mediated Gene Therapies for Neurodegenerative Diseases. *Brain Sci* 10:119.

- Piersanti S, Cherubini G, Martina Y, Salone B, Avitabile D, Grosso F, et al: Mammalian cell transduction and internalization properties of lambda phages displaying the full-length adenoviral penton base or its central domain. *J Mol Med* 2004, 82(7):467-476.
- Pranjol M. Z., and Hajitou A. (2015). Bacteriophage-derived vectors for targeted cancer gene therapy. *Viruses*, 7:268–284. doi: 10.3390/v7010268.
- Prondzynski M., Mearini G., and Carrier L. (2018). Gene therapy strategies in the treatment of hypertrophic cardiomyopathy. *Pflugers Arch-Eur. J. Physiol. P.*, 1–9. doi: 10.1007/s00424-018-2173-5.
- Qiu J., Kakkola L., Cheng F. (2005). Human circovirus TT virus genotype 6 expresses six proteins following transfection of a fulllength clone. *J. Virol.*, 79:6505–10.
- Raes L., Stremersch S., Fraire J. C., Brans T., Goetgeluk G., Munter S. D., Hoecke L. V., Verbeke R., Hoecke J. V., Xiong R., Saelens X., Vandekerckhove B., Smedt S. D., Raemdonck K., and Braekmans K. (2020). Intracellular Delivery of mRNA in Adherent and Suspension Cells by Vapor Nanobubble Photoporation. *Nano-Micro Lett.*, 12:185.
- Rajotte D., Arap W., Hagedorn M., Koivunen E., Pasqualini R., and Ruoslahti E. (1998). Molecular heterogeneity of the vascular endothelium; revealed by in vivo phage display. *J. Clin. Invest.*, 102(7):430-437.
- Ramamoorth M., and Narvekar A. (2015). Non Viral Vectors in Gene Therapy- An Overview. *Journal of Clinical and Diagnostic Research.*, Vol-9(1): GE01-GE06.
- Ramezani A., and Hawley R. G. (2002). Overview of the HIV-1 lentiviral vector system. *Curr. Protoc. Mol. Biol.*, 16: 16.21.
- Ren S., Li M., Smith, J. M., DeTolla L. J., and Furth P. A. (2002). Low volume jet injection for intradermal immunization in rabbits. *BMC Biotechnol.*, 2, 10.

Robbins P. D., Tahara H., and Ghivizzani S. C. (1998). Viral vectors for gene therapy. TIBTECH., VOL.16

Robinton D. A., and Daley G. Q. (2012). The promise of induced pluripotent stem cells in research and therapy. *Nature* 481 : 295-305.

Roingeard P., Raynal P., Eymieux S., and Blanchard E. (2019). Virus detection by transmission electron microscopy: Still useful for diagnosis and a plus for biosafety. *Rev. Med. Virol.*, 29:e2019.

Romano G. (2012). Development of safer gene delivery systems to minimize the risk of insertional mutagenesis-related malignancies: a critical issue for the field of gene therapy. *ISRN Oncol.*, 2012:616310

Romano G., Marino I. R., Pentimalli F., Adamo V., and Giordano A. (2009). Insertional mutagenesis and development of malignancies induced by integrating gene delivery systems: implications for the design of safer gene-based interventions in patients. *Drug News Perspect.*, 22(4):185-196.

Roossinck M. J. (2015). Move over, bacteria! Viruses make their mark as mutualistic microbial symbionts. *J Virol.*, 89:6532–5. doi: 10.1128/jvi.02974–14

Roossinck M. J., and Bazan E. R. (2017). Symbiosis: viruses as intimate partners. *Annu Rev Virol.*, 4:123–39. doi: 10.1146/annurev-virology-110615–042323

Rosario K., Duffy S., and Breitbart M. (2012). A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. *Arch. Virol.*, 157:1851–71.

Rosenberg S. A., Aebersold P., Cornetta K., Kasid A., Morgan R. A., Moen R., Karson E. M., Lotze M. T., Yang J. C., and Topalian S. L. (1990). Gene transfer into human's immunotherapy of

- patients with advanced melanoma, using tumour-infiltrating lymphocytes modified by retroviral gene transduction. *N. Engl. J. Med.*, 323, 570–578.
- Sarker S., Terron M. C., Khandokar Y. (2016). Structural insights into the assembly and regulation of distinct viral capsid complexes. *Nat. Commun.*, 7:13014.
- Rybicki E. P. (2020). Plant molecular farming of virus-like nanoparticles as vaccines and reagents. *Wiley Interdiscip. Rev.*12
- Schaack J. (2005). Adenovirus vectors deleted for genes essential for viral DNA replication. *Front Biosci.*, 10:1146-1155.
- Sessa M., Lorioli L., Fumagalli F. (2016). Lentiviral haemopoietic stem-cell gene therapy in early-onset metachromatic leukodystrophy: an ad-hoc analysis of a non-randomised, open-label, phase 1/2 trial. *Lancet.*, 388(10043):476-487.
- Sheyn D., Kimelman-Bleich N., Pelled G., Zilberman Y., Gazit D., Gazit Z. (2008). Ultrasound-based nonviral gene delivery induces bone formation in vivo. *Gene Ther.*, 15(4), 257-266.
- Sidhu S. S. (2000). Phage display in pharmaceutical biotechnology. *Curr. Opin. Biotechnol.*, 11(6):610-616.
- Sidhu S. S. (2001). Engineering M13 for phage display. *Biomol. Eng.*, 18 : 57 -63.
- Singh S., Khan I., Khim S. (2017). Safe and effective gene therapy for murine Wiskott-Aldrich syndrome using an insulated lentiviral vector. *Mol. Ther. Methods Clin. Dev.*, 4:1-16.
- Smeal S. W., Schmitt M. A., Pereira R. R., Prasad A, and Fisk J. D. (2017). Simulation of the M13 life cycle I: Assembly of a genetically structured deterministic chemical kinetic simulation. *Virology* 500 259–274
- Smith A. E. (1995). Viral vectors in gene therapy. *Annu Rev Microbiol.*, 49:807-838.

- Smith G.P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, 228(4705):1315-1317.
- Somia N., and Verma I. M. (2000). Gene Therapy: Trials and tribulations. *Nature Reviews*.
- Spandole S., Cimponeriu D., Berca L. M. (2015). Human anelloviruses: an update of molecular, epidemiological and clinical aspects. *Arch. Virol.*, 160:893–908.
- Sperinde J. J., Choi S. J., and Szoka F. C. (2001). Jr: Phage display selection of a peptide DNaseII inhibitor that enhances gene delivery. *J. Gene Med.*, 3(2):101-108.
- Stein U., Walther W., Stege A., Kaszubiak A., Fichtner I., and Lage H. (2008). Complete in vivo reversal of the multidrug resistance phenotype by jet-injection of anti-MDR1 short hairpin RNA-encoding plasmid DNA. *Mol. Ther.*, 16(1), 178-186.
- Suwan K., Yata T., Waramit S., Pryzstal J. M., Stoneham C. A., Bentayebi K., Asavarut P., Chongchai A., Pothachareon P., Lee K., Topanurak S., Smith T. L., Gelovani J. G., Sidman R. L., Pasqualini R., Arap W., and Hajitou A. (2019). Next-generation of targeted AAVP vectors for systemic transgene delivery against cancer. *PNAS*, Vol. 116 (37), 18571-18577
- Takahashi K., and Yamanaka S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126 : 663-676.
- Takahashi K., Iwasa Y., Hijikata M., and Mishiro S. (2000). Identification of a new human DNA virus (TTV-like mini virus, TLMV) intermediately related to TT virus and chicken anemia virus. *Arch. Virol.*, 145:979–993.
- Takahashi K., Tanabe K., Ohnuki M., Naritta M., Ichisaka T., Tomoda K. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131 : 861-872.

- Tariq H., Batool S., Asif S., Ali M., and Abbasi B. H. (2022). Virus-Like Particles: Revolutionary Platforms for Developing Vaccines Against Emerging Infectious Disease. *Frontiers in Microbiol.*, Vol. 12
- Thomas C. E., Ehrhardt A., and Kay M. A. (2003). progress and problems with the use of viral vectors for gene therapy. *Nature*. Vol. 4
- Titomirov A. V., Sukharev S., and Kistanova, E. (1991). In vivo electroporation and stable transformation of skin cells of newborn mice by plasmid DNA. *Biochim. Biophys. Acta*, 1088(1), 131-134.
- Travis J. (2015). Genetic engineering. Germline editing dominates DNA summit. *Science*., 350:1299–300. doi: 10.1126/science.350.6266.1299
- Turner M., Leslie S., Martin N. G., Peschanski M., Rao M., Taylor C. J. (2013). Toward the development of a global induced pluripotent stem cell library. *Cell Stem Cell* 13 : 382-384.
- Uchida M., Natsume H., Kobayashi D., Sugibayashi K., and Morimoto Y. (2002). Effects of particle size, helium gas pressure and microparticle dose on the plasma concentration of indomethacin after bombardment of indomethacin-loaded poly-L-lactic acid microspheres using a Helios (TM) gun system. *Biol. Pharm. Bull.*, 25(5), 690-693.
- Usmani S. S., Bedi G., Samuel J. S., Singh S., Kalra S., Kumar P., Ahuja A. A., Sharma M., Gautam A., and Raghava G. P. S. (2017). THPdb: Database of FDA-approved peptide and protein therapeutics. *PLoS ONE*. 12: e0181748. doi: 10.1371/journal.pone.0181748.
- Van V. K. M., Blouin V., Brument N., Agbandje-McKenna M., and Snyder R. O. (2008). The role of the adeno-associated virus capsid in gene transfer. *Methods Mol Biol.*, 437:51-91.

- Vanwezenbeek P., Hulsebos T. J. M., and Schoenmakers J. G. G. (1980). Nucleotide sequence of the filamentous bacteriophage M13-DNA genome-comparison with phage-Fd. *Gene* 11(1–2), 129–148.
- Verma I. M., and Weitzman M. D. (2005). GENE THERAPY: Twenty-First Century Medicine. *Annu. Rev. Biochem.*, 74:711–38 doi: 10.1146/annurev.biochem.74.050304.091637
- Vile R. G., and Russell S. J. (1995). Retroviruses as vectors. *Br. Med. Bull.*, 51(1):12-30.
- Virgin H. W., Wherry E. J., and Ahmed R. (2009). Redefining chronic viral infection. *Cell*, 138:30–50.
- Viswanathan P., Betapudi V., Kumar M. S., Rasheedi S., Vratil S., Bashyam M. D., and Hasnain S. E. (2003). The homologous region sequence (hr1) of *Autographa californica* multinucleocapsid polyhedrosis virus can enhance transcription from non-baculoviral promoters in mammalian cells. *J. Biol. Chem.*, 278:52564–71. doi: 10.1074/jbc.M309351200
- Vogt P. K. (2012). Retroviral oncogenes: a historical primer. *Nat Rev Cancer.*, 12(9):639-648.
- Walsh G. (2014). Biopharmaceutical benchmarks. *Nat Biotechnol.*, 32:992–1000. doi: 10.1038/nbt.3040
- Walther W., Siegel R., Kobelt D., Knosel T., Dietel M., Bembenek A., Aumann J., Schleef M., Baier R., Stein U., and Schlag P. M. (2008). Novel jet-injection technology for nonviral intratumoural gene transfer in patients with melanoma and breast cancer. *Clin. Cancer Res.*, 14(22), 7545-7553.
- Wang D., Tai P. W. L., and Gao G. (2019). Adeno-associated virus vector as a platform for gene therapy delivery. *Nat. Rev. Drug Discov.*, 18(5):358-378.
- Wang W., Li W., Ma N., and Steinhoff G. (2013). Non-viral gene delivery methods. *Curr. Pharm. Biotechnol.*, 14(1):46-60.

- Weiss R. A., and Tailor C. S. (1995). Retrovirus receptors. *Cell*. 82(4): 531-533
- Wendell D. M., Hemond B. D., Hogan N. C., Taberner A. J., and Hunter I. W. (2006). The effect of jet parameters on jet injection. *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 1, 5005-5008.
- Wilson J. M., Flotte T. R. (2020). Moving Forward After Two Deaths in a Gene Therapy Trial of Myotubular Myopathy. *Hum. Gene. Ther.*, 31:695-696.
- Wirth T., Parker N., and Yla-Herttuala S. (2013). History of gene therapy. *Gene.*, 525:162–9. doi: 10.1016/j.gene.2013.03.137
- Wold W. S., and Toth K. (2013). Adenovirus vectors for gene therapy, vaccination and cancer gene therapy. *Curr. Gene Ther.*, 13(6): 421-433.
- Wong S. (2022). Construction and characterization of a hybrid phage gene delivery platform. (Doctoral Thesis). University of Waterloo, Waterloo, Canada.
- Yacoby I., and Benhar I. (2008). Targeted filamentous bacteriophages as therapeutic agents. *Expert Opin. Drug Deliv.*, 5(3).
- Yan Z., Zak R., Zhang Y., and Engelhardt J. F. (2005). Inverted terminal repeat sequences are important for intermolecular recombination and circularization of adeno-associated virus genomes. *J. Virol.*, 79(1):364-379.
- Yang N. S., Burkholder J., Roberts B., Martinell B., and McCabe D. (1990). In vivo and in vitro gene transfer to mammalian somatic cells by particle bombardment. *Proc. Natl. Acad. Sci. USA*, 87(24), 9568-9572.
- Yang Q., and Catalano C. E. (2020). ATP serves as a nucleotide switch coupling the genome maturation and packaging motor complexes of a virus assembly machine. *Nucleic Acids Research*, Vol. 48, No. 9.

- Yin H., Kanasty R. L., Eltoukhy A. A., Vegas A. J., Dorkin J. R., and Anderson D. G. (2014). Non-viral vectors for gene-based therapy. *Nat. Rev. Genet.*, 15:541–55. doi: 10.1038/nrg3763
- Yoo S. Y., Jin H., Choi D. S., Kobayashi M., Farouz Y., Wang S., and Lee S. (2015). M13 Bacteriophage and Adeno-Associated Virus Hybrid for Novel Tissue Engineering Material with Gene Delivery Functions. *Advanced Healthcare Materials*, Vol. 1.
- Yoshida Y., and Yamanaka S. (2011). iPS cells: a source of cardiac regeneration. *J. Mol. Cell. Cardiol.*, 50 : 327-332.
- Young R., Wang I. N., and Roof W. D. (2000.) Phages will out: strategies of host cell lysis. *Trends Microbiol.*, 8(3),120–128.
- Yu J., Vodyanik M. A., Smuga-Otto K., Antosiewicz-Bourget J., Frane J. L., Tian S. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science*, 318: 1917-1920.
- Zhao N., Kamijo K., Fox P., Oda H., Morisaki T., Sato Y. (2019). A genetically encoded probe for imaging nascent and mature HA-tagged proteins in vivo. *Nat. Commun.*,10:2947
- Zheng H., Ye L., Fang X., Li B., Wang Y., Xiang X. (2007). Torque teno virus (SANBAN isolate) ORF2 protein suppresses NF-kappaB pathways via interaction with ikappaB kinases. *J. Virol.* 81, 11917–11924. doi: 10.1128/JVI. 01101-07.
- Zhou Y., Do D. C., Ishmael F. T., Squadrito M. L., Tang H. M., Tang H. L. (2017). Mannose receptor modulates macrophage polarization and allergic inflammation through miR-511-3p. *J. Allergy Clin. Immunol.* 141, 350–364.e8. doi: 10.1016/j.jaci.2017.04.049
- Zu H., and Gao D. (2021). Non-viral Vectors in Gene Therapy: Recent Development, Challenges, and Prospects. *The AAPS Journal*, 23:78 DOI: 10.1208/s12248-021-00608-7.

APPENDICES

Appendix A: TTV19 DNA sequences

atggcctggggctggtggcgccgctggcgagatggcctgcccgcggttgagaagaagacgtagacgtcgccccctgcgccgcagaaggg
ctggccgtcctgcgcgcaggtaccgcccagggcgtaccgtgcgcacaagacgccgtagggtggggaaggcgctgtaccgtcgccggtggcgt
aggcgtacctacgtgcgtaagggccgcccaccgtaagaagaagaagcgactgacctgcgccagtgaggcagcctgcgacaaggcgctgctgtac
catcaccgggttacctccctatcgtcttctgcgggtcacacaaaggtaacaagaactacgccctgcattctgacgactacacgccccagggccaacc
cttcgggtggagctctctgaccacctccttctccctgaaggttctctcgaccaacaccagaggggcctgaacaagtggctccttcctaaccgaccaa
cttgacttggctcgttaccgcggtgcaaattcatttctaccgcaccaagcagaccgactggatcgacagtagacatcagcgaaccatacaag
ctggacaagtactcatgccccaaactaccacctggcaatctgatcaaggctaagcacaagttcctgatcccttcatacgaaccaaccaagggg
aagacagaagatcatcgtgaagatcccaccgcctgacctgttcgtggacaaaatggtacacacagaggaggacttgtgctcggtgaacctggtgtcatt
ggccgtgtccgcccctccttctgcacccttcggctccccacagaccgacaacccctgttacaccttcagggtgctgaaggagttctactacca
ggctatcggttcatgtctaccgacgagtcggcaactacgtgttcaacgtcctctacgaggagaactcatactgggaatcgaacatcactcccttc
tacgtgatcaacgttaagaagggttagcaataccagggactacatgtcaccaaagatctccgactcacacttccgcaacaagggtcaacaccaacta
caactggtacacatacaacgctaagagccacaagaacgacttgacgaactcaggcgctgcctacttcaagcagttgactaccgagggccctcag
cagacctctagtgaaggggtacgcctcccagtggaacaacccctaccactgacgcttacgagtagtatcacctgggcatgttcttaccatcttctcgc
ccccgaccgcccgtgctcgccttcccttgcgcctaccaggacgtaacctacaacccactgatggacaagggcgctcggaaccacgtgtggtt
ccaatacaacacaaagggtgatacacagttgatcgtcaccgggtggttagttgcaaggctcacatcgaggatattcctctgtggggccgcatttctacgg
ctactccgatttcatcagtcaggctgggccccttctgtggacgctgagaccgtgggctgatctgcgtgatctgcccctacacaaagcctcccat
gtacaacaagaccaacccctatgatgggatacgtgttctacgaccgcaacttcgggtgacggaaagtggattgacggacgtggaaaaatcgagcca
tactggcaggtgcgctggcgctcccgaaatgctgttccaggagaccgtgatggctgacatcgtacagacaggtcccttcttacaaggacgagct
gaagaactcaacgctgggtggccaagtacaagttctacttcacatggggagggaacatgatgttccagcagactatcaagaacccctgcaagacg
gacggctgacccacagactccgatagacacccctcgcggtatccagggtggctgacccagaacagatgggcccagatgggtgttccactcttctc
actggcgccgtggttacctgtctgaaggtgccatcaagcgtctgcacgagaagccttagactacgaatcttacttcacccagccaaagagaccac
gcatcttcccaccgaccgaagctgcggaggggagagttccgtgagcccgagaagggttcttactccgaggaagaacgctcccaagcttcagctg
aagagcaggctaccgaggagactgttcttctgctcaagagaagactgagagagcagagaaaagttggaacaacagctgcagttcctcactcgcga
aatgttcaagaccagggccgactgcacatcaaccctatgctcctgagccagagataagaaaacctgtatttctcagtcaccacatcaccatcaccat

FIG A1. TTV 19 codon optimized capsid protein sequence

tcggagccgcgcgtacgccttcgggtgcgcgcggcacctcagacccccgctcgtgctgacacgctcgcgcgtgtcagaccacttc
gggctcgcgggggtcgggaattttgctaaacagactccgaggtgctcttggaactgagtgggcatatcagtaacgaaagtgagtgg
ggccagacttcgcataaggcctttatcttccattggatagtggaagaggggttgcataacgcggacccggttttacccttcgg
actacaaaatggccgttccgtgacgtcacggccgcattttaagtaaggcggaagcagctcggcatatacaaaatgtcggcggagc
acttcgggttacccaaaatgaaggccaatatcttcgggtcataggtcacacctacgtcacaagtcacgtggggagggttggcgata
gcccggagtcaatcctccacgtggcctgtcacgtgacgtgtacgtcacagtcgccattttctttacaaaatggccgacttcctcctt
ttttaaaaataacggctcggcggcggcgcgcgcgtgcgcgcgcgcgcgggggggctgccgcccccccggggaaagggggg
ggcccccccggggggggggtttgcccccgcccccccccg

FIG A2. TTV 19 intergenic sequence containing DNA binding signal.

Appendix B: Media and buffer compositions

TABLE B1. Composition of media and buffers used for experiments.

Media/Buffer Name	Composition
LB Broth	12.5 g of LB (containing tryptone, yeast extract, and NaCl) in 500 mL of MilliQ water
LB Agar	12.5 g of LB (containing tryptone, yeast extract, and NaCl), 6.5 g of agar in 500 mL of MilliQ water
1X Wash Buffer (TBS-T)	20 mM Tris HCl (pH = 7.4), 150 mM NaCl, and 0.1% (w/v) Tween 20 in 1L of MilliQ water
Blocking Buffer (5% Skim milk)	5 g of skim milk in 100 mL of TBS-T
10X Running Buffer	30 g Tris base, 144 g glycine, and 10 g SDS in 1L of MilliQ water

LB: Luria Burtani; SDS: sodium dodecyl sulfate; TBS-T: tris-buffered saline/tween

Appendix C: Supplementary qPCR Data

Amplification Plot (Rn vs. Cycle)

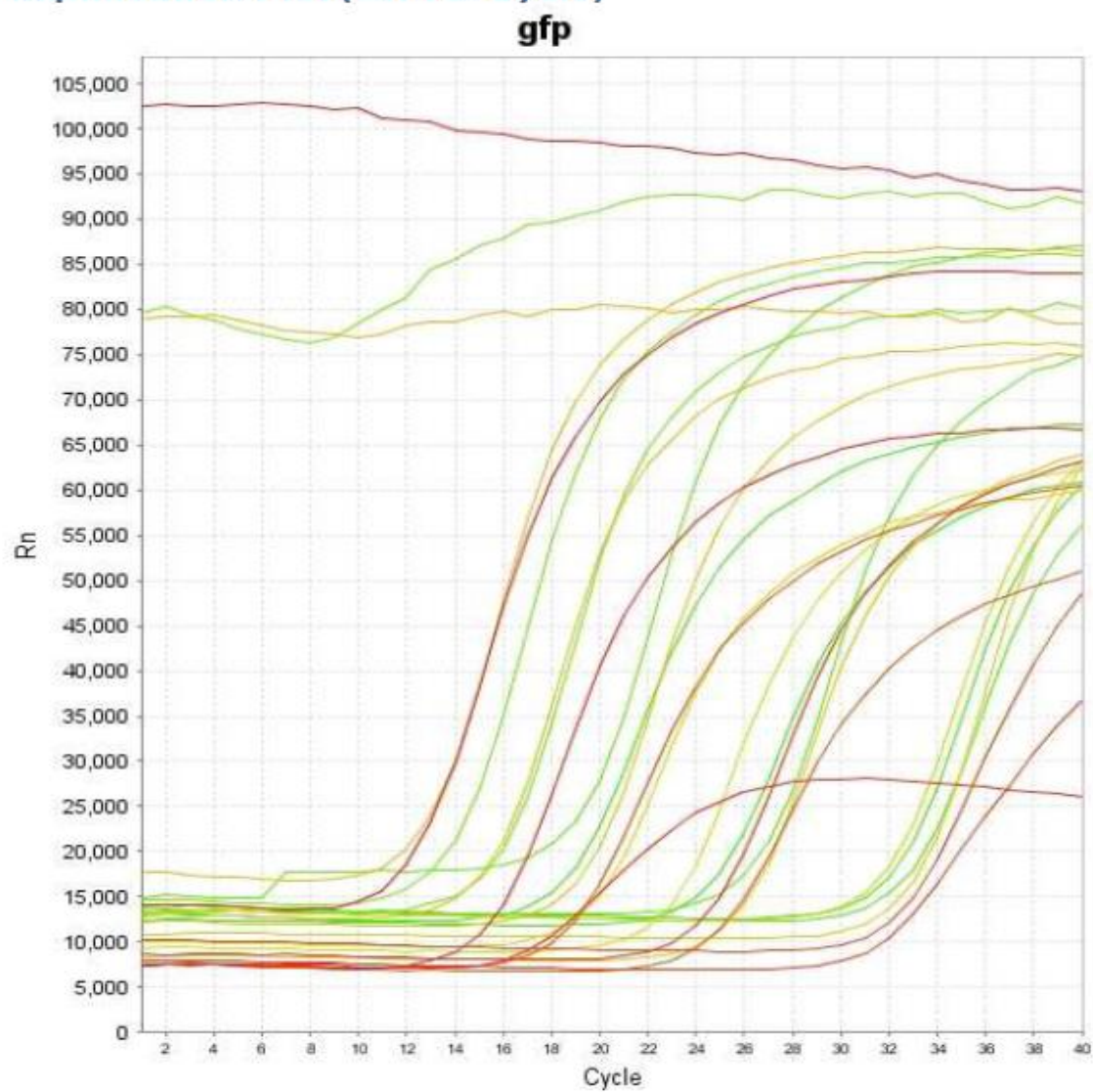


FIG C1: Amplification Plot for *gfp* amplicon.

Appendix D: TTV19 ORF1 protein cleavage site

TABLE D1: List of enzymes capable to cleave TTV19 capsid protein.

Name of enzyme	No. of cleavage sites	Positions of cleavage sites
Arg-C proteinase	78	24 25 27 28 32 33 35 36 37 38 39 40 41 44 45 46 47 50 53 54 56 57 58 59 62 64 65 66 67 70 71 72 74 75 78 79 80 84 87 89 94 98 105 106 107 162 178 180 188 236 238 312 358 383 384 432 436 539 550 560 562 621 627 630 643 652 653 663 680 682 695 707 726 727 729 732 743 762
Asp-N endopeptidase	36	132 133 157 171 174 192 198 206 231 248 252 259 285 308 345 353 377 413 430 442 450 466 482 494 505 537 542 547 573 584 618 623 625 635 649 669
Asp-N endopeptidase + N-terminal Glu	75	9 132 133 157 171 174 192 198 201 206 231 248 252 258 259 285 296 308 309 320 321 326 345 353 377 380 392 400 413 416 430 442 450 466 481 482 494 497 499 505 507 537 542 547 553 563 568 573 584 585 618 623 625 635 637 649 657 665 669 671 687 690 692 695 697 703 704 705 712 713 717 718 729 734 743
BNPS-Skatole	23	20 22 23 26 29 34 68 77 100 167 194 255 326 367 408 459 487 546 557 561 602 644 651
CNBr	14	18 348 422 450 524 531 532 565 572 606 607 640 745 757
Chymotrypsin-high specificity	106	13 14 20 22 23 26 34 55 68 73 77 82 100 113 118 128 135 142 151 157 167 179 184 186 187 194 198 204 209 214 226 231 251 255 256 276 280 290 292 298 299 300 305 314 316 320 325 326 333 334 347 357 365 367 368 370 386 387 404 408 416 418 423 427 441 446 459 460 462 487 490 491 493 496 504 519 525 534 536 537 541 546 556 557 561 567 581 583 596 598 599 600 602 608 644 646 649 651 655 671 674 675 694 702 740 746
Chymotrypsin-low specificity	186	1 2 3 4 5 6 12 13 14 18 20 22 23 26 34 43 55 68 73 77 82 88 95 97 100 113 118 121 128 130 131 135 142 146 151 153 156 157 160 164 167 174 176 179 184 186 187 194 198 204 206 209 214 219 224 226 227 231 250 251 255 256 261 266 269 276 277 280 290 292 295 298 299 300 305 314 316 319 320 325 326 333 334 347 348 356 357 365 367 368 370 375 379 380 382 386 387 390 404 408 416 418 419 420 422 423 427 428 441 446 449 450 457 459 460 462 470 480 486 487 490 491 493 496 501 504 512 519 525 531 532 534 536 537 541 546 556 557 561 565 566 567 572 581 583 587 592 596 598 599 600 602 606 607 608 640 644 646 647 649 651 655 656 664 665 669 671 674 675 694 702 722 723 724 728 734 738 740 741 745 746 752 753 757 758 759
Clostripain	78	24 25 27 28 32 33 35 36 37 38 39 40 41 44 45 46 47 50 53 54 56 57 58 59 62 64 65 66 67 70 71 72 74 75 78 79 80 84 87 89 94 98 105 106 107 162 178 180 188 236 238 312 358 383 384 432 436 539 550 560 562 621 627 630 643 652 653 663 680 682 695 707 726 727 729 732 743 762
Factor Xa	2	550 621

Formic acid	36	133 134 158 172 175 193 199 207 232 249 253 260 286 309 346 354 378 414 431 443 451 467 483 495 506 538 543 548 574 585 619 624 626 636 650 670
Glutamyl endopeptidase	39	10 202 259 297 310 321 322 327 381 393 401 417 482 498 500 508 554 564 569 586 638 658 666 672 688 691 693 696 698 704 705 706 713 714 718 719 730 735 744
Iodosobenzoic acid	23	20 22 23 26 29 34 68 77 100 167 194 255 326 367 408 459 487 546 557 561 602 644 651
LysC	48	85 90 91 92 93 123 126 154 166 183 190 205 208 221 223 225 240 244 254 296 339 340 351 360 373 376 388 402 452 465 478 521 527 545 552 584 588 595 597 613 617 662 667 679 699 725 733 747
LysN	48	84 89 90 91 92 122 125 153 165 182 189 204 207 220 222 224 239 243 253 295 338 339 350 359 372 375 387 401 451 464 477 520 526 544 551 583 587 594 596 612 616 661 666 678 698 724 732 746
NTCB (2-nitro-5- thiocyanobenzoic acid)	11	107 118 181 210 261 288 438 476 513 516 615
Pepsin (pH1.3)	118	11 12 13 14 97 114 117 118 129 130 141 145 146 150 151 152 153 155 157 163 169 173 174 175 176 183 184 186 205 206 218 219 250 251 260 261 265 266 268 269 275 276 279 291 292 294 295 297 304 305 315 316 318 319 332 356 357 379 381 387 389 419 420 422 423 426 427 437 448 460 469 470 485 489 490 495 496 500 503 511 512 535 536 540 565 566 567 580 587 591 592 598 600 607 608 646 648 656 668 674 675 693 694 721 722 723 724 733 737 738 739 740 741 746 751 752 758 759
Pepsin (pH>2)	197	11 12 13 14 20 21 22 23 97 99 112 114 117 118 127 129 130 134 141 145 146 150 151 152 153 155 157 163 166 167 169 173 174 175 176 179 183 184 186 187 193 194 197 198 203 205 206 208 209 218 219 231 250 251 254 255 260 261 265 266 268 269 275 276 279 290 291 292 294 295 297 299 300 304 305 313 315 316 318 319 320 324 325 326 332 334 346 347 356 357 364 365 366 367 368 369 370 379 381 387 389 403 407 408 415 416 417 418 419 420 422 423 426 427 437 440 441 445 448 458 460 461 462 469 470 485 487 489 490 491 492 493 495 496 500 503 511 512 518 525 533 534 535 536 537 540 545 546 555 557 565 566 567 580 582 583 587 591 592 595 596 598 600 601 602 607 608 644 646 648 650 651 656 668 670 671 673 674 675 693 694 702 721 722 723 724 733 737 738 739 740 741 746 751 752 758 759
Proline- endopeptidase	10	42 51 216 279 433 563 622 629 668 681
Proteinase K	342	10 12 13 14 19 20 22 23 26 29 31 34 43 48 52 55 60 61 63 68 73 77 81 82 83 95 96 97 100 103 104 109 110 111 113 114 116 117 118 122 128 129 130 135 136 142 145 146 148 149 151 153 155 156 157 164 167 169 174 176 177 179 184 185 186 187 189 192 194 195 198 200 202 204 206 209 214 219 220 222 226 227 228 231 233 241 242 243 245 250 251 252 255 256 257 259 261 264 266 267 269 270 271 273 274 276 277 280 285 290 291 292 294 295 297 298 299 300 302 303 305 307 308 310 314 315 316 318 319 320 321 322 325 326 327 330 331 333 334 335 336 338 344 347 352 357 361 363 365 367 368 369

		370 372 379 381 382 385 386 387 390 391 392 393 398 401 404 405 408 409 410 412 413 415 416 417 418 420 423 425 426 427 428 429 434 437 440 441 444 445 446 449 454 458 459 460 462 464 466 468 470 471 472 473 479 481 482 484 486 487 488 489 490 491 493 496 497 498 500 501 504 505 507 508 509 510 512 513 515 516 519 520 525 528 534 535 536 537 541 546 547 553 554 556 557 559 561 564 566 567 569 570 571 573 575 576 578 581 583 586 587 591 592 593 594 596 598 599 600 601 602 608 611 612 618 623 632 634 635 638 644 645 646 649 651 655 656 658 660 661 664 666 669 671 672 674 675 676 683 684 687 688 689 690 691 693 694 696 698 702 704 705 706 710 712 713 714 716 717 718 719 720 721 722 723 724 728 730 734 735 738 740 741 742 744 746 748 750 752 754 758 759
Staphylococcal peptidase I	34	10 202 259 297 310 321 327 381 393 401 417 482 498 500 508 554 564 569 586 638 658 666 672 688 691 693 696 698 704 713 718 730 735 744
Thermolysin	171	11 13 17 18 30 42 47 51 60 82 94 95 96 102 109 115 116 117 128 129 141 144 145 150 152 154 155 156 163 173 176 183 184 185 194 205 218 219 221 225 226 240 241 242 250 251 263 265 266 268 269 270 272 273 275 276 279 291 293 294 301 302 304 306 314 315 317 318 329 332 334 335 337 347 351 356 360 371 384 386 389 404 419 421 422 425 426 427 439 448 449 453 457 459 465 469 470 471 478 480 485 487 488 489 496 503 504 509 511 512 514 515 523 530 531 534 535 540 546 552 558 565 566 570 571 572 575 580 591 592 593 597 599 605 606 607 611 631 633 634 639 644 645 648 655 659 660 663 668 674 682 689 709 711 715 720 721 722 723 727 733 737 739 740 745 749 751 753 756 757 758
Thrombin	1	550
Tobacco virus protease	etch (TEV) 1	15
Trypsin	104	24 25 27 28 32 33 35 44 47 53 54 56 59 62 64 67 70 72 74 75 78 80 84 85 87 89 90 91 92 93 94 98 105 107 123 126 154 162 166 178 180 183 188 190 205 208 221 223 225 236 238 240 244 254 296 312 339 340 351 358 360 373 376 383 384 388 402 436 452 465 478 527 539 545 550 552 560 584 588 595 597 613 617 627 630 643 652 653 662 663 679 682 695 699 707 725 726 727 729 732 733 743 747 762