The WDR11 protein as a novel host factor for AAV transduction

A thesis submitted to

The Indian Institute of Science Education and Research Pune in partial fulfilment of the requirements for the BS-MS Dual Degree Programme

by

Shivani Nachiket Deshpande

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

Pashan, Pune 411008, India

Date: 1st April 2023

Under the guidance of

Supervisor: Dr. Chuck Bailey, Prof. John Rasko

Gene and Stem Cell Therapy Program, Centenary Institute, University of Sydney

From May 2022 to March 2023

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

Certificate

This is to certify that this dissertation entitled “The WDR11 protein as a novel host factor for AAV transduction” towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research represents study/work carried out by Shivani Nachiket Deshpande at the Gene and Stem Cell Therapy Program at the Centenary Institute, University of Sydney under the supervision of Prof. John Rasko and Dr. Charles Bailey (Gene and Stem Cell Therapy Program), during the academic year 2022-2023.

Prof. John Rasko
Head, GSCT
Centenary Institute, University of Sydney

Dr. Charles Bailey
Head, Cancer and Gene Regulation, GSCT
Centenary Institute, University of Sydney

Dr. Girish Ratnaparkhi
Professor and Dean, Academics
Department of Biology, IISER Pune

Committee:

Supervisor: Prof. John Rasko, Gene and Stem Cell Therapy Program, Centenary Institute

Co-supervisors: Dr. Charles Bailey, Dr. Bijay Dhungel, Gene and Stem Cell Therapy Program, Centenary Institute

TAC Member: Dr. Girish Ratnaparkhi, Department of Biology, IISER Pune

Declaration

I hereby declare that the matter embodied in the report entitled “The WDR11 protein as a novel host factor for AAV transduction” are the results of the work carried out by me under the Gene and Stem Cell Therapy Program at the Centenary Institute, University of Sydney under the supervision of Prof. John Rasko and Dr. Charles Bailey. This work has not been submitted elsewhere for any other degree.

Shivani Nachiket Deshpande

Date: 1st April 2023

Table of Contents

List of tables

List of figures

Abstract

Gene therapy has emerged in the last decade as a widely studied and implemented group of unique therapeutic modalities. Though still young and investigational in nature, it has generated significant traction, with many translational technologies making their way to the clinic as licensed treatments for genetic diseases. For *in vivo* gene therapies, adeno-associated virus (AAV) is among the most widely studied and promising viral vectors, due to its low immunogenicity and diverse tissue tropism. AAV transduction and hence the efficiency of genetic transfer is influenced by host trafficking and processing machinery, through proteins defined as host factors. This study aims at understanding the role played by one such host factor, the WD-repeat family protein WDR11 in AAV transduction. WDR11 was knocked out in different cell lines using CRISPR/Cas9 and these knockout cells were studied for differences in AAV transduction. Genomic knockout of WDR11 increased AAV transduction by 1.5 fold in HeLa cells, and 3-3.5 fold in eHAP cells, an effect that was reversed by ectopic overexpression of WDR11 in knockout cells. This study also showed that WDR11 KO increases the cell surface expression of AAVR, a receptor shown to be critical for transduction by multiple AAV serotypes. It is hence possible that WDR11 KO causes leads to increased localisation of AAVR at the cell surface, and together assisted by compensatory pathways leads to increased AAV trafficking in the cell. Overall, this study aims to study WDR11 as a potential restriction factor to gain a deeper understanding of AAV biology.

Acknowledgments

As my 9 months in Sydney and 5 years at IISER Pune are coming to an end, I have a feeling words will not be enough to thank and express my feelings for everyone who has gotten me over the finish line. They have loved me, taught me, and nurtured my passion and love for science, and they deserve much more than I can ever convey in this small section.

To my Sydney family, Aniruddha Kaka, Sushma Kaku, Ananya, and Bruno. You have never let me miss home, and I will remain forever grateful for all the love you have given me. To Ramah, I will always consider it so lucky that we met again in this beautiful city, and I had a wonderful friend to explore it with.

To Chuck, your experience and kindness are incomparable. Thank you for your constant guidance and feedback, and for teaching me to go where the science led me. To Bijay, I hope any lab you're in continues to have your wild, unencumbered, chaotic spirit. Your brilliance, your friendship, and your open door make you a great mentor, and I hope we get to catch up over a pizza in New York soon. To John, thank you for always asking the tough questions and for taking a chance on me, Sydney changed my life. To Habib and Oliver, who had their own theses and projects to work on, but always took the time to talk to me, help me with my experiments, and have stimulating discussions over tea with me. I will miss your company. To Sarah, my murder-mystery soul sister, for never letting me feel like I was alone when my experiments failed. To Chirag, thank you for always making sure I was okay away from home. For taking the time to be my friend, even after having your hands full with Ruhani, and for being there to remind me that India is always with me wherever I go. To Mehdi, for always lending me his positive spirit and endless optimism, even when things looked bleak.

To the RAs that made GSCT my family. Cynthia, there is no one like you. In spirit, my sister, mother and friend. You're brilliant, beautiful, and kind. Every day at GT felt like home because of you, and I mean it when I say, you might just know everything there is to know on the planet about how to make experiments work. To Julie, thank you for teaching me to be disciplined and patiently showing me how to do the littlest things. You are a force to be reckoned with, and I loved every stimulating conversation we had over tea. To Rajini, thank you for teaching me how to do beautiful western blots. For telling me it's okay when experiments fail, and that I'm doing enough. You inspire

me to work harder and do better. To Althea, I loved having someone fun to bond over Old Monk and cats! I hope we get to do both when you visit someday. It was a privilege to be a part of GSCT, and their warmth and guidance will always stay with me.

To Kundan, my first lab mentor. You are remarkable, selfless, and brilliant. Thank you for teaching me every bit of molecular biology I learnt at IISER, and for always putting your students before yourself. To Girish Sir, thank you for taking a chance on a strange second-year student before anyone else did. Deadpan humor and honesty make a great scientist.

To my friends Mihir, Jatin, Adithyan, Chinmay, Shruthi, Akanksha, Aishwarya and fellow Kangaroo Suryo who spent the last 9 months being scattered all across India and the world, we did it. You've booked permanent spots in my life and that somehow means permanent spots in this document that marks the end of our amazing journey at IISER.

To my boyfriend Siddharth, you are my rock and my cheerleader. Even from the other end of the world, you make me laugh, take my stress away, and make me feel loved. We did it, and I'm so honoured and proud to graduate with you.

Lastly, to my parents Nachiket and Nisha, and my sister Jui. You have cheered me on, lifted me up, and given me every bit of love and opportunity since the day I was born. I have the most wonderful parents in the world, and the best friend in a sister anyone could ever ask for. I was so lucky to be so close to them, and I have missed them terribly every single day I have been away from home.

Contributions

Contributor name	Contributor role
Prof. John Rasko, Dr. Charles Bailey, Dr. Bijay Dhungel	Conceptualization Ideas
Dr. Charles Bailey, Dr. Bijay Dhungel, Shivani Deshpande	Methodology
-	Software
Shivani Deshpande	Validation
Shivani Deshpande	Formal analysis
Shivani Deshpande	Investigation
Gene and Stem Cell Therapy Program, Centenary Institute	Resources
-	Data Curation
Shivani Deshpande	Writing - original draft preparation
Dr. Charles Bailey	Writing - review and editing
Shivani Deshpande	Visualization
Dr. Charles Bailey, Dr. Bijay Dhungel	Supervision
Dr. Charles Bailey, Prof. John Rasko	Project administration
Dr. Charles Bailey, Prof. John Rasko	Funding acquisition

1. Introduction and literature review

1.1 Gene therapy

The human body is an incredibly complex entity encoded from a genetic blueprint. Our DNA carries vast amounts of critical information which direct the correct development and differentiation that is required. Dysregulation in the reading of this information or defects in the base sequence can cause various genetic diseases. These diseases typically have poor prognosis and few, if any, treatment options. For example, cystic fibrosis is a monogenic rare disease caused by possession of two recessive copies of the *CFTR* (cystic fibrosis transmembrane conductance regulator) gene (Chen *et al.*, 2021). Simple mutations in a single gene cause a serious life-shortening disease. On the other hand, diseases with genetic causes also include cancer- which is caused by uncontrolled cellular proliferation associated with mutations in as many as 550 different genes and genomic instability (Martínez-Jiménez *et al.*, 2020).

The concept of gene therapy started with the first effective genetic modification of mammalian cells using DNA isolated from other healthy cells performed by Waclaw and Elizabeth Szybalski in 1962 (Dulak, 2021). The basic definition of gene therapy would then be treatment of a disease caused by defective genes by simply supplying a correct copy of those genes.

In subsequent years, the discovery of techniques like calcium phosphate, liposomes, and dendrimers for transfection of plasmid DNA into cells, and advances in techniques like recombinant DNA cloning which allowed ready manipulation of nucleic acids progressed the field to allow for delivery of specific genes to mammalian cells (Dulak, 2021). Structural and biological studies on viruses allowed researchers to modify these viruses and use them as biological vehicles, which gave us an efficient way to genetically modify mammalian cells *in vivo* (Wirth *et al.*, 2013).

These discoveries were rapidly applied to patient treatments in the early 90s, but the outcomes were disappointing with trials either not showing therapeutic benefits, or incidents like the death of Jesse Gelsinger, a clinical trial patient being treated for ornithine transcarbamylase deficiency who experienced multi-organ failure due to an inflammatory reaction to a therapeutic adenovirus vector (Wirth *et al.*, 2013). In 1996,

a National Institutes of Health (NIH) Board recommended that taking these therapeutic technologies to the clinic was premature, as researchers had limited knowledge of the nature and interactions of the viral vectors being used (Kimmelman, 2003). Over the next decade, there was a systematic reinvention of the study and development of these technologies by focusing on understanding vector biology and target cell-vector interactions so that safer gene delivery techniques could be developed to effectively target tissues (Alhakamy *et al.*, 2021).

In a much broader sense, genetic therapies now encompass applications including CAR-T cell therapies (modification of T-lymphocytes to improve selective killing of cancer cells) (Aghajanian *et al.*, 2022), siRNA (short interfering RNA)/antisense RNA-based therapeutics using exon skipping or inclusion mechanisms (Winkle *et al.*, 2021), *in vivo* correction of genetic defects by the targeted delivery of gene editing components like the CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats) system to diseased cells, and mRNA-based vaccines for infectious diseases. Alzheimer's disease is a neurodegenerative disorder involving accumulation and aggregation of the amyloid beta protein. In a recent approach by Park et al. (Park *et al.*, 2019), delivery of CRISPR/Cas9 ribonucleoprotein complexes to diseased neurons has been tested using amphiphilic peptides so that this defect can be corrected at the genetic level.

The gene therapy field has seen development in massive strides, as is reflected in the large number of nucleic acid therapies and clinical trials currently in use or in development (Table 1.1,1.2).

The term gene therapy applies to both germline and somatic therapy. However, therapeutic alterations to the genome at the germline level that can be passed along to future generations come with significant ethical concerns. Somatic genetic alterations are restricted to the target patient and hence are the only form of gene therapy approved (Sayed *et al.*, 2022).

1.1.1 Genome editing

Genome editing refers to the addition, removal, or modification of target DNA in a living organism. Zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALEN), and CRISPR/Cas9 are the dominant gene editing technologies currently in use.

1.1.1.1 ZFNs and TALEN

ZFNs have a DNA-binding domain and a *FokI* nuclease in a DNA cleavage domain. Each ZF will bind to a specific codon in the target sequence, so ZFNs have a characteristic binding sequence of 18-36 nucleotides, which are then cleaved by *FokI* (Alhakamy *et al.*, 2021).

TALEN repeat domains are around 35 amino-acid sequences originally identified in *Xanthomonas* bacteria. These repeats can be engineered into multimeric arrays to recognize specific DNA sequences, and then fused with non-specific nuclease domains to create precise molecular 'scissors', which can edit genes in yeast, plants, rats, and human somatic and pluripotent cells (Reyon *et al.*, 2012). However, targeting specific genes using ZFNs and TALENs is time consuming and costly due to the specific requirements of construct design (Cox *et al.*, 2015).

1.1.1.2 CRISPR/Cas9

The CRISPR-Cas9 system is a bacterial RNA-mediated adaptive defense system that has been modified to create precise genomic edits in DNA. Cas9 is an RNA-guided DNA endonuclease that is coupled with a chimeric RNA construct called an sgRNA (single guide RNA). The sgRNA guides the Cas9 nuclease to a precise genomic location through complimentary base pairing and mediates cutting of the DNA next to a specific nucleic acid motif called the protospacer adjacent motif (PAM) (Jinek *et al.*, 2012). This system can be used to generate targeted edits at the genetic level in mammalian cells by creating double-stranded breaks (DSBs) in genomic DNA (Ran *et al.*, 2013).

DSBs are repaired by two major pathways: non-homologous end joining (NHEJ) or homology-directed repair (HDR). NHEJ can lead to error prone insertion or deletions (indels) at the site of the break (defective repair) (Hoeijmakers, 2001). Using double-

stranded breaks, NHEJ can be activated and exploited for a therapeutic benefit by introducing indels into genes with deleterious gain-of-function mutations or dominant-negative effects (Cring and Sheffield, 2022). For example, a study utilized NHEJ DNA repair to knockdown the expression of a mutant version of a gene (*myocilin)* and showed increased protection from glaucoma in mice (Jain *et al.*, 2017). In cases where genetic haploinsufficiency or defective protein production leads to a disease phenotype, the HDR pathway can be exploited to introduce the desired therapeutic edits into the target gene. Recently, there has also been a significant number of studies that use the deactivated version of the Cas9 nuclease coupled to either transcriptional activators or repressors to modulate gene activity using the targeting function of the RNA guide (Vojta *et al.*, 2016).

Gene therapy can be broadly categorized into two approaches- where the therapeutic nucleic acid is introduced into the patient directly *in vivo,* or the *ex vivo* approach wherein patient-sourced stem or progenitor cells are genetically modified in a closed system ('cell factories') and re-introduced into the patient (Mallik *et al.*, 2022).

1.1.2 *Ex vivo* gene therapy

1.1.2.1 CAR-T (Chimeric antigen receptor- T cell) therapy

CAR-T cell therapy involves the genetic modification of autologous immune (killer T) cells to treat aggressive and treatment-resistant forms of cancer. CARs are engineered receptor-like molecules that reprogram T cells to target specific chosen antigens (Rivière and Sadelain, 2017). A variety of methods are used to introduce CAR constructs into T cells, including non-viral based DNA transfection, gamma-retroviruses, and lentiviruses (Maus *et al.*, 2014). CAR-T cell therapies have shown great promise in clinical trials and use, highlighted by the number of FDA-approved therapies using these constructs (Table 1.1).

1.1.2.2 HSC (Hematopoietic stem cell) editing

Hematopoietic stem cells are self-renewing cells that give rise to all cell lineages in the blood. In this kind of *ex vivo* therapy, as progenitor cells undergo cell division and differentiation, the therapeutic gene has to be stably integrated into the genome to be expressed long-term by all daughter cells (Anguela and High, 2019). For example, this approach has been explored with respect to ADA-SCID (adenosine deaminase-severe

combined immunodeficiency) where patient-derived HSCs corrected with a retroviral construct supplying a functional copy of ADA leading to the generation of multiple cell lineages showing gene correction and a restored immune system (Kumar *et al.*, 2016).

1.1.3 *In vivo* gene therapy

Vectors used for gene therapy are characterized as viral or non-viral vectors (Fig. 1.1). The most prominent barriers to delivering genes to specific tissues are poor efficiency of gene transfer, unwanted immune reactions and toxicity (Alhakamy *et al.*, 2021). Non-viral vectors are widely regarded as safer, simpler and more cost effective than viral vectors, but they lack the specificity and efficiency for use in the clinic. On the other hand, viral vectors carry risks of toxicity and immunogenicity as a cost of the increased targeting efficiency (Ibraheem *et al.*, 2014).

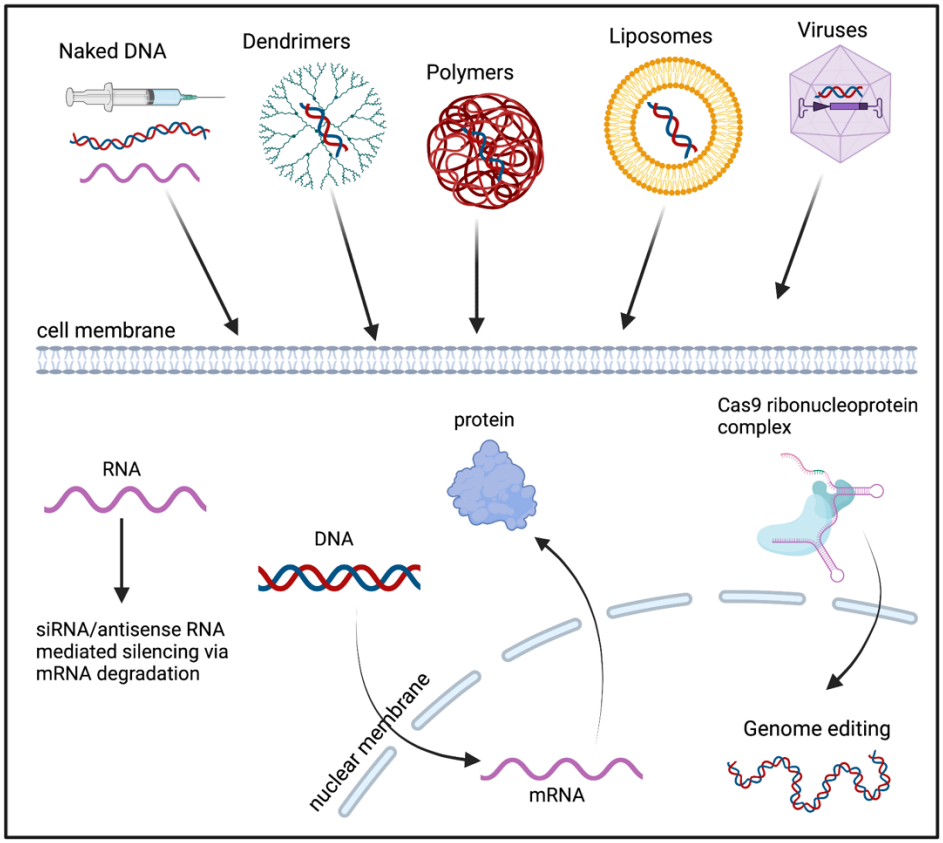

Figure 1.1: Various modes of in-vivo therapeutic gene delivery. Studied and adapted vehicles include both non-viral and viral vectors carrying DNA, RNA as well as genome editing ribonucleoprotein complexes.

1.1.3.1 Viral vectors

Viral vectors are the most commonly used tools for gene therapy (Dulak, 2021). Retroviruses, lentiviruses, adenoviruses and adeno-associated viruses (AAV) have been most widely adapted as gene delivery vehicles. Considerations while choosing a viral vector include the size of the gene, efficiency of expression, duration of treatment, immunogenicity, scalability, integration into host DNA and risk of infection (Alhakamy *et al.*, 2021). Adenoviruses and retroviruses have a higher carrying capacity (7.5 kB and 8 kB respectively) as compared to AAV with 5 kB. Adenoviruses and AAVs do not integrate into the genome, but typically remain as episomes. Retroviruses integrate into the genome, which confers the advantage of permanent induction of therapeutic gene expression (Cring and Sheffield, 2022). Adenoviruses have also been shown to elicit strong immune responses due to pre-existing natural adenovirus immunity in humans (pre-existing anti-adenoviral neutralizing antibodies in humans dampen the efficacy of the vector). Lentiviruses are also able to integrate into the host genome and transduce both dividing and non-dividing cells but carry a high risk of insertional mutagenesis (Bulcha *et al.*, 2021).

Due to its wide tropism, low immunogenicity, low integration frequency, episomal maintenance and long-term stable gene expression, AAVs are the preferred mode for *in vivo* viral gene transfer, as seen by the overwhelming percentage (~50%) of currently active clinical gene therapy trials that use AAV as the gene delivery vector (Zhao *et al.*, 2022). The role of AAV as a gene delivery vector is expanded on in subsequent sections.

1.1.3.2 Non-viral vectors

Non-viral vectors confer several advantages over viral vectors. They are less cytotoxic, less immunogenic due to being non-biological entities, and carry no risk of insertional mutagenesis. However, in comparison to viral vectors, these molecules offer less specificity, and carry the risk of being degraded on their journey to the target tissue (Mohammadinejad *et al.*, 2020). Non-viral systems currently in development include naked DNA injection with physical methods like gene gun, electroporation, magnetofection and hydrodynamic delivery, and vehicles for *in-vivo* delivery like lipids and liposomes, polymers, peptides, polysaccharides and inorganic nanoparticles (Yin *et al.*, 2014). For example, cationic liposomes encapsulating siRNA against the

CYP1A1 (cytochrome P450) gene (which has been implicated in enhanced tumor growth) have been successfully shown to silence *CYP1A1* in mice (Mohammadinejad *et al.*, 2020).

Table 1 and 2: List of approved *ex vivo* (1) and *in vivo* (2) genetic/nucleic acid therapies collated as of December 2022 (Alhakamy *et al.*, 2021; Dulak, 2021; Arabi *et al.*, 2022; Mallik *et al.*, 2022)

Abbreviations: EMA: European Medicines Agency, FDA- American Food and Drug Administration, SFDA- State Food and Drug Administration, China

1.1.4 Table 1: List of approved *ex-vivo* gene therapy treatments

Trade name (general name)	**Manufacturer**	**Genetic modification**	**Vector used**	**Year of approval**	**Approving agency**
Strimvelis	Orchard Therapeutics	Adenosine deaminase (ADA) for Severe combined immunodeficiency (SCID) in HSCs	Retrovirus	2016	EMA, UK
Kymriah (Tisangenlecleucel)	Novartis Therapies	CD19 CAR-T cell for lymphoblastic leukemia	Lentivirus	2017	FDA, EMA, UK, Australia (+4 others)
Yescarta (Axicabtagene Ciloleucel)	Kite Pharmaceuticals	CD19 CAR-T cell for B cell lymphoma	Retrovirus	2017	FDA, EMA, UK (+3 others)
Invossa	Kolon TissueGene	TGF-β1 in chondrocytes for arthritis	Retrovirus	2017	Korea
Tecartus	Kite Pharmaceuticals	CD19 CAR-T cell for mantle cell lymphoma	Retrovirus	2020	FDA, EMA, UK
Libmeldy (Atidarsagene autotemcel)	Orchard Therapeutics	ARSA (arylsulfatase) gene in HSCs for leukodystrophy	Lentivirus	2020	EMA, UK
Breyanzi (lisocabtagene maraleucel)	Celgene	CD19 CAR-T cell for B cell and follicular lymphoma	Retrovirus	2021	FDA, Japan
Abecma	Bluebird Bio	BCMA (B-cell maturation antigen) CAR-T cell for multiple myeloma	Retrovirus	2021	FDA, EMA, UK (+3 others)
Carvykti	Legend Biotech	BCMA (B-cell maturation antigen) CAR-T cell for multiple myeloma	Lentivirus	2022	FDA

1.1.5 Table 2: List of approved *in vivo* gene therapy treatments

Trade name	**Manufacturer**	**Target gene and disease**	**Vector used**	**Year of approval**	**Approved by**
Gendicine	Shenzhen SiBiono GeneTech	TP53 for head and neck cancer	Recombinant Adenovirus (rAd-p53)	2003	SFDA (China)
Oncorine	Shanghai Sunway Biotech	Oncolytic virus for nasopharyngeal cancer	E1B-deleted Adenovirus (H101)	2005	SFDA (China)
Rexin-G	Epeius Biotechnologies	Mutant cyclin-G1 for soft tissue sarcoma and osteosarcoma	Retrovirus	2007	Phillipines
Neovasculgen (Cambiogenplasmid)	Human Stem Cells Institute	VEGF for peripheral vascular disease	Plasmid	2011	Russian Healthcare Ministry, Ukraine
Imlygic (Talimogene Laherparevec)	Amgen Pharmaceuticals	Oncolytic immunotherapy via addition of GM-CSF (immunostimulant)	HSV1 (Herpes simplex virus 1)	2015	FDA, EMA, UK, Australia
Exondys 51 (Eteplirsen), Vyondys 53 (Golodirsen)	Serapta Therapeutics	Antisense-RNA for *dystrophin* gene for Duchenne muscular dystrophy	RNA	2016, 2019	FDA
Spinraza	Ionis Pharmaceuticals	Antisense RNA for *SMN2* to treat Spinal Muscular Atrophy	RNA	2016	FDA, EMA, UK, Canada, Japan (+9 others)
Luxturna (Voretigene Neparvovec-rzyl)	Spark Therapeutics	RPE65 for inherited retinal dystrophy	AAV2	2017	FDA, EMA, UK, Australia (+2 others)
1. Tegsedi 2. Onpattro	1. Ionis Pharma 2. Alnylam	Antisense oligo/ds-siRNA against *transthyretin* for hereditary amyloidosis	RNA	2018	EMA, UK, FDA (+3 others)
Zolgensma	Novartis Therapies	SMN1 for spinal muscular atrophy in children	AAV9	2019	FDA, EMA, UK, Australia (+6 others)
Spikevax (Moderna COVID-19 vaccine)	Moderna Therapeutics	mRNA for stabilized spike protein of SARS-CoV-2 for COVID-19 vaccination	mRNA	2020	FDA, EMA, UK (+14 others)
Comirnaty	BioNTech	Lipid nanoparticle enveloped mRNA for SARS-CoV-2 spike protein for COVID-19 vaccination	mRNA	2020	FDA, EMA, UK (+21 others)
Viltepso	NS Pharmaceuticals	Anti-sense DNA for dystrophin pre-mRNA	DNA	2020	FDA, Japan

1.1.6 Limitations of gene therapy

For most diseases, providing a functional copy of a defective gene is not useful without modulation of its expression at the proper level, as high levels of expression can cause genotoxicity. Proteins used for therapeutic correction, like a AAV-*BbsI* construct used for treatment of Bardet-Biedl syndrome type I have been shown to cause overexpression toxicity in mouse studies (Seo *et al.*, 2013). There can also be toxicity associated with using specific promoters for expression, like in the case of ubiquitous promoters like CMV and retinal pigment epithelium (RPE) specific promoters which caused retinal toxicity following injection (Xiong *et al.*, 2019).

Genome editing, especially with CRISPR/Cas9 has significant safety concerns associated with possible off-target effects, which can lead to unwanted consequences like genome instability or deleterious mutations in essential genes (Luo, 2019). Diseases that make good targets for gene therapy are usually monogenic (caused by mutations in a single gene) like cystic fibrosis or sickle cell anemia. However, most diseases have complex phenotypes and are heterologous, which makes them difficult targets. For example, autosomal recessive retinitis pigmentosa has 25 known genes independently implicated as a genetic cause (den Hollander *et al.*, 2010). Gene therapies also need to go through a very rigorous, costly and difficult process to gain regulatory approval, which means they remain inaccessible to most people with these rare diseases (Cring and Sheffield, 2022).

1.2 Adeno-associated virus (AAV)

1.2.1 AAV as a gene therapy vector

Adeno-associated virus (AAV) is a helper-dependent virus belonging to the *Parvoviridae* family. It is a part of the genus Dependovirus and requires the presence of a helper virus such as herpes simplex virus (HSV) or adenovirus to carry out proliferative replication and infection. In the absence of the helper virus, AAV persists episomally in the cell, or integrates at specific sites in the host genome (Wu *et al.*, 2006). Though it is naturally found in several human and primate species, it does not cause disease to the best of our knowledge.

AAV is a single-stranded DNA virus (genome length 4.7kB) that can be sense or anti-sense. It encodes genes for replication (*rep*) and capsid assembly (*cap*). There are 4 genes within the genome required for replication named after their molecular masses, Rep78, Rep68, Rep52 and Rep40. The rep and cap genes are flanked by ITRs (inverted terminal repeats) on either side of the genome, which constitute the viral origins of replication and the cue for packaging. The natural AAV genome can integrate into the human genome into a locus called *AAVS1* on chromosome 19 and exists latently (Wang *et al.*, 2019). AAV is made up of an icosahedral capsid roughly 26 nm in diameter made up of 3 capsid proteins VP1, VP2 and VP3 made from a single ORF (open reading frame) in the *cap* gene via alternative splicing and a rare start codon (ACG), which means they have a common C-terminus. Assembly-activating protein (*aap)* is encoded in a frameshifted ORF within the same gene and is required for capsid assembly (Li and Samulski, 2020). The structure of the AAV genome is represented in Figure 1.2. The finished AAV virion is composed of 60 VP subunits with a ratio of 1:1:10 of VP1:VP2:VP3.

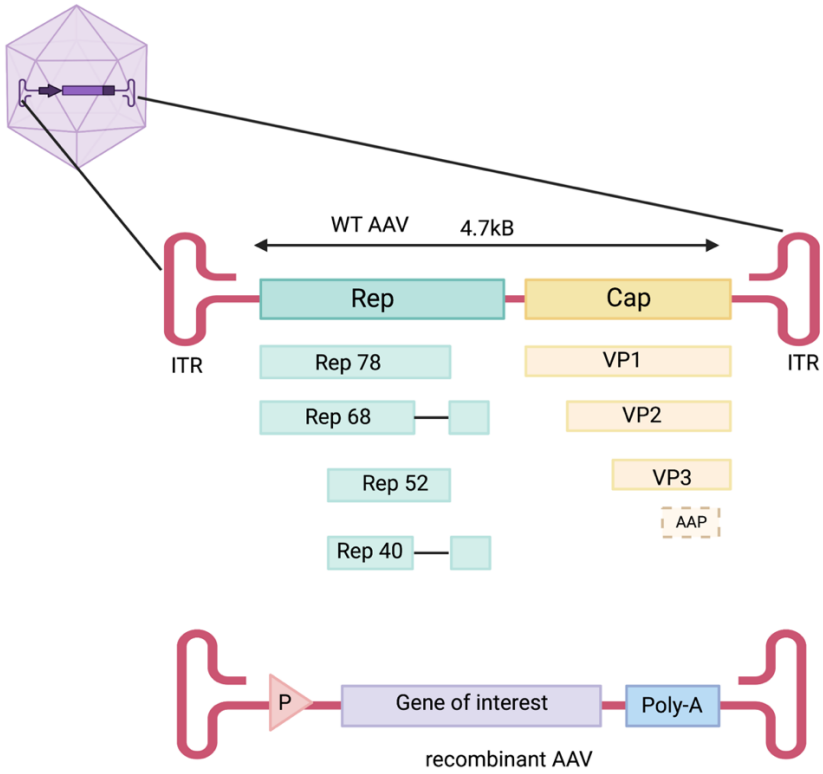

Figure 1.2: Structure of the wild-type and recombinant AAV genome. Various subunits VP1, VP2 and VP3 are encoded by the *cap* genes and replication proteins are encoded by the *rep* genes. In recombinant capsids, these two genes are replaced by the therapeutic gene of interest.

Recombinant AAV (rAAV) has a capsid structure and sequence that is identical to the wild-type (WT) AAV. However, they encapsidate genomes in which the AAV protein-coding sequences have been removed and replaced with therapeutic genes of interest (Wang *et al.*, 2019). The only genomic regions remaining are the inverted terminal repeats (ITRs) on either end which are needed for viral genome replication and packaging (Fig. 1.2). The removal of the replication and capsid genes from the genome increases the packaging capacity of the virion and makes it less toxic for *in vivo* applications. The genomic integration of the WT AAV genome is due to sequence similarities between the *rep* genes and the *AAVS1* locus, which is the natural site of integration. However, since the *rep* gene is removed in recombinant AAV, the risk of integration is negligible (Wang *et al.*, 2019).

Table 1.3 (contents of table adapted from (Li and Samulski, 2020; Pupo *et al.*, 2022)): Different AAV serotypes and their tissue tropism

Capsid name	**Origin**	**Tissue tropism**
AAV1	Monkey	Musculoskeletal, central nervous, retina, pancreas, heart, liver
AAV2	Human	Liver, central nervous, musculoskeletal, kidney, retina
AAV3	Human	Musculoskeletal, stem cells, liver, central nervous
AAV4	Non-human primate	Central nervous, retina, lung, kidney
AAV5	Human	Musculoskeletal, central nervous, lung, retina
AAV6	Human	Musculoskeletal, heart, lung, retina, central nervous
AAV7	*rhesus macaque*	Musculoskeletal, retina, central nervous, liver
AAV8	*rhesus macaque*	Liver, musculoskeletal, central nervous, retina, pancreas, heart
AAVrh.8	*rhesus macaque*	Musculoskeletal, central nervous
AAV9	Human	Liver, musculoskeletal, central nervous, lung, pancreas, retina, kidney, heart
AAV10	Cynomolgus monkeys	Heart, lung, liver, kidney
AAVrh.10	*rhesus macaque*	Liver, heart, musculoskeletal, lung, pancreas, central nervous, kidney
AAV11	Cynomolgus monkeys	Musculoskeletal, kidney, spleen, lung, heart
AAV12	Non-human primate	Salivary glands, nasal
AAVrh32.33	*rhesus macaque*	-

1.2.2 The art of AAV capsid design

Each subunit in the assembled AAV capsid have 9 variable regions on the surface. These regions determine the tropism, differential transduction efficiency with respect to cell type, and intracellular trafficking of the vector. Modifications in these regions

can influence the efficiency of transduction and the ability of neutralizing antibodies to bind to the virus (Li and Samulski, 2020). There are 13 identified broad serotypes of AAV (isolated from various organisms), which form the basis of most *in vitro* and *in vivo* studies (Table 1.3). Phylogenetic analysis has grouped these into 6 clades A-F. When viral protein structures of various capsids are mapped to the amino acid sequence, the variable regions correspond to different conformations in essential structures that are involved in tissue tropism, receptor binding and differential antigenic profile (Pupo *et al.*, 2022). Tropism, or cell type-serotype specificity arises due to the specific interactions between capsid-specific structures on the surface and cellular receptors and host machinery (Pipe *et al.*, 2019). With this background, we can logically follow the progression of the field, which has moved to the concept of a "receptor footprint" (Zolotukhin and Vandenberghe, 2022) possessed by different serotypes, that can be combined and exploited for advantages with respect to creating capsids more efficient at evading the immune system, transducing certain target organs preferentially and with higher efficiency. Innovative capsid engineering approaches have stimulated significant interest in developing novel gene therapy modalities.

Various techniques have been developed to genetically modify the AAV capsid, which include rational design and directed evolution. Using our knowledge of AAV binding and tropism, capsids showing enhanced transduction can be designed. Error-prone PCR can be used to introduce point mutations into capsid genes and gene shuffling can be used to initiate homologous recombination and form hybrid capsids. Chimeric capsids can also be made using targeted recombination (Li and Samulski, 2020). For example, by substituting a hexapeptide motif, a chimeric capsid containing components from AAV2 and AAV8 was created. This capsid showed enhanced transduction in vascular tissues (AAV2i8), and this chimera was further modified by adding a galactose-binding footprint from AAV9 to form AAV2i8G9. AAV2i8G9 targeted both the liver and muscular system with equally enhanced efficiency (Zolotukhin and Vandenberghe, 2022). Phylogenetic analysis has also been used to generate a library of "ancestral capsids" to possibly uncover lost efficient designs or novel mechanisms and pathways (Li and Samulski, 2020).

1.2.3 AAV: The Optimus Prime of viral gene therapy

As detailed in previous sections, despite the pursuit of clinically relevant constructs using various viral and non-viral methods of gene delivery, AAV remains the dominant choice as an *in vivo* vector, with around 137 active clinical trials using more than 20 distinct serotypes of AAV (Zhao *et al.*, 2022). This is due to its wide tissue tropism (Table 1.3) providing diversity in the organs and diseases that can be targeted, its ability to transduce both dividing and non-dividing cells, and its excellent safety profile. As discussed in the capsid design section, AAV design is modular and both the capsid and the therapeutic gene contained can be readily modified using molecular biology techniques (Pupo *et al.*, 2022). AAV also shows a low risk of genomic integration and long-term gene expression by existing latently in episomal form. Long-term follow up (~15 years) in Haemophilia B patients treated with AAV2 harboring the Factor IX gene has demonstrated stable Factor IX expression and bleed-free events in most patients (George *et al.*, 2020).

1.2.4 Current issues with AAV-mediated gene therapy

A drawback with AAV as a gene delivery vector is the limitations in carrying capacity. rAAVs optimally carry genes under 5kB, and hence therapeutic genes have to be designed carefully to consider both the target gene, but also including elements necessary for its optimum expression like promoters and polyadenylation sequences (Wang *et al.*, 2019). Strategies like deletion of functionally redundant regions in the transgene and providing the transgene split into two overlapping segments utilizing the host splicing machinery can circumvent this.

Despite being hailed as the one of the safest viral vectors due to its low immunogenicity, AAVs have been shown to activate an immune response in animal models, and high titers have been shown to cause renal failure, liver toxicity and low platelet levels (Cring and Sheffield, 2022). High doses of rAAV (>$5*10^{13}$ vg/kg) trigger inflammatory responses by CpG sequences within its genome binding to TLR9 (toll-like receptor), PAMPs (pathogen associated molecular pattern) within the capsid binding to TLR2 and a potential immune response generated by dsRNA (Ertl, 2022). These complement responses have resulted in anemia, low blood platelet counts and kidney damage despite patients being supplemented with steroids. As far as the

adaptive immune system is concerned, cells transduced with AAV and other antigen-presenting cells present immunogenic viral peptides to cytotoxic (CD8+) T cells via MHC Class I, which then prompts these killer T-cells to clear transduced hepatic cells and cause inflammation in the target tissue. These leads to decreased efficacy of gene transfer (Verdera *et al.*, 2020).

Data from clinical trials shows a dose-dependent relationship with AAV vector immunogenicity. Low doses of the viral vector are likely to cause milder inflammation and prevent loss of target gene expression, possibly via lesser MHC Class I presentation (Verdera *et al.*, 2020). Immune responses against AAV and the associated toxicity is the result of our highly tuned immune system which carefully monitors and eliminates threats. Our best bet to reduce the collateral damage suffered by AAV vectors is developing capsids that can transduce cells with higher efficiency or achieve higher protein expression, which would allow to lower the dose required for therapeutic benefit (Ertl, 2022). To focus on the first approach, we need to better understand interactions between the virus and host cells and tissues of various types. The widely advantageous tissue tropism that makes AAV so successful as a vector arises from the combined effects of differential cell surface binding, uptake, intracellular processing, and interactions between capsid proteins and host cell components. It is therefore critical for us to gain a deeper understanding of the structural and functional determinants of AAV's interactions with and within the host cell. Once we identify the interactions and host proteins that give rise to tropism, these can be used to overcome current limitations for AAV vectors.

1.3 AAV trafficking within the cell

AAV undergoes a long and complicated journey from the surface of the cell where the capsid first attaches, to its nucleus where its genome is released. This process of internalization and transport is rapid, with 13% of rAAV particles being internalized in just over a second, and at least one particle being detected inside the nucleus of half of the cells in as little as 15 minutes (Seisenberger *et al.*, 2001). The steps and mechanisms for this process are detailed below.

1.3.1 Surface attachment and internalization

The journey begins with the AAV capsid binding to glycans or glycoconjugates at the surface of the cell. These are termed as “primary receptors”. AAVs binding to these proteoglycans also have secondary attachment factors, and the idea is that primary receptor binding induces favourable conformational changes in the capsid that promotes its binding to their co-receptors and uptake (Riyad and Weber, 2021). For example, the AAV2 capsid utilizes heparan sulfate proteoglycan (HSPG) as the primary receptor and integrins or growth factor receptors as secondary receptors (Summerford and Samulski, 1998).

Several studies have attempted to dissect the mechanism of endocytosis used by AAVs, pre-dominantly using AAV2 as the exemplar. The importance of clathrin-mediated endocytosis was put forth as a hypothesis due to evidence showing that overexpression of a dominant-negative form of dynamin (an essential protein in the pathway) decreased AAV transduction. This was contradicted by another study which showed that a dominant-negative mutant of Rac1 (effector in macropinocytosis, has been shown to inhibit clathrin-mediated endocytosis) also inhibited AAV transduction, and overexpressing it had a positive effect (Riyad and Weber, 2021). Blocking macropinocytosis has been shown to increase AAV transduction in HuH7 and HepG2 cells and decrease it in HeLa cells (Dhungel *et al.*, 2021), while one study has shown AAV2 transduction to be independent of clathrin, caveolin and dynamin, instead relying on the clathrin-independent carriers and GPI-enriched endocytic compartment (CLIC/GEEC) endocytic pathway (Riyad and Weber, 2021). The only consensus is hence that various clathrin and caveolin dependent and independent mechanisms for internalization co-exist, and their usage depends on the capsid-receptor combination and the cell type being studied (Dhungel *et al.*, 2021). As the first steps in the process of transduction, receptor attachment and internalization continue to be studied as important determinants of tissue-specific tropism.

1.3.2 Intracellular trafficking and sorting

Following its entry into the cell, AAV now internalized with the endocytosed vesicle can take a variety of trafficking pathways, which will decide its fate with respect to successful transduction or degradation by the cell’s proteasomal degradation

machinery. Internalized vesicles first fuse with the early endosome (EE) also called the sorting endosome (SE). Evidence for AAV's travel starting at the EE include a dominant-negative mutant RAB5 (a key component of EEs) reducing transduction by >60% (Liu *et al.*, 2013), and colocalisation with EE-specific markers like EEA1(Liu *et al.*, 2013). Cargo received by the EE is trafficked to one of three locations (Fig. 1.3): late endosomes (LEs) and subsequently lysosomes for degradation, the slow or fast recycling pathway to the plasma membrane (through recycling endosomes or REs), or via retrograde trafficking to the trans-Golgi network (TGN) (Naslavsky and Caplan, 2018), which we will examine one by one.

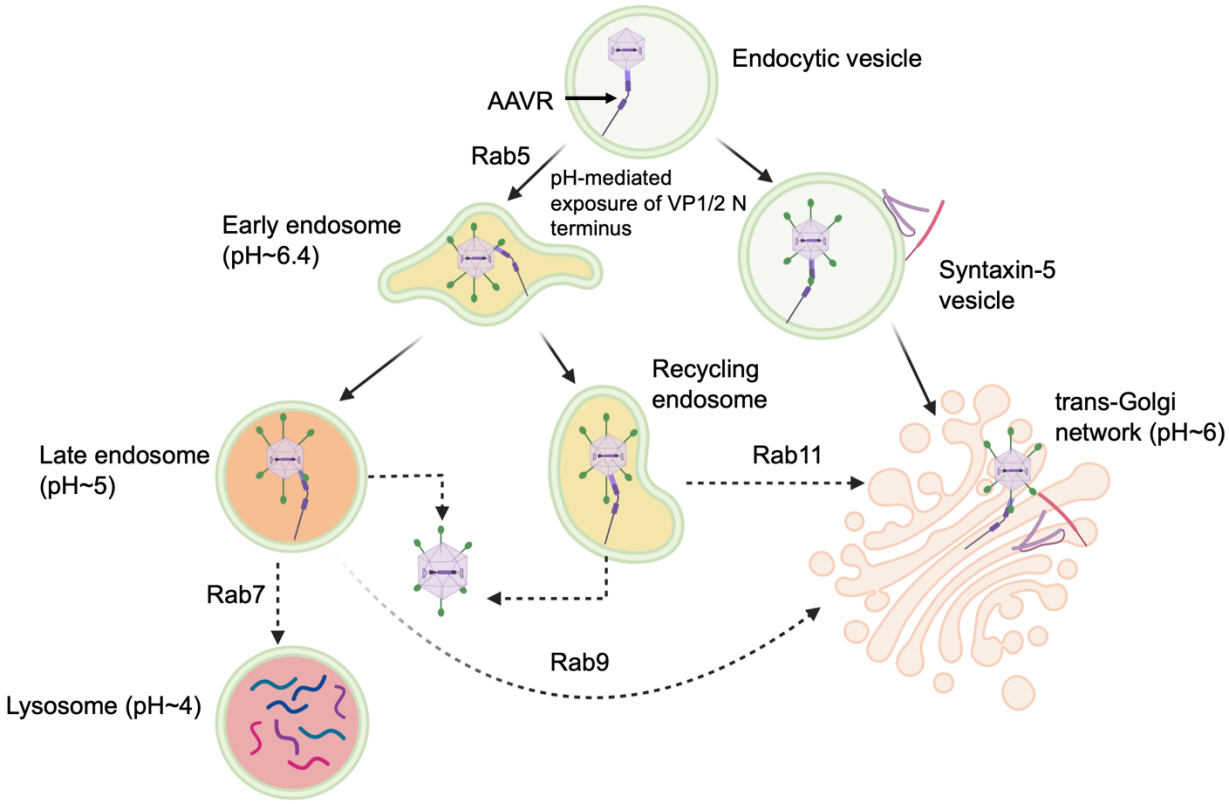

Figure 1.3: Endocytic trafficking of AAV. AAV internalized via endocytosis can be trafficked to three locations: late endosomes and lysosomes, recycling endosomes, or the TGN. Figure adapted from (Riyad and Weber, 2021).

The retrograde transport system is responsible for shuttling specific retrieved lipids and proteins from the plasma membrane to the Golgi or the ER, and has been shown to be used by toxins like shiga toxin, ricin and cholera toxin to ultimately reach the cytosol (Nonnenmacher *et al.*, 2015). The fast accumulation of AAV in the Golgi has

been shown by several studies (Nonnenmacher *et al.*, 2015), pointing to Golgi localisation as an important step in the process of transduction. Treatment with drugs disrupting the Golgi like brefeldin A or Golgicide A have been shown to block AAV transduction to a significant extent (Zengel and Carette, 2020). An interesting idea comes from the lower AAV2 transduction efficiency shown in NIH3T3 cells as compared to HeLa. It has been shown that there is a difference between endocytic processing in the two cell lines, and NIH3T3 cells showed impaired AAV2 transport to the Golgi (Riyad and Weber, 2021). This observation supports the hypothesis that the accumulation of AAV in the Golgi apparatus correlates with the success of transduction and poor efficiency of transduction in certain cell types could be explained by impairment in the retrograde pathway from the endosome to the Golgi. Using siRNA knockdown, the specific non-canonical retrograde transport pathway mediated by Syntaxin-5 (Nonnenmacher *et al.*, 2015) was shown to be the common and most efficient pathway used by most AAV serotypes.

The efficiency of AAV transduction has not convincingly been connected to the late endosome or to recycling endosomes, but there has been some evidence of a dose-dependent accumulation of transient AAV2 in the LE or RE compartments (Riyad and Weber, 2021). More importantly, modulators of systems like the ubiquitin-protease system and autophagy can lead to degradation of the capsid, which reduces the efficiency of transduction (Dhungel *et al.*, 2021). Inhibition of the ERAD (endoplasmic reticulum-associated protein degradation) pathway and limiting the ubiquitinylation of capsids by mutating critical residues have been used to increase AAV transduction (Dhungel *et al.*, 2021).

1.3.3 Escape into the cytoplasm and nuclear entry

After trafficking to the Golgi, the capsid has to escape into the cytosol in order to facilitate its entry into the nucleus. Structural data has shown that a domain that is critical for endosomal escape of AAV called PLA2 (phospholipase A2) encoded by the N-terminal domains of the VP1 and VP2 capsid proteins (Zengel and Carette, 2020) is hidden inside the capsid. PLA2 domain exposure is triggered by the acidic pH in the endosome. Trans-complementation of PLA2 via infection with AAV showing functional PLA2 activity can also help in effective endosomal escape (Riyad and Weber, 2021).

Additional factors might be required for this escape process like the discovered host factor GPR108 and capsid self-proteolytic ability (Dhungel *et al.*, 2021).

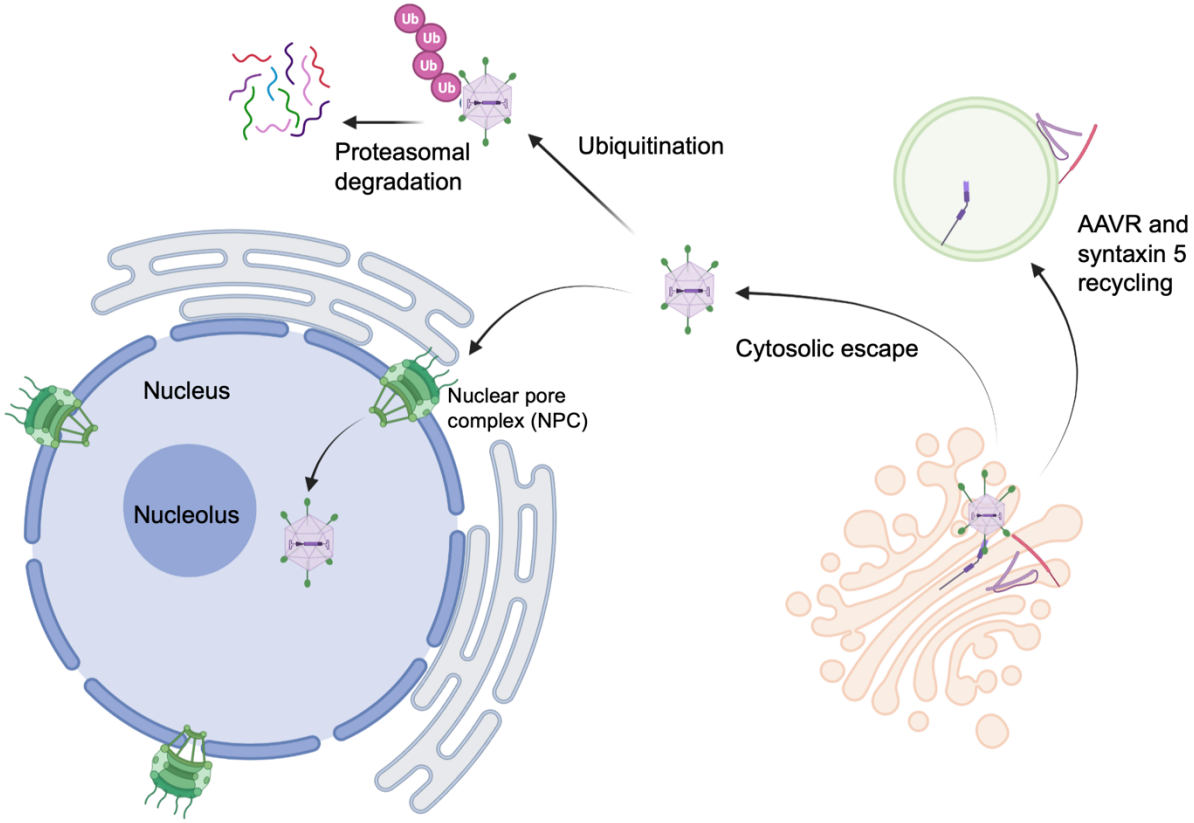

Figure 1.4: Cytoplasmic escape and nuclear entry by AAV. Structural changes in AAV aid its escape into the cytoplasm, following which it enters the nucleus using nuclear pore complexes. Some AAV virions may also be degraded from the cytoplasm via ubiquitination and subsequent proteasomal degradation.

Post-escape, AAV can enter the nucleus via nuclear pore complexes (NPCs) selectively (17% of docking virions enter successfully) (Riyad and Weber, 2021). AAV is reportedly able to bind to importin ß1 (owing to an NLS (nuclear localisation signal)-like motif formed by basic regions BR1-4 at the N-terminus of capsid proteins VP1/2 (Dhungel *et al.*, 2021). Post-entry, genome release and uncoating is a poorly understood phenomenon. Capsid particles accumulate inside the nucleolus and undergo genome release in the nucleoplasm according to the nuclear processing pathway currently proposed (Dhungel *et al.*, 2021).

1.4 AAVR as a host factor for AAV transduction

To find host factors in cells critical for AAV transduction, an unbiased haploid genetic screen was performed in 2016 (Pillay *et al.*, 2016) and the most significant hits were characterized. The most prominent gene to come out of this screen was KIAA0319L. Now also known as AAVR (adeno-associated virus receptor), KIAA0319L is a 150kDA protein which the authors showed to be essential for the transduction of a wide variety of AAV serotypes. AAVR is a glycosylated membrane protein that contains five PKD (polycystic kidney disease) domains. These are Ig-like domains that have been implicated in cell adhesion and attachment, and are often exploited by viruses for cellular entry (Pillay *et al.*, 2017). Further studies using a combination of cryo-electron microscopy data and evolutionary analysis have shown a serotype-dependent interaction pattern with the PKD1-5 domains, with AAV5 requiring PKD1 and others interacting with PKD2 for transduction (Dhungel *et al.*, 2021). AAVR-independent transduction has been observed in the capsids AAV4 and rh32.33, hinting towards an alternate pathway of endocytosis and trafficking independent of AAVR (Dhungel *et al.*, 2021).

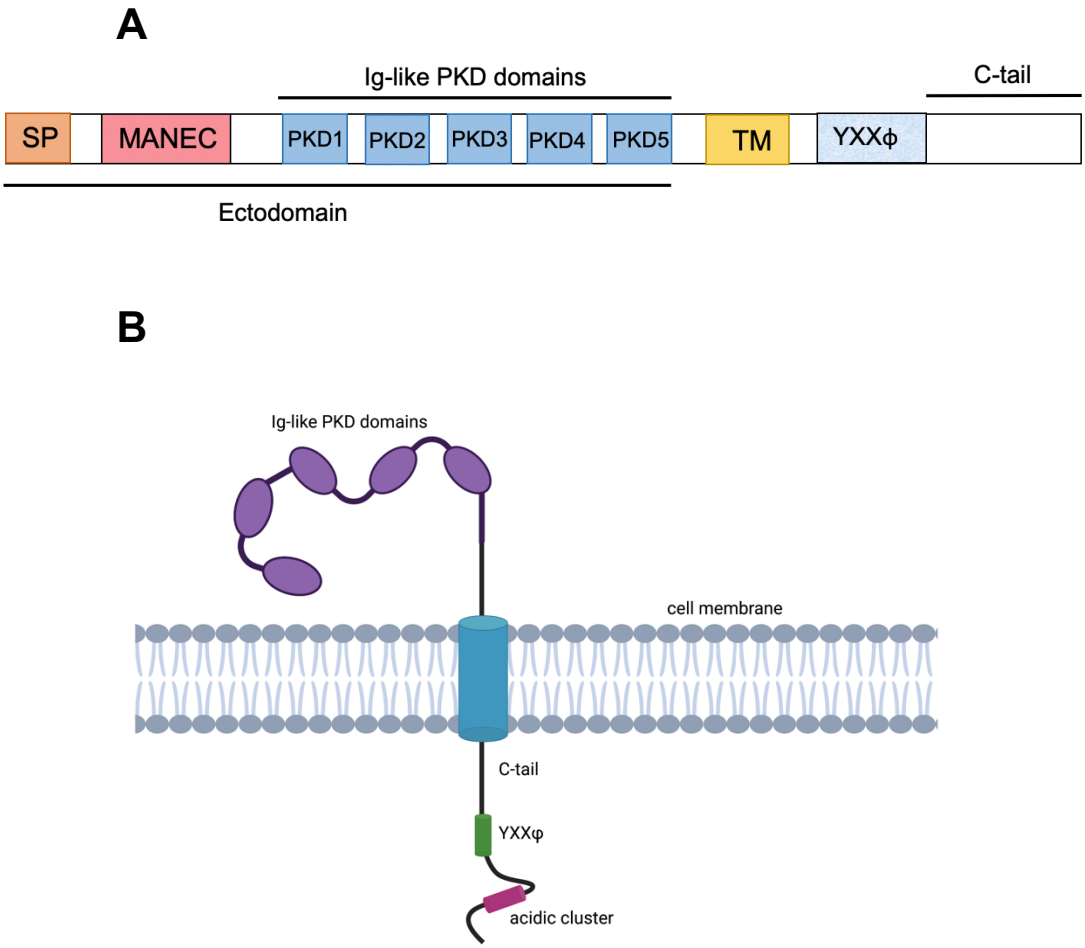

Figure 1.5: Schematic of the AAVR locus and protein domains. (A) Schematic of the AAVR genetic locus containing SP: signal peptide, MANEC: motif at N-terminus with 8 cysteines, TM: transmembrane, and an acidic cluster motif. (B) Simplified schematic of AAVR protein at the cell membrane.

AAVR is a receptor containing a C-terminal tail (Fig. 1.5) that shows motifs acting as cues for internalization and deletion of the C-tail leads to increased localisation of AAVR at the plasma membrane (Pillay *et al.*, 2016). This internalization motif initiates interactions with the clathrin-associated adaptor proteins AP-1 and A-2, and are such motifs are also found on recycled receptors like CI-MPR (Zengel and Carette, 2020). Deletion of the C-terminal tail prevents AAVR from being internalized and AAV transduction is unable to occur, and replacement of this motif with the internationalization motif from CI-MPR reinstates its ability to be transduced, which confirms that efficient internalization of AAVR is a critical step in successful AAV transduction (Pillay *et al.*, 2016).

As discussed earlier, AAV enters the cell through endocytosis and follows a route via retrograde trafficking pathways to the TGN. AAVR has been shown to follow a similar route and trafficking kinetics to AAV and localizing to the TGN, which suggests it might play a role in the correct sorting and routing of the capsid (Pillay *et al.*, 2016; Zengel and Carette, 2020). However, deletion of the C-tail and replacement with well-characterized endocytic motifs that are reported to traffic to non-TGN compartments also rescued the loss in transduction to some extent, which suggests that trafficking to the TGN might not be as strict a requirement as thought (Pillay *et al.*, 2016).

Knocking out AAVR in diverse cell types results in impaired transduction, but binding at the surface remains intact, which provides evidence for the functioning of AAVR, less as a receptor, and more as a "post-attachment factor" which influences endocytosis and possibly trafficking of AAV (Dhungel *et al.*, 2021). It also reminds us that there is a lot about the AAV-AAVR interaction that is unknown, and whether this interaction occurs at the surface or during endosomal trafficking is still unanswered.

1.5 The WDR11 protein as a novel host factor for AAV transduction

1.5.1 The WDR11 protein and its implications in disease

A member of the WD-repeat gene family, WDR11 was first described by Chernova et al (Chernova *et al.*, 2001) in 2001. It was discovered in a study attempting to characterize genetic loci implicated in glioblastomas, at the chromosomal location 10q25-26. The gene was aptly named *WDR11* as it was the 11th WD-repeat containing gene discovered and contains 12 WD repeats.

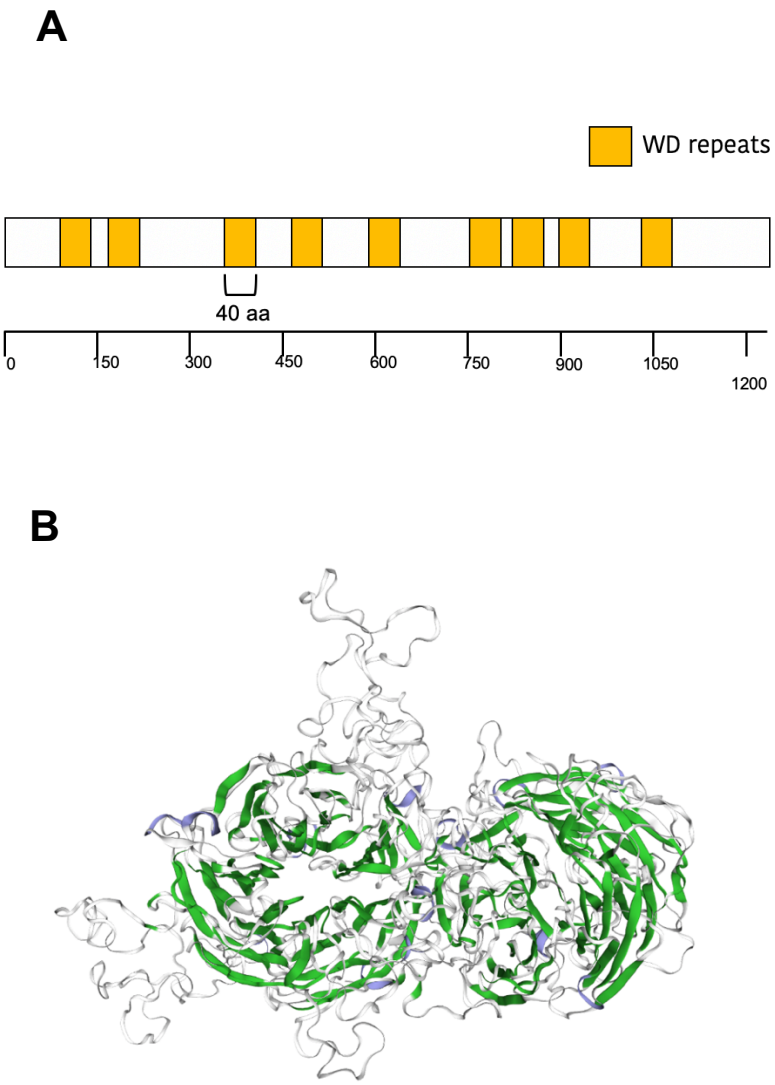

Figure 1.6: WDR11 locus and protein structure. (A) Schematic of the WDR11 genetic locus. (B) Beta-propeller structures formed by WD-repeats visualised using SWISS-MODEL.

WD repeats are loosely conserved units of around 40 amino acids that collectively fold into a symmetrical propeller-like structure. This propeller structure can take part in protein-protein interactions by forming complexes with other proteins (Chernova *et al.*, 2001). The WD-repeat family genes have been implicated in processes like signal transduction, transcriptional regulation, splicing, vesicular fusion and ciliogenesis, and members of this family contain between 4-16 WD-repeat motifs (Fig. 1.6) (Chernova *et al.*, 2001).

WDR11 is a 1224 amino acid (137 kDA) protein whose WD domains form two ß-propeller structures in each protein chain. Predictions on the basis of its yeast homolog

Actin Interacting Protein 1 (AIP1) show possible actin-binding ability (Kim *et al.*, 2010). Disruptions in WDR11 have been found in human glioma and breast cancer cells and patients with idiopathic hypogonadotropic hypogonadism (IHH), Kallman syndrome and pituitary hormone deficiency, all conditions involving endocrine signaling abnormalities (Taylor and Mossman, 2015). WDR11 has been shown to have definitive interaction with EMX1, a homeobox transcription factor that dictates differentiation fates for cells in the developing nervous system and the development of olfactory neurons (Kim *et al.*, 2010), and involvement in trafficking of GLI3 during ciliogenesis, a process that further downstream mediates the transcriptional activation of various genes in the hedgehog signalling pathway (Kim *et al.*, 2018). Though it has been predicted to be involved in autophagy, tumor suppression and endocytosed vesicle trafficking, the exact molecular mechanisms surrounding its diverse functions are largely unknown.

1.5.2 WDR11 in intracellular trafficking

WDR11 has been shown to localize to TGN and has been shown to interact with viral components in the event of HSV-1 (herpes simplex virus) HSV-1 infection (Taylor and Mossman, 2015). An interesting breakthrough in elucidating the function of WDR11 in endosome to Golgi trafficking came from a study conducted by Navarro Negredo et al (Navarro Negredo *et al.*, 2018). The authors used a forward genetic approach to screen for genes involved in the endosome-to-TGN trafficking of acidic cluster-containing proteins, which revealed WDR11 as a significant factor. Acidic clusters are sorting signals found in proteins that are trafficked via clathrin-coated vesicles between the TGN and endosomes.

Cargo trafficking within the cell is carefully regulated by distinct “postal addresses” possessed by various organelles within the cell in the form of a distinct lipid and protein composition for easy identification by trafficking complexes. Selected cargo is packaged into vesicles and carried via the cytoskeleton to the target organelle, where the vesicle docks with the help of tethering proteins and fuses with the acceptor membrane (Navarro Negredo *et al.*, 2018). For trafficking between the endosomes and the Golgi, this process happens with the help of clathrin coats and cargo-binding

adaptor proteins like the AP-1 complex, which is a significant mediator of endosome-to-TGN trafficking of clathrin-coated vesicles.

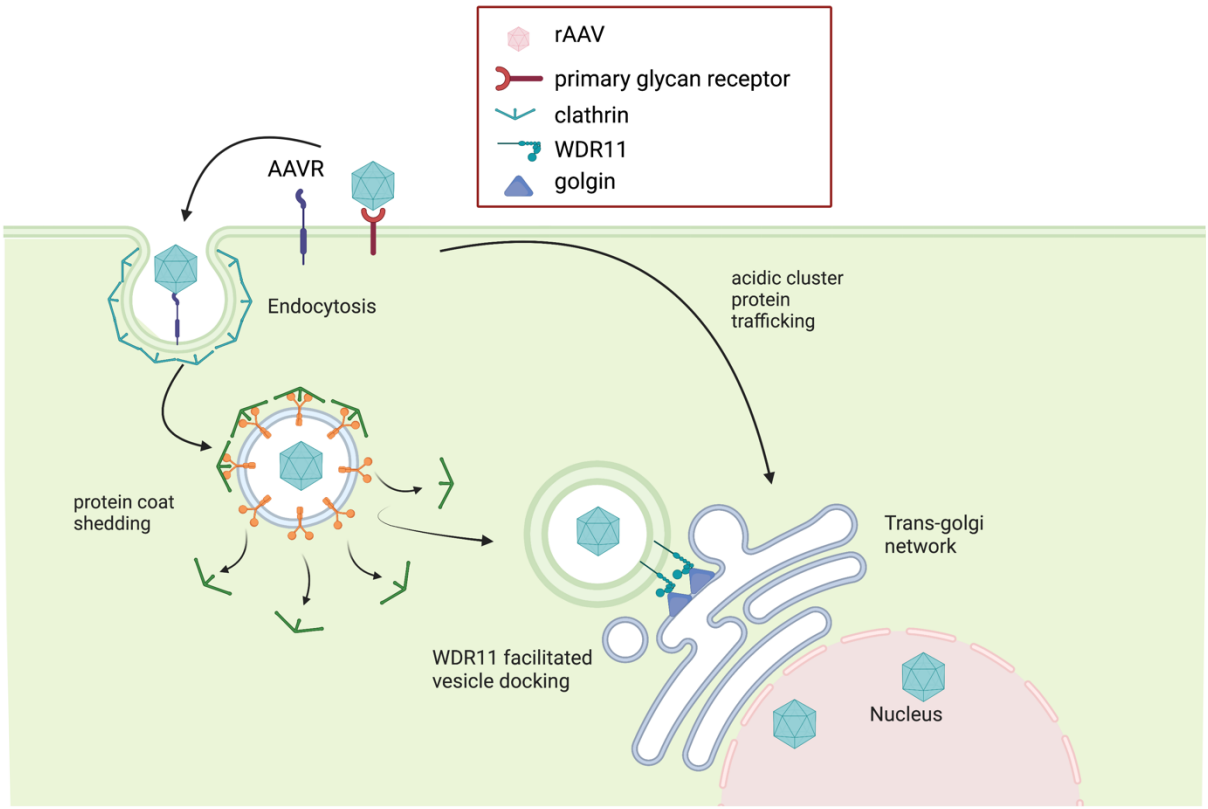

Figure 1.7: WDR11-dependent AAV trafficking from the surface of the cell to the nucleus. Post primary glycan receptor attachment, AAV particles attach to "post attachment" factors like AAVR and are internalized. AAV containing vesicles are carried along cytoskeletal tracks and delivered to the TGN in a WDR11 complex-dependent manner through its association with Golgin 245.

WDR11 was found to be part of a stable complex with two other proteins C17orf75 and FAM91A1 (Navarro Negredo *et al.*, 2018) which appears to be specific for AP-1-generated vesicles. WDR11 and FAM91A1 were shown to be indispensable for proper localisation. According to the current model, the WDR11 complex is responsible (after its recruitment onto AP1-derived vesicles) for the further recruitment of a tethering protein TBC1D23 (Shin *et al.*, 2017). This tether is then used by the complex to dock on to the TGN via the activity of two Golgin proteins Golgin-97 and Golgin-245 (Fig

1.7). Interestingly, *WDR11* knockout *in vitro* was shown to increase susceptibility of the cells to ricin, a toxin that as mentioned in previous sections utilizes the endosome-to-TGN route (Bassik *et al.*, 2013). This is a paradoxical finding considering that WDR11 is the very complex that might facilitate this endosome-to-TGN transport, but it might be explained by the compensation of trafficking by other redundant pathways which will be elaborated on in the Discussion section.

1.5.3 WDR11 and AAV

WDR11 knockout *in vitro* was shown to cause increased localisation of acidic cluster-containing proteins at the plasma membrane (Navarro Negredo *et al.*, 2018). More interestingly, an increased signal at the surface was seen for two endogenous AP-1 cargo proteins, one of which was KIAA0319L or AAVR, a protein of significant interest to us due to its implication in AAV transduction as detailed above. The study demonstrated the mislocalisation and plasma membrane distribution of AAVR despite overall protein levels in the cell remaining unchanged (Navarro Negredo *et al.*, 2018). This provokes an interesting question in the context of AAV transduction: could WDR11 be a novel host factor affecting AAV transduction via altering AAVR distribution and hence capsid trafficking within the cell?

Preliminary experiments and knockdown and knockout studies by Alex Sherman, a Master's student in the Gene and Stem Cell Therapy Program showed convincingly that shRNA-mediated knockdown of WDR11 *in vitro* increases AAV transduction for a wide variety of cell-types and serotypes (Sherman, 2021). This project aims to elucidate the role played by WDR11 in altering AAV transduction to some extent and to hopefully propose a potential mechanism by which it is doing so.

1.6 Hypothesis and aims

1.6.1 Hypothesis

WDR11 influences intracellular rAAV trafficking and acts as a restriction factor for rAAV transduction.

1.6.1 Aims

1. Generate WDR11 KO cells in a variety of cell lines.
2. Examine the effect of WDR11 overexpression in WT and WDR11 KO cells.
3. Examine cell surface levels of AAVR in WDR11 KO cells.
4. Examine the subcellular localisation of AAVR and WDR11.

2. Materials and methods

2.1 Materials and reagents

2.1.1 Table 2.1: Solutions and buffers used in this study

Buffer name	Composition
Cell lysis buffer	20 mM Tris-Cl, 150 mM NaCl, 1% (v/v) Triton X-100, 0.5% (w/v) sodium deoxycholate, 0.1% (w/v) SDS
PBS (phosphate buffered saline) 1x	PBS (phosphate buffered saline) 10x (Life technologies, CA): 1.37 M NaCl, 27 mM KCl, 100 mM Na_2HPO_4, 18 mM KH_2PO_4, diluted to 1x in MilliQ water at pH 7.4 and autoclaved before use
Anode I buffer (western blot)	0.3M Tris, 10% (v/v) methanol in triple distilled water (TDW), pH 10.4
Cathode buffer (western blot)	25 mM Tris-Cl, 40 mM glycine, 10% (v/v) methanol in TDW, pH 9.4
Anode II buffer (western blot)	25 mM Tris, 10% (v/v) methanol in TDW, pH 10.4
PBST	1x PBS + 0.1% (v/v) Tween® 20
FACS buffer	Sterile 1x PBS + 5% (v/v) FBS
Phosphate buffer (300 mM)	195 mM Na_2HPO_4, 105 mM NaH_2PO_4 (pH 7.05)
HEPES Buffer (2X)	140 mM NaCl, 50 mM HEPES (pH 7.05)
TE buffer ($1/10^{th}$)	1mM Tris-Cl, 0.1 mM EDTA (pH 7.6)
TAE buffer (1x)	40 mM Tris-acetate, 1mM EDTA (pH 8.6)

2.1.2 Table 2.2: Antibodies used in this study

Abbreviations: WB: Western blot, IF: Immunofluorescence

Antibody name	Clonality	Host	Target	Catalogue no.	Source	Concentration
Primary						
WDR11	Polyclonal	Rabbit	Human	ab93871	Abcam	1:5000 (WB)
WDR11 (BWDR2)	Polyclonal	Rabbit	Mouse	TA590853	Thermo Fisher	1:5000 (WB)
KIAA0319L	Polyclonal	Mouse	Human	ab105385	Abcam	1:5000 (WB) 1:500 (IF)
GM-130	Polyclonal	Rabbit	Human	Ab52649	Abcam	1:500 (IF)
GAPDH	Monoclonal	Mouse	Human	Ab8245	Abcam	1:5000 (WB)
Secondary						
Alexa Fluor 488	Polyclonal	Goat	Mouse	A11001	Invitrogen	1:1000 (IF)
Alexa Fluor 488	Polyclonal	Goat	Rabbit	A11034	Invitrogen	1:1000 (IF)
Alexa Fluor 594	Polyclonal	Goat	Rabbit	A11012	Invitrogen	1:1000 (IF)
Horseradish peroxidase (HRP)	Polyclonal	Donkey	Mouse	AP192P	Merck	1:5000 (WB)
Horseradish peroxidase (HRP)	Polyclonal	Donkey	Rabbit	AP182P	Merck	1:5000 (WB)

Fluorescent dyes and conjugated antibodies						
Haemagglutinin (HA) epitope HRP-conjugated	Monoclonal	Mouse		#2999	CellSignaling Technologies	1:5000 (WB)
DAPI				D1306	Invitrogen	1:5000 (IF)
Alexa Fluor 674 Phalloidin				A22287	Thermo Fisher	1:1000 (IF)

2.2 Cell culture

2.2.1 Culturing conditions

All cells were cultured at 37oC/5% CO2 in 75 cm2 canted tissue culture flasks with filter caps. (Corning, NY). During experiments, cells were also cultured in Corning® 6,12,24 or 96 well-plates. Cells were passaged when they reached around 75-80% confluency to avoid overgrowth. Media was removed and cells were washed once with 1x PBS. TrypLE™ Express (Invitrogen) was used to detach the cells. Post-detachment, the TrypLE™ was deactivated with media and an appropriate number of cells were transferred to a new flask/plate with fresh complete media.

2.2.2 Cryopreservation and thawing

Cells were cryopreserved at a low passage number for long-term storage in liquid nitrogen. After detachment with TrypLE™, cells were pelleted and resuspended at a concentration of 1 x 10^6 cells/mL in media containing 10% (v/v) DMSO (Sigma-Aldrich) as a cryopreserving agent. This cell suspension was aliquoted into 1mL cryovials and cooled to -80^o (gradually using a Corning® CoolCell tank) and then moved to liquid nitrogen tanks for long-term storage. To thaw cryopreserved cells, DMSO-containing cell mixture was warmed to 37^o. Cells were then pelleted at 1500 rpm for 5 minutes. The supernatant was removed, and cells were re-suspended in 10mL of fresh complete media before being transferred to a T75 flask.

2.2.3 Table 2.3: Cell lines and respective media used

Cell line	Origin	Media and supplements
HEK293T	Embryonic kidney	D10: DMEM (Dulbecco's Modified Eagle Media) with 10% (v/v) foetal bovine serum (FBS) + 1% (v/v) Penicillin-Streptomycin
eHAP (engineered haploid)	Myeloid leukemia cell-derived line	IMDM (Iscove's Modified Dulbecco's Medium) with 10% (v/v) FBS + 1% (v/v) Penicillin-Streptomycin
HeLa	Cervical cancer cell line	D10: DMEM (Dulbecco's Modified Eagle Media) with 10% (v/v) FBS + 1% (v/v) Penicillin-Streptomycin

2.3 Western blot

2.3.1 Cell lysis and protein extraction

Cells were counted (1 x 10^6) and collected in 1.5 mL Eppendorf® tubes. Cells were pelleted by centrifugation at 12,000 rpm for 5 minutes. Media was removed and cells were re-suspended in 250µL lysis buffer (Table 2.1) with the addition of 1x PIC (protease inhibitor cocktail III, A.G Scientific, San Diego, CA). All lysates were kept on ice. Samples were then sonicated for 10 cycles (30s on, 30s off) in a BioRuptor® Plus Sonicator (Diagenode, NJ) to fragment the membrane.

2.3.2 SDS-PAGE (polyacrylamide gel electrophoresis)

Protein samples (20µL) were mixed with 1x NuPAGE® LDS buffer and 10mM Dithiothreitol (DTT). Samples were then incubated at 85^{o}C for 10 minutes and then loaded on a 4-12% BOLT™ Bis-Tris 10,12 or 15 well gel (Invitrogen). A Precision Plus Protein™ Standard ladder was loaded alongside samples and used to determine protein sizes (BioRad, CA). Gel was assembled into the XCell SureLock® Mini Cell and filled with 500 mL of 1x BOLT™ SDS running buffer. Gel was run at a constant voltage of 160 V for 60 minutes.

2.3.3 Transfer

Gel was removed from the cassette and equilibriated in cathode buffer (Table 2.1) for 5 minutes. Immobilon Polyvinylidene fluoride (PVDF) membrane (Merck) was cut to the dimensions of the gel, activated with 100% methanol by immersing for 10 seconds, rinsed in triple distilled water (TDW) and equilibriated in Anode II buffer (Table 2.1) for 5 minutes. Sheets of filter paper were cut to the dimensions of the gel and membrane and soaked in the Anode I, Anode II and Cathode buffers for 5 minutes, and a transfer

stack (Figure 2.1) was assembled. The transfer was run in a semi-dry blotter (BioRad) at a constant current of 0.08 A for 60 minutes.

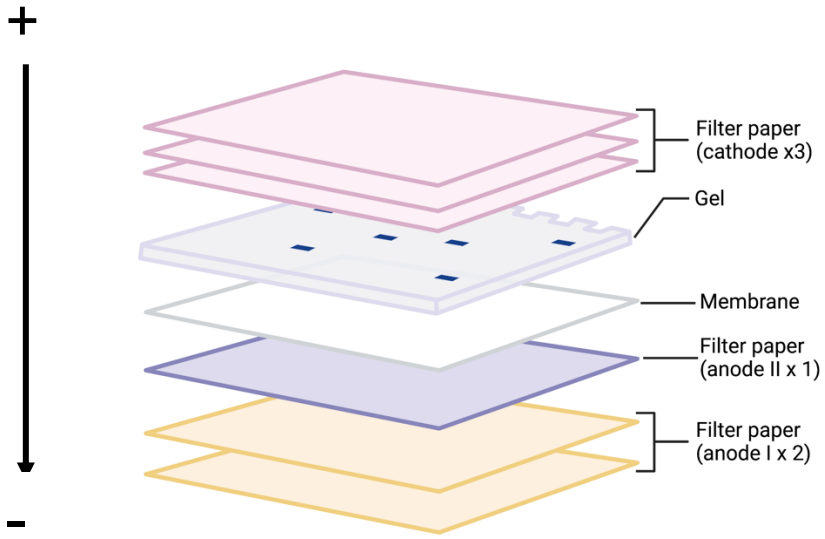

Figure 2.1: Western blot transfer stack. Arrow represents the direction of electric current.

2.3.4 Blotting and imaging

Following the transfer, the protein-bound PVDF membrane was incubated in blocking buffer (5% (w/v) skim milk in PBST) for one hour. Membrane was then incubated overnight with the appropriate antibody (Table 2.2) in 2.5% (w/v) skim milk in PBST. The next day, the membrane was washed thrice for 5 minutes each with PBST before incubation of the appropriate secondary antibody in 2.5% (w/v) skim milk in PBST for 2 hours. Post incubation, the membrane was again washed thrice for 5 minutes each with PBST. Membrane was imaged by preparing a SuperSignal West Pico chemiluminescent substrate as per the manufacturer's instructions in the BioRad ChemiDoc Touch Imaging System.

2.4 Immunofluorescence

Cells were seeded overnight (5 x 10^4) in an 8-well µ-slide (Ibidi, Martinsried, Germany). Cells were washed with room temperature PBS and fixed for 15 minutes at room temperature with 4% (v/v) paraformaldehyde (Thermo Fisher). Cells were 43ermeabilized with 0.5% (v/v) Triton® X-100 (Sigma-Aldrich) in PBS for 5 minutes. Cells were blocked with 3% (w/v) Bovine Serum Albumin (BSA) (Bovogen, Australia) for 30 minutes. Cells were incubated with appropriate primary and secondary antibodies (Table 2.2) in the blocking solution for one hour each at room temperature and covered with aluminum foil to avoid photobleaching. DAPI was included in the penultimate PBS wash to stain the nucleus. Ibidi Mounting Media (Ibidi, Germany) was added dropwise to the wells till the entire surface of the well was covered. Slides were imaged using a 93x glycerol immersion objective on the Leica SP8 confocal microscope at the Centenary Institute, University of Sydney. Images were processed using Fiji software.

2.5 Molecular cloning

2.5.1 Agarose gel electrophoresis

Agarose gel (1-2% (w/v), Astral Scientific Australia) was prepared in 1x TAE buffer (Table 2.1). GelRed (1 µL, Biotium, CA) was added to the melted agarose to visualise DNA. DNA was mixed with 1x Purple Gel Loading Dye (New England BioLabs) and electrophoresed at 80-110V for 45 minutes – 1 hour. DNA ladders (1kB and 100 bp, Promega, WI) were used to compare DNA band sizes and the gel was imaged using the G-box UV Transilluminator (Syngene, Cambridge).

2.5.2 Bacterial transformation and plasmid recovery

Chemically competent DH5α *E. coli* cells were generated within the lab and stored in 100 µL aliquots at -80°C. Bacteria were grown in Luria-Bertani (LB) broth made up of 2% (w/v) LB powder (Life technologies) in water, or on LB-Agar plates made by supplementing LB with 1.5% (w/v) bacteriological grade agar (Bacto, Australia). Bacterial transformation was performed using the heat shock method. Competent cells were thawed on ice, mixed with plasmid DNA and incubated on ice for 20 minutes. Heat shock was performed at 42° for 90 seconds followed by incubation on ice for 5 minutes. LB broth (750µL) was added for recovery at 37° for 45 minutes. Transformed

bacteria were streaked on LB agar plates containing appropriate selection antibiotic depending on the plasmid (Ampicillin: 100µg/mL, Kanamycin: 50µg/mL) and incubated at 37° for 16 hours. Post-incubation, single bacterial colonies were picked using a pipette tip under sterile conditions and grown in 2mL LB containing the appropriate antibiotic overnight while shaking at 37°C. Plasmid DNA was purified using the NucleoSpin® Plasmid purification kit (Macharey-Nagel, Germany) according to the manufacturer's instructions. For more amounts of concentrated DNA, starter cultures were used to inoculate 100 mL of sterile LB containing the appropriate antibiotic and grown overnight while shaking at 37°C. Plasmid DNA was then purified using the NucleoBond® Xtra Midi Plus Kit (Macharey-Nagel, Germany) according to the manufacturer's instructions.

2.5.3 Table 2.4: List of PCR and cloning reagents

Reagent	Source
dNTP	Bioline, UK
Hot Phusion®	New England Biolabs
5x HF buffer	New England Biolabs
DMSO	Sigma-Aldrich
5x GC buffer	New England Biolabs
NotI	New England Biolabs
XbaI	New England Biolabs
AgeI	New England Biolabs
MscI	New England Biolabs
CutSmart® Buffer	New England Biolabs
T4 DNA Ligase	New England Biolabs
T4 DNA Ligase buffer	New England Biolabs

2.5.4 Table 2.5: List of PCR primers

Primer	Sequence (5'-3')	Purpose
T7E1 WDR11 5'	ACACACACGTCACATTTGGG	Determining indel in KO cells
T7E1 WDR11 3'	GACATGCTTCTCCACTTCGC	Determining indel in KO cells
M13Rev	CAGGAAACAGCTATGAC	Sanger sequencing
TagBFP 5'*XbaI*	TCTTTCTTTGGTCTAGACATGAGCGAGCTGATTAAGG	PCR amplification-tagBFP
TagBFP 3'*AgeI*	TAGTGAATACCGGTCAGTTAAGCTTGTGCCCCAGTTTG	PCR amplification-tagBFP
pcDNA3.1 revseq	TAGAAGGCACAGTCGAGGC	Sanger sequencing
WPRE revseq	CAGCGTATCCACATAGCGT	Sanger sequencing
Mcherry forseq	CCATTATGATGCCGAAGTC	Sanger sequencing

2.5.5 Construction of pFUW-WDR11 lentiviral overexpression vector

A cDNA encoding the full length *WDR11* open reading frame was cloned into the pcDNA3.1 vector with an N-terminus HA tag and C-terminus HA tag (referred to as *HA-WDR11* and *WDR11-HA* respectively) by Cynthia Metierre in the Gene and Stem Cell Therapy Program at the Centenary Institute. *HA-WDR11* was then sub-cloned into the pFUW-mCherry-2A-C-HA vector (obtained from Cynthia Metierre, Gene and Stem Cell Therapy Program) by Alex Sherman (GSCT Program). To clone *WDR11-HA* into the pFUW vector, both the pFUW vector and pcDNA3.1-WDR11-HA were digested with *NotI* and *XbaI-HF* restriction enzymes in CutSmart® Buffer for 1 hour at 37^{o}C. The digested inserts and vector were run on an agarose gel and purified using the NucleoSpin® Gel and PCR Cleanup kit according to manufacturer's instructions. The *WDR11-HA* insert was ligated into the vector at a molar ratio of 1:3 with T4 DNA Ligase in T4 DNA Ligase Buffer for 30 minutes at room temperature. A ligation mixture with only the digested vector present was also incubated as a control. The ligation mixtures were then used to transform competent *E. coli* cells. DNA from successful clones was prepared and purified as described above. Correct insertion of the insert was confirmed by digestion with *MscI* in CutSmart® Buffer for 1 hour at 37^{o}C. Samples showing the expected band sizes 6736 bp, 4202 bp, 1416 bp, 1411 bp were

sequenced using primers WPRE revseq and Mcherry forseq as detailed in Table 2.5 at the Australian Genome Research Facility (AGRF).

2.5.6 Construction of BFP-WDR11 fusion protein (pcDNA3.1 vector)

A plasmid containing the full-length sequence of *Tag-BFP* was obtained from Rajini Nagarajah. Tag-BFP was amplified using PCR with primers designed to target the N-terminus of WDR11 under the following cycling conditions (30 cycles) and reaction composition in an Eppendorf Mastercycler®:

Component	**Volume (µL)**
dH_2O	14.5
5 X GC buffer	5
Forward primer- tagBFP 5' *XbaI*	1.25
Reverse primer- tagBFP 3' *AgeI*	1.25
dNTP	0.5
DMSO	0.75
Hot Phusion®	0.25
Template	1
Total reaction volume	25

Cycling step	**Temperature (°C)**	**Duration**
Denaturation	95	10s
Annealing	55	20s
Extension	72	30s

To clone Tag-BFP into the pcDNA3.1-WDR11-HA vector, both the insert and the vector were digested with *XbaI* and *AgeI* in CutSmart® Buffer for 1 hour at 37°C. The digested inserts and vector were run on an agarose gel and purified using the NucleoSpin® Gel and PCR Cleanup kit according to manufacturer's instructions. Ligation and purification of ligated plasmid DNA was carried out as in section 2.5.5. Correct insertion of the insert was confirmed by digestion with *XbaI* and *AgeI* in CutSmart® Buffer for 1 hour at 37°C. Samples showing the expected band sizes were sequenced using the primer pcDNA3.1revseq (Table 2.5) at the Australian Genome Research Facility (AGRF).

2.6 CRISPR/Cas9-mediated knockout of WDR11 in multiple cell lines

The WDR11 protein was knocked out in 3 cell lines- HEK293T, eHAP and HeLa using CRISPR-Cas9. Three single guide RNAs (sgRNAs) were designed by Dr. Bijay Dhungel (Gene and Stem Cell Therapy Program, Centenary Institute) and validated by Alex Sherman (Gene and Stem Cell Therapy Program, Centenary Institute). A guide RNA sequence sgRNA3 (accgGGATAGCAAGCAGTAAATCG) was cloned into Addgene plasmid #52628-mCherry (provided by Dr. Charles Bailey) by Divya Gokal (Gene and Stem Cell Therapy Program, Centenary Institute). A second plasmid containing a non-targeting scramble guideRNA (scr, caccgGATTGTATGAACGCAATAGC) was used as a CRISPR control, and cells treated with this construct were used as WT cells for all future experiment.

Plasmids expressing sgRNA3-mCherry and Cas9-eGFP were co-transfected into wild-type (WT) HEK293T cells using Lipofectamine™ 3000 and WT eHAP cells using Turbofectin® as detailed in section 2.9. Doubly positive cells expressing both mCherry and eGFP were deposited as single cells into a 96-well plate using fluorescence activated single-cell sorting (FACS). These distinct clones were expanded as monoclonal populations. Plates were incubated for 10 to 14 days and scanned under a light microscope to identify colonies originating from a single cell. Colonies were allowed to grow and expanded once they reached confluency.

In HeLa cells, a similar workflow was used with sgRNA3-puromycin (encoding gene for resistance to puromycin) and Cas9-blasticidin (conferring resistance to blasticidin) being co-transfected into WT cells using Lipofectamine™ 3000 as detailed in section 2.9. Cells were grown in double-selection media containing 2 µg/mL puromycin and 10 µg/mL blasticidin till control (untransfected) cells were dead. The bulk population obtained after selection was screened for loss of WDR11 expression. After western blot confirmation, limiting cell dilution was performed to isolate single cells in a 96-well plate, which were expanded as monoclonal populations. Cells were detached from plates as described in section 2.2.1, counted and diluted in D10 media (Table 2.3) at a concentration of 5 cells/mL. Cell solution was seeded into each well of a 96-well plate using a multi-channel pipette. Colonies were expanded as described above.

WDR11-knockout clones were validated by western blot for loss of protein expression. One clone each from HeLa, HEK293T and eHAP were further validated at the genomic level for *WDR11* KO. Genomic DNA was isolated from cells using the Extracta™ genomic DNA extraction kit (Quanta Biosciences) as per the manufacturer's instructions.

A region of 490 bp flanking the sgRNA target site was amplified using the following PCR cycling conditions (30 cycles) and reaction composition in an Eppendorf Mastercycler®:

Component	Volume (µL)
dH_2O	14.1
5 X HF buffer	5
Forward primer- T7E1 WDR11 5'	1.25
Reverse primer- T7E1 WDR11 3'	1.25
dNTP	0.5
DMSO	0.75
Hot Phusion®	0.15
Template	2
Total reaction volume	25

Cycling step	Temperature (°C)	Duration
Denaturation	95	10s
Annealing	60	20s
Extension	72	20s

The amplified product was purified using the NucleoSpin® Gel and PCR Cleanup kit according to manufacturer's instructions and cloned directly into a pTOPO vector using the ZeroBlunt® TOPO® PCR cloning kit (Life Technologies). These plasmids were then transformed into DH5α *E. coli* and plasmid DNA was extracted. These samples were then sent for Sanger sequencing at the Australian Genome Research Facility (AGRF) in Sydney with primer (m13 Rev, Table 2.5)

2.7 Lentiviral constructs

2.7.1 Virus preparation

HEK293T cells were seeded (4 x 10^6) in a 10 cm plate and allowed to attach overnight. On the next day, media on cells were changed to fresh chloroquine (25µM) -containing D10 (DMEM+10% FBS+ 1% PenStrep). Transfection solutions as detailed below were prepared using calcium phosphate transfection reagents (at room temperature).

Solution A – Components	Volume
PO_4^{3-}	4.3 µL
2x HEPES	860 µL

Solution B- Components		Volume/amount
2.5 M $CaCl_2$		86 µL
1/10th TE		774 µL
Cargo plasmid		10 µg
Viral packaging plasmids	pRSV-Rev	1.5 µg
	pMD2-g/prre	4.5 µg
	pMD2-VSV-G	2.5 µg

Viral packaging plasmids and the lentiviral vector DNA plasmid were added to solution B. While vortexing solution A, solution B was added dropwise, and the resulting mixture was added dropwise to the seeded cells. 24h post transfection, media was changed to fresh 2 mM Sodium butyrate-containing media. Viral supernatant was collected 48h post transfection, filtered through a 0.45 µm filter and snap-frozen using liquid nitrogen to be stored at -80^{o}C.

2.7.2 Viral transduction

Cells were seeded (5 x 10^5 in a 6-well plate) in media containing 8 µg/mL of Polybrene® (Merck). 400 µL of virus as prepared in the earlier section was added dropwise to the cells and plates were spun at 1,500 rpm for 1.5 hours at 22^{o}C and then moved to regular incubation at 37^{o}C. Cells were analysed for expression of the desired protein after 48h.

2.8 Adeno-associated virus (AAV) transduction

Purified AAV vectors harboring eGFP were purchased from the Vector and Genome Engineering Facility (VGEF) at the Children's Medical Research Institute (CMRI), Sydney. The vectors were purified with two rounds of cesium chloride (CsCl)-based gradient centrifugation and titred with eGFP-specific primers.

Cells were seeded in 24-well plates at the density of 5 x 104 cells per well and allowed to adhere. AAVs were added according to dosages mentioned in Table 2.6. Cells were incubated for 48 h post addition of AAV and subsequent flow cytometric analysis was performed.

2.8.1 Table 2.6: AAV dose by capsid for various cell lines

Cell type	Serotype	MOI (vg/cell)
HeLa	AAV2	1×10^3
	AAV8	1×10^4
eHAP	AAV2	1×10^3
	AAV8	5×10^4
HEK293T	AAV2	1×10^3
	AAV8	1×10^3

2.9 Transfection

2.9.1 Lipofectamine transfection

Cells were seeded (5×10^4 in a 24-well plate, or 5×10^5 in a 6-well plate) and allowed to adhere overnight. Plasmid DNA, P3000™ Enhancer Reagent (Invitrogen) Opti-MEM™ (Gibco, MA), and Lipofectamine 3000™ Transfection Reagent (Invitrogen) were combined in the respective ratios as mentioned below. Both solutions were mixed, vortexed and incubated at room temperature for 20 minutes before dropwise addition to the cells. Cells were further incubated at 37°C for 48 hours before experiments.

Culture vessel	Number of cells	Dilution of DNA	Lipofectamine™ 3000
24-well plate	5×10^4	25 µL Opti-MEM™ + 1 µg DNA +1.5 µL P-3000™	25 µL Opti-MEM™+ 1 µL Lipofectamine™ 3000
6-well plate	5×10^5	125 µL Opti-MEM™ + 4 µg DNA + 8 µL P-3000™	125 µL Opti-MEM™ + 5 µL Lipofectamine™ 3000

2.9.2 Turbofectin transfection

eHAP cells were seeded (5×10^4) in 24 well plates and allowed to adhere overnight. 1 µg of plasmid DNA was diluted in 50 µL Opti-MEM™ and vortexed gently. 3 µL of Turbofectin 8.0 (Origine, MD) was mixed in by pipetting and the mixture was incubated for 15 minutes. The transfection mixture was added to cells dropwise and the cells were grown at 37°C for 48 h before any analysis.

2.10 Flow cytometry

2.10.1 AAV transduction assay

Cells were washed with PBS, detached with TrypLE Express™ (Invitrogen, MA), pelleted, washed and resuspended in FACS buffer (Table 3). DAPI (Invitrogen) was added (1:1000) and fluorescence was measured (MCherry or GFP) on the BD LSR Fortessa™ flow cytometer.

2.10.2 AAVR surface expression assay

HEK293T and eHAP cells (5 x 10^5) were pelleted, washed and resuspended in chilled (4°) PBS. Cells were incubated in anti-AAVR antibody (Table 4) in 5% FCS in PBS for 1 hour at room temperature in a 1.5 mL Eppendorf™ tube. Post one wash with cold PBS, cells were incubated with secondary antibody Alexa Fluor 488 Goat Anti-Mouse (Table 4) at room temperature for 30 minutes. Post one more wash with ice cold PBS, cells were resuspended in 200µL FACS buffer (Table 3) and AF 488 expression was measured using flow cytometry (BD LSR Fortessa™).

2.11 Statistical analysis and visual representation

All experiments were performed in triplicates. All data represented as mean ± SD (standard deviation) calculated using Microsoft Excel and GraphPad Prism v.9. Unpaired student's t-test was used to determine significance (ns = not significant, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001).

2.11.1 Table 2.7: Software used for data analysis and visualization

Software	**Source**
BioRender®	BioRender
BD FACSDiva™	BD Biosciences
Fiji (Fiji is Just ImageJ)	Open source
FlowJo v10.2	FlowJo LLC
GraphPad Prism v9	GraphPad Software Inc
ImageLab™ v6.1	Bio-Rad
Leica Application Suite x (LAS X)	Leica Microsystems
Benchling	Overwatch Research Limited

3. Results

3.1 Successful knockout of *WDR11 in vitro* in HEK293T and eHAP cells

To study the impact of *WDR11* on AAV transduction, WDR11 KO (knockout) cells were created using CRISPR/Cas9 gene editing, the workflow for which is demonstrated in Fig. 3.1. A single guide RNA (sgRNA3) was designed *in silico* to target exon 4 of the *WDR11* gene. Plasmids expressing sgRNA3-mCherry and Cas9-eGFP were co-transfected into wild-type HEK293T and eHAP cells. Doubly positive cells were isolated using fluorescence activated single-cell sorting (FACS) and expanded as monoclonal populations in a 96-well plate. A plasmid containing a non-targeting scrambled sequence sgRNA (scr) was used as a CRISPR control. Cells were subjected to the same clonal selection process for use as wild-type (WT) cells in all future experiments.

Western blot was performed to screen for successful loss in WDR11 protein expression, using the scr clones as a positive control. Clones #1, #2 and #3 for both HEK293T and eHAP showed no band at 137 kDA, the expected size of the WDR11 protein, while the scr clone shows a positive band (Fig. 3.1). GAPDH was used as a loading control in the western blot to ensure uniform loading of all samples.

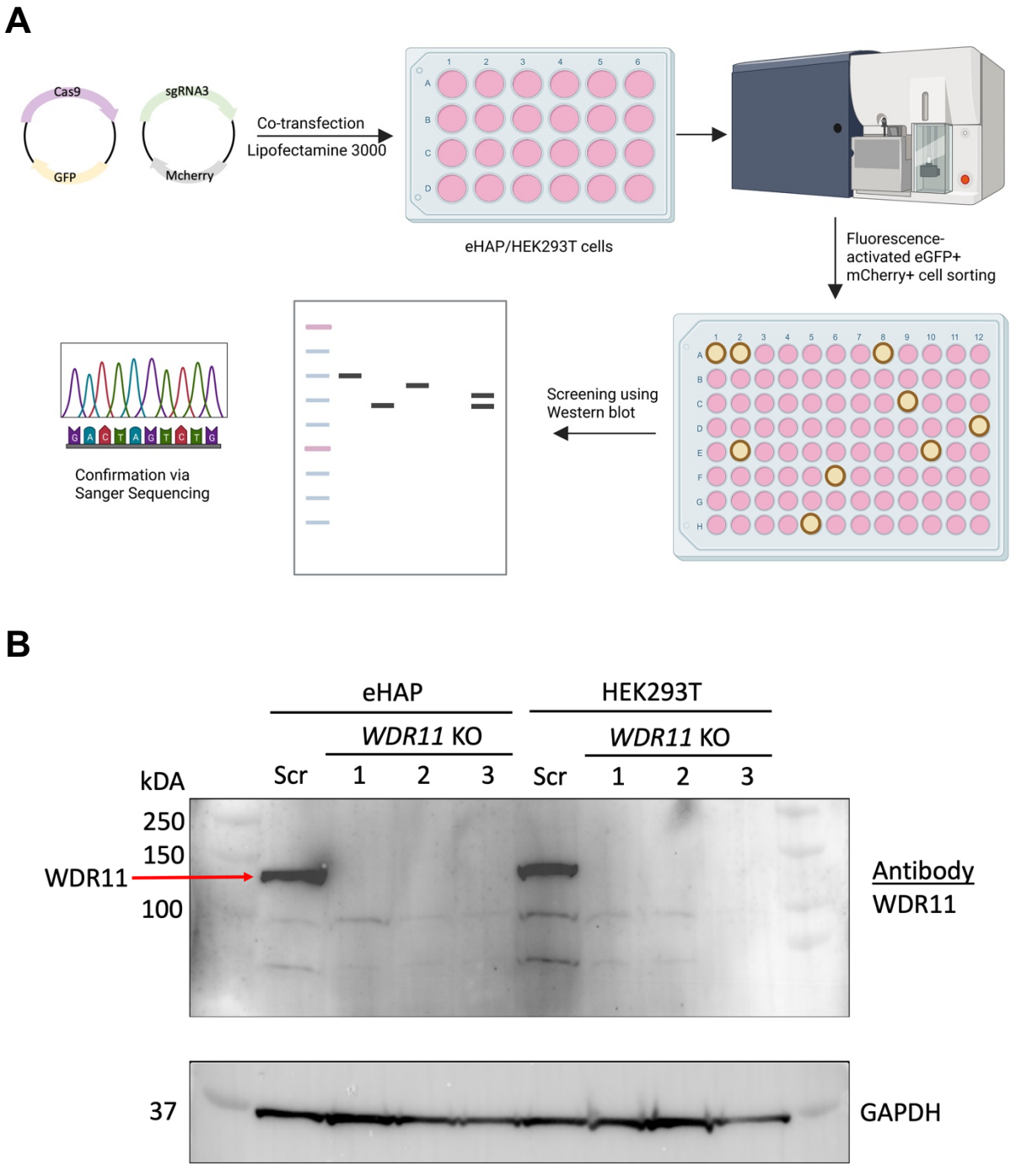

Figure 3.1: Successful knockout of *WDR11* in HEK293T and eHAP cells. (A) Workflow for CRISPR-Cas9 mediated knockout of *WDR11* in HEK293T and eHAP cells. Plasmids encoding i) GFP and Cas9 and ii) mCherry and sgRNA (sgRNA3 or scr) were co-transfected into cells. Doubly positive single cells were sorted into individual wells in 96-well plates, grown and screened for loss of WDR11 expression using western blot. Results were confirmed by Sanger sequencing for one clone each. (B) Western blot depicting successful knockout in 3 clones each in eHAP and HEK293T cell lines. Both were probed with the Abcam antibody.

Clones exhibiting loss of WDR11 expression (clone #2 for eHAP and clone #2 for HEK293T) were selected for confirmation using Sanger sequencing. Genomic DNA was isolated from the clones and a 490bp region flanking the sgRNA targeting region was amplified by PCR using a proof-reading polymerase. Blunt-ended amplicons were cloned into a TOPO cloning vector. The recombinant vectors were used transformed into chemically competent *E. coli* bacteria and 10 colonies for each cell type from the subsequent transformation were sequenced. Alignment of the cloned sequences obtained for each putative *WDR11* KO cell-line (Fig. 3.2) demonstrated successful CRISPR/Cas9 mediated gene editing.

A single base-pair insertion was observed in the eHAP *WDR11* KO cell line consistent with its haploid status (Fig 3.2). In HEK293T cells, a single base-pair deletion, a 7-base pair deletion, and a single base pair insertion was observed, which is expected due to the known polyploidy of HEK293T cells. These insertions and deletions (indels) act as frameshift mutations and lead to out-of-frame translation of the WDR11 protein, hence proving successful knockout of WDR11 in these cell lines.

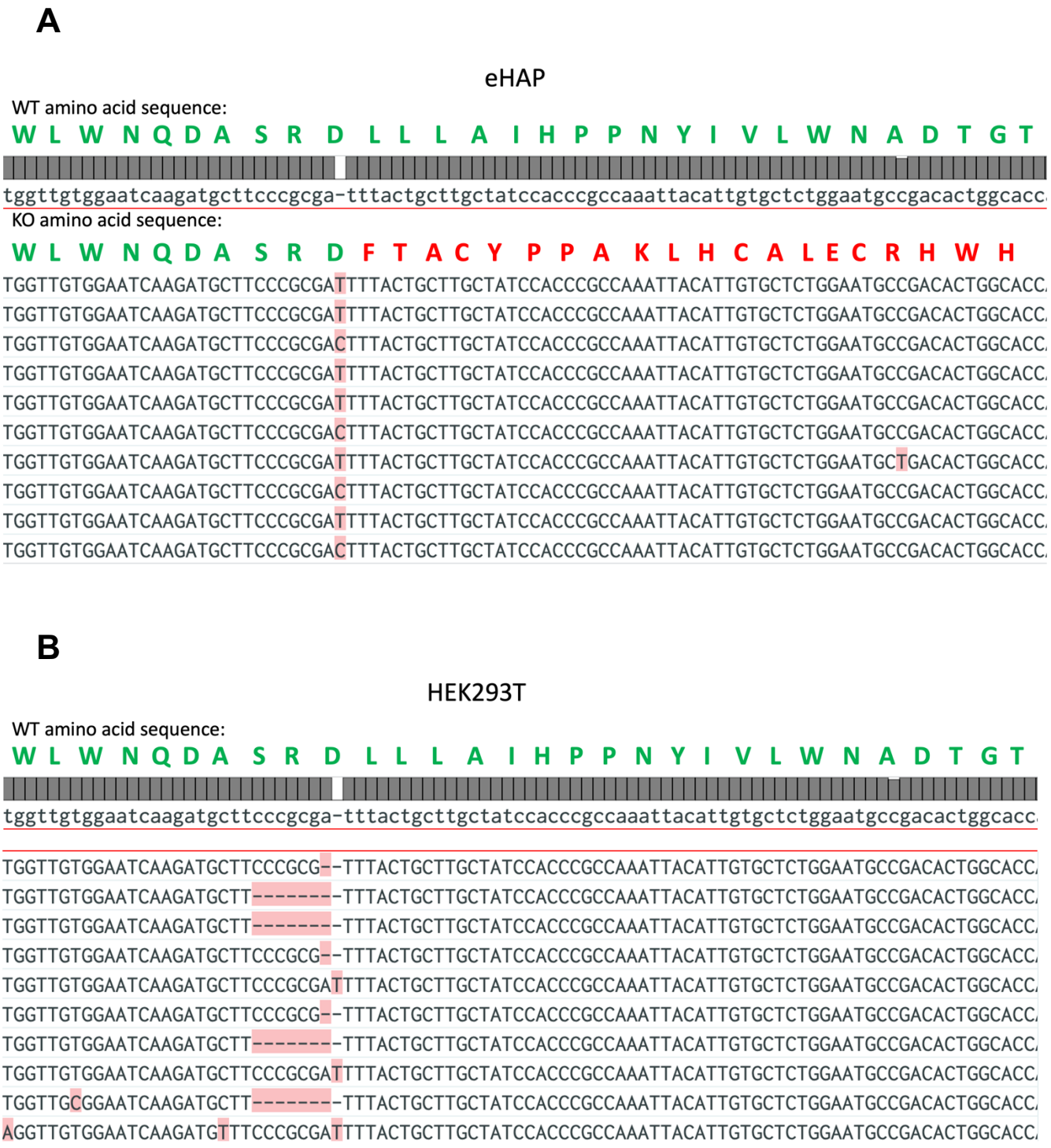

Figure 3.2: Genetic composition of *WDR11* KO alleles in eHAP and HEK239T cells confirmed by Sanger sequencing. Post screening with western blot, WDR11 KO clones from both cell lines were selected, their genomic DNA was extracted and the sgRNA targeting region was PCR-amplified. Amplicons were cloned into a TOPO cloning vector and Sanger sequenced. Sequences were aligned to the wild-type sequencing using Benchling. Green: Canonical amino acid translation, red: resultant frameshifted translation.

3.2 Successful knockout of *WDR11 in vitro* in HeLa cells

Similar to the workflow used for eHAP and HEK293T cell lines, two identical populations of wild-type HeLa cells were co-transfected with a plasmid co-expressing an sgRNA and a puromycin resistance gene and a plasmid co-expressing Cas9 and a blasticidin resistance gene. (Fig. 3.3). A plasmid containing a non-targeting (scramble) guideRNA was transfected as a CRISPR control, and these cells were used as WT cells for all future experiments. Cells were grown in double-selection media containing puromycin and blasticidin till control (untransfected) cells were dead. The selected population of cells was checked for loss of WDR11 protein expression using western blot, which showed successful knockout in the bulk population (Fig. 3.4 A).

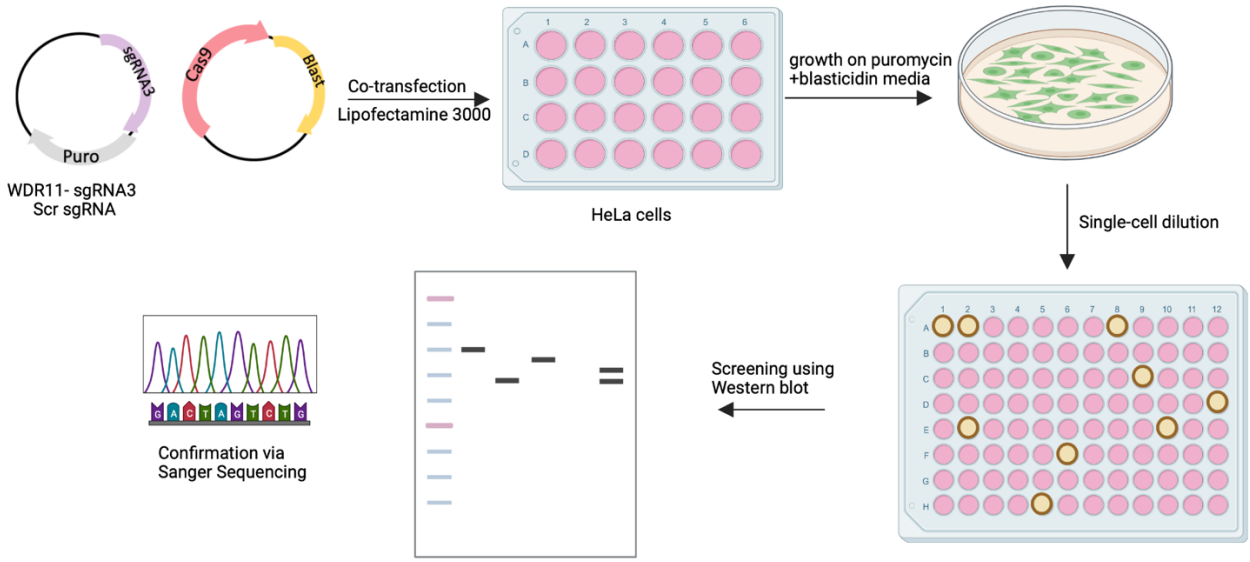

Figure 3.3: Workflow for CRISPR/Cas9-mediated knockout of *WDR11* in HeLa cells. Plasmids encoding i) Puromycin resistance and sgRNA (sgRNA3 or scr) and ii) Blasticidin resistance and Cas9 were co-transfected into cells using Lipofectamine 3000. Cells were co-selected for blasticidin and puromycin resistance, limiting cell dilution was performed into single wells in 96-well plates, and clones were grown and screened for loss of WDR11 expression using western blot. Results were confirmed by Sanger sequencing for one clone each.

The bulk populations were grown in normal media to recover from antibiotic selection and were then serially diluted to obtain monoclonal cell populations in a 96-well plate.

These single clones were expanded and screened for loss of WDR11 expression using western blot. The scrambled clones generated were used a positive control and GAPDH was used as a loading control to ensure uniform loading of samples. Clones #13, #14, #15, #17 and #18 exhibited successful *WDR11* KO (Fig. 3.4 B).

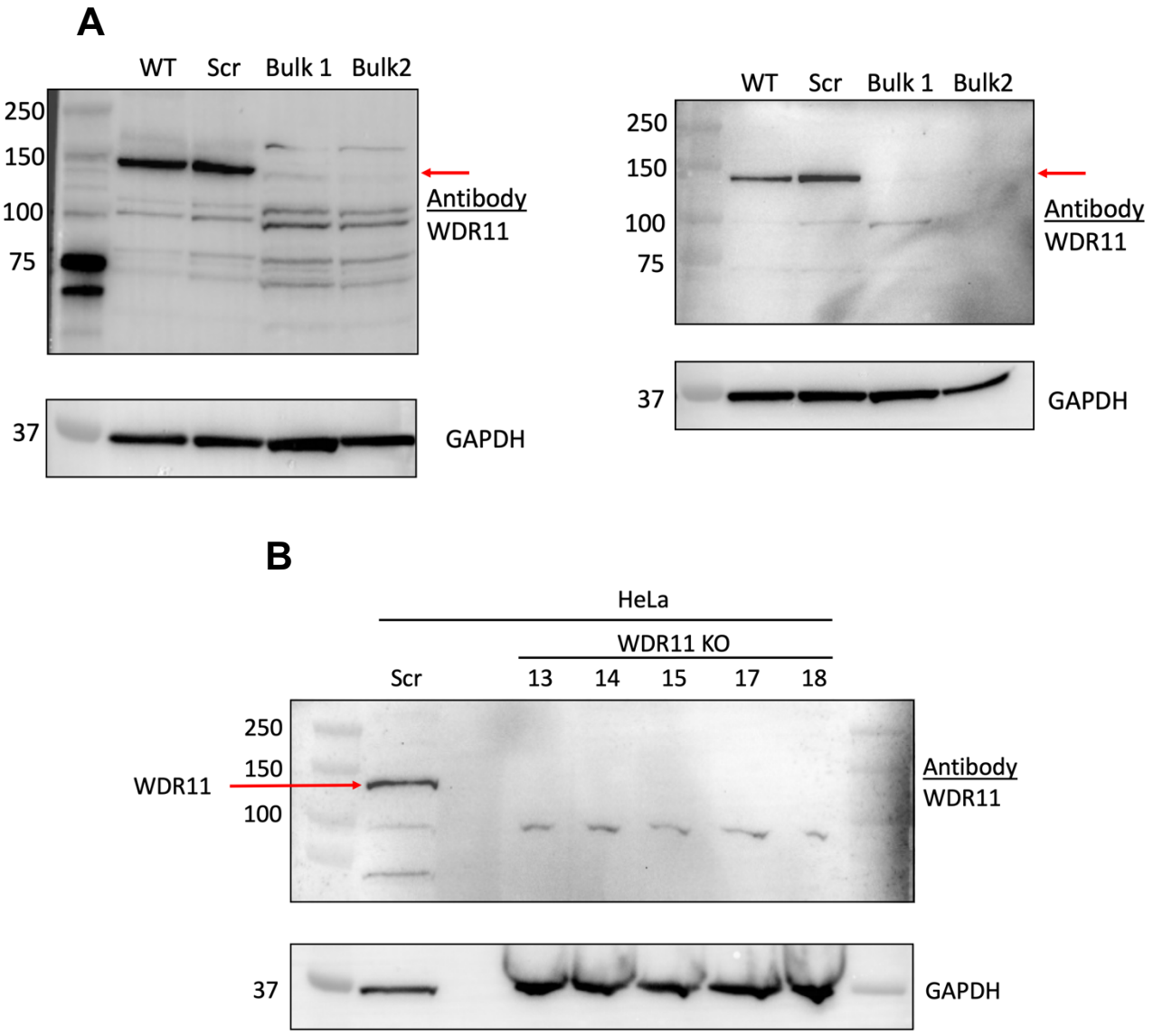

Figure 3.4: Western blot confirmation of *WDR11* KO in HeLa cells. (A) Western blot depicting successful knockout in the CRISPR-Cas9 treated bulk populations using two different antibodies for WDR11. Left: ThermoFisher, Right: Abcam. Red arrow indicates WT WDR11 expression. (B) Western blot depicting successful knockout in 5 clones in HeLa (Abcam antibody).

Genomic DNA was extracted from clone #13 and a 490-bp region flanking the sgRNA3 targeting site was amplified using PCR. The amplicons were cloned into a TOPO

cloning vector and Sanger-sequenced. A 22bp deletion and 3 single base pair deletions at different locations were observed (Fig. 3.5). HeLa is known to be a mostly triploid cell line and the different indels are indications of successful knockout of WDR11 in multiple alleles.

Figure 3.5: Confirmation of successful *WDR11* KO in HeLa cells. Post screening with western blot, a KO clone was selected, and genomic DNA was extracted. The sgRNA3-targeting region was PCR amplified and cloned into a TOPO cloning vector. 15 bacterial colonies from the resulting transformation were grown in culture, and plasmid DNA from 13 colonies grown successfully was Sanger sequenced. The sequences were aligned to the wild-type sequence using Benchling. Green: Canonical amino acid translation.

3.3 Construction of pFUW-mCherry-WDR11-HA overexpression construct

To enable exogenous expression of WDR11 via lentiviral transduction or transfection, the WDR11 open reading frame was previously cloned into the lentiviral mCherry-2A-HA expressing vector by Alex Sherman such that the WDR11 protein was tagged at the N-terminus with an HA epitope tag. With this construct co-expression of both mCherry and WDR11 was observed. Using this construct, the WDR11 insert was digested and cloned into the same vector such that the protein was HA-tagged at the C-terminus (both constructs shown in Fig 3.6). Ligation mixture was used to transform competent *E. coli,* the resulting colonies were grown, and their plasmid DNA was harvested. Insertion was verified by restriction digestion with *MscI* with expected band sizes as given in Fig. 3.6. Plasmids that showed the correct digestion pattern were verified to contain the correct insert using Sanger sequencing.

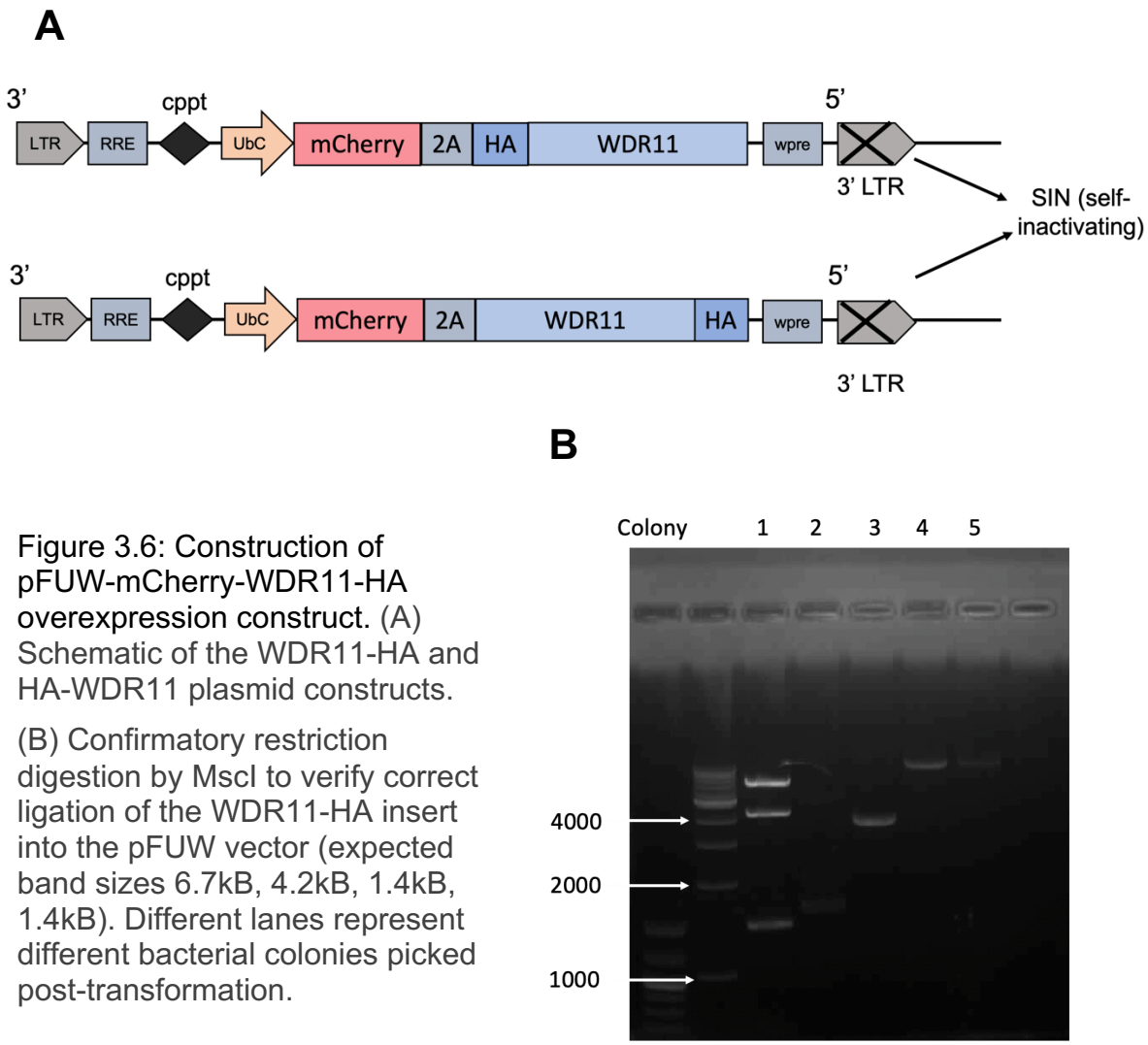

Figure 3.6: Construction of pFUW-mCherry-WDR11-HA overexpression construct. (A) Schematic of the WDR11-HA and HA-WDR11 plasmid constructs.

(B) Confirmatory restriction digestion by MscI to verify correct ligation of the WDR11-HA insert into the pFUW vector (expected band sizes 6.7kB, 4.2kB, 1.4kB, 1.4kB). Different lanes represent different bacterial colonies picked post-transformation.

Lentivirus was prepared from both the N-tagged and C-tagged WDR11 constructs, hereafter referred to as HA-WDR11 and WDR11-HA respectively. Wild-type and WDR11 KO HEK293T cells were transduced with both the lentiviruses to check for expression and detection of endogenous WDR11 and the HA tag, and untransfected cells were used as a negative control in both cases. Cells were harvested post transduction and checked for protein expression using western blot (Fig 3.7). GAPDH was used as a loading control.

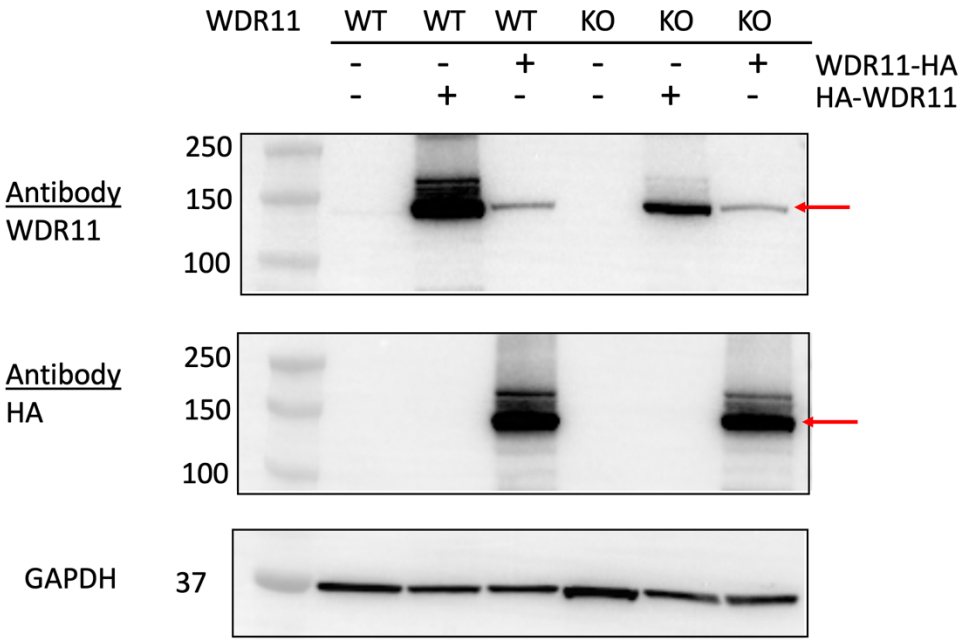

Figure 3.7: Confirmation of WDR11 overexpression in HEK293T cells. WT and WDR11 KO HEK293T cells were transduced with the pFUW-HA-WDR11 and pFUW-WDR11-HA constructs. Protein expression was checked using two separate antibodies: an antibody targeting an epitope in the C-terminus region of WDR11 (Abcam), and a fluorescent HRP-conjugated antibody targeting the HA tag. Red arrow indicates WT expression of WDR11.

As seen in Fig. 3.7, overexpressed WDR11-HA was not detected efficiently by the anti-WDR11 antibody, while the HA tag in HA-WDR11 could not be detected by the anti-HA antibody. However, endogenous as well as overexpressed WDR11 was detected by the anti-WDR11 antibody in the case of HA-WDR11. Since all following

experiments require a robust readout of WDR11 in the cell (both endogenous and exogenous), the N-terminus tagged construct pFUW-2A-HA-WDR11 was chosen as the vector for exogenous lentiviral overexpression of WDR11.

3.4 WDR11 KO *in vitro* leads to an increase in AAV transduction that is reversed by compensation

Having generated WDR11 KO cell lines as well as a construct to stably overexpress WDR11, an experiment was performed to study the effect of WDR11 KO on AAV transduction. WT and WDR11 KO cells were transduced with pFUW-mCherry-2A-HA-WDR11 and incubated for 48 hours. Cells were also transduced with an empty vector pFUW-mCherry as a control. Cells were evaluated 48 hours later for mCherry expression using flow cytometry and WDR11 protein expression using western blot. These lentivirus-transduced cells were seeded at a density of 50,000 cells per well in a 24-well plate and transduced with fixed doses of eGFP-containing AAV2 and AAV8 virions (Section 2.7). AAV transduction efficiency was measured 48 hours later using flow cytometry. Transduction efficiency was measured as the percentage of GFP-positive cells. This experiment was performed in both HeLa and eHAP cells. The flow cytometric gating strategy is highlighted in Fig. 3.8.

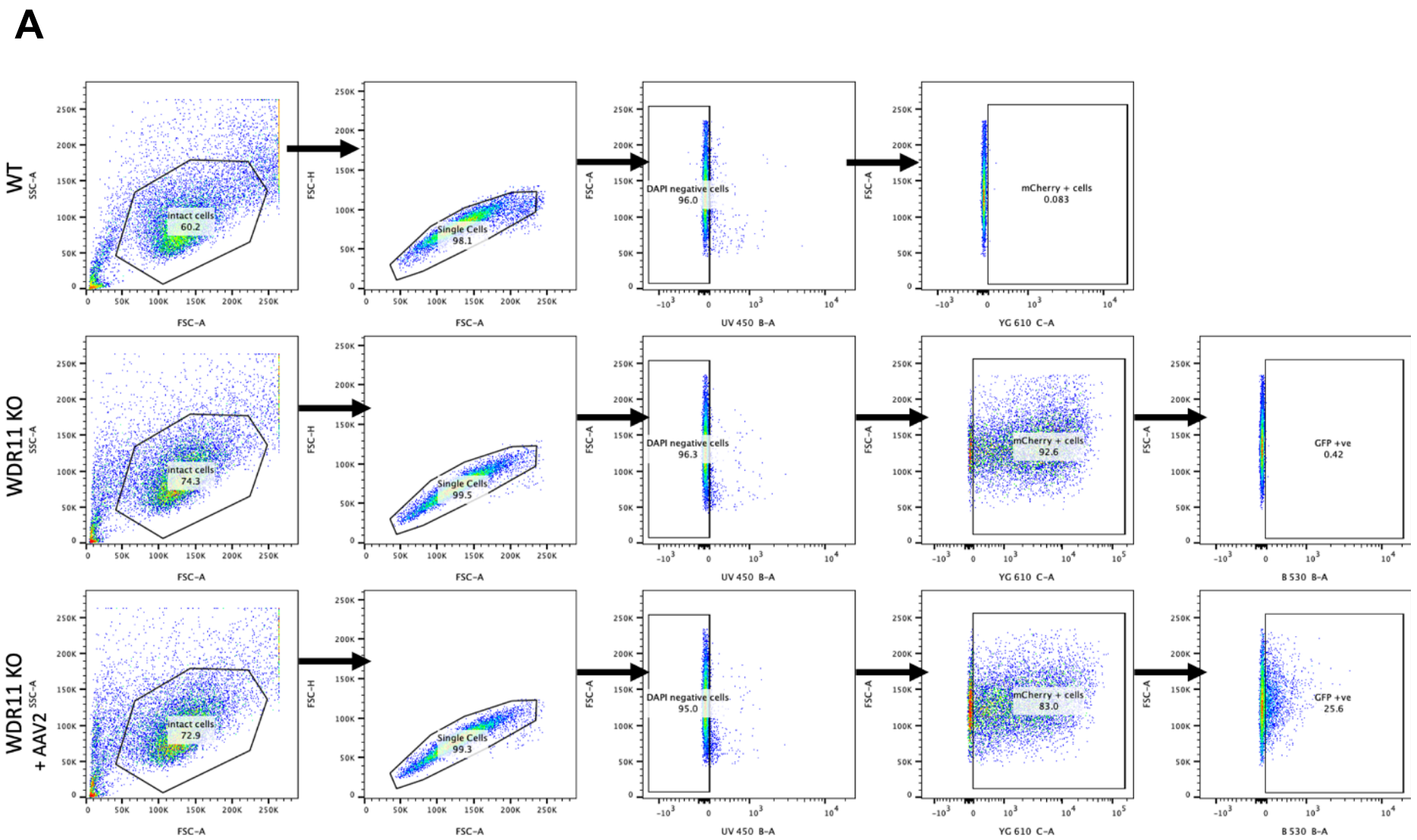

Figure 3.8: Gating strategy for flow cytometric analysis of AAV transduction in the presence and absence of WDR11. (A) Selecting for a treated population. An FSC-A (forward scatter area) vs SSC-A (side scatter area) measurement was used to identify intact cells, followed by an SSC-A vs FSC-H (forward scatter height) measurement to identify single cells and exclude doublets. DAPI staining was used to exclude dead cells. mCherry-positive cells were gated for (YG-610 channel) using untransduced WT and WDR11 KO cells as an unstained control.

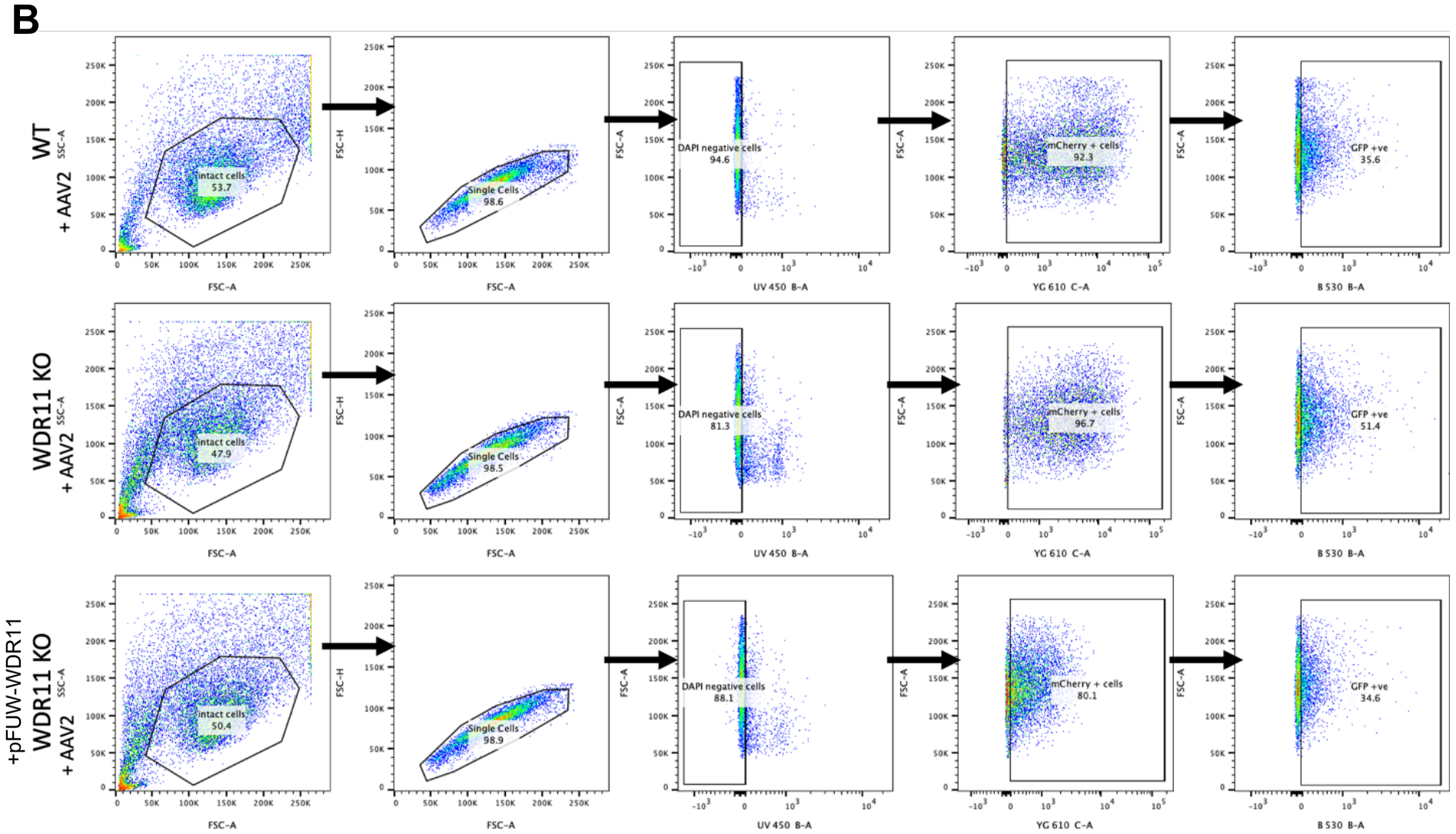

Figure 3.8: Gating strategy for flow cytometric analysis of AAV transduction. (B) AAV transduction was analyzed in WT and WDR11 KO cells treated with an empty pFUW mCherry vector (control) or the pFUW-mCherry-2A-HA-WDR11 vector (treatment). Percentage transduction was analyzed by measuring eGFP expression (B-530B channel).

Knockdown of WDR11 showed an increase in AAV transduction for both AAV2 and AAV8 in HeLa and eHAP cell lines. WDR11 KO led to a ~1.5-fold increase in AAV2 and AAV8 transduction in HeLa cells (Fig.3.9) which was statistically significant. Overexpression of WDR11 in WT HeLa cells had no significant effect on AAV transduction but led to a reversal of the observed phenotype in WDR11 KO cells, where AAV transduction showed levels similar to WT cells.

In eHAP cells, WDR11 KO caused a 3.5-fold increase in AAV2 transduction and 3-fold increase in AAV8 transduction. Lentiviral overexpression in WDR11 KO cells led to reduction of AAV transduction to WT levels, while overexpression in WT cells had no significant effect (Fig. 3.10).

Overall, loss of WDR11 leads to an increase in AAV2 and AAV8 transduction in two different cell lines. Compensating for this lack of WDR11 through exogenous overexpression reverses this effect.

We also observe that the effect on transduction is more pronounced in the case of eHAP cells than HeLa cells. An explanation for this observation is that lower transduction efficiency in WT cells possibly allows for greater fold increase in transduction in the KO cells due to a higher availability of untransduced cells.

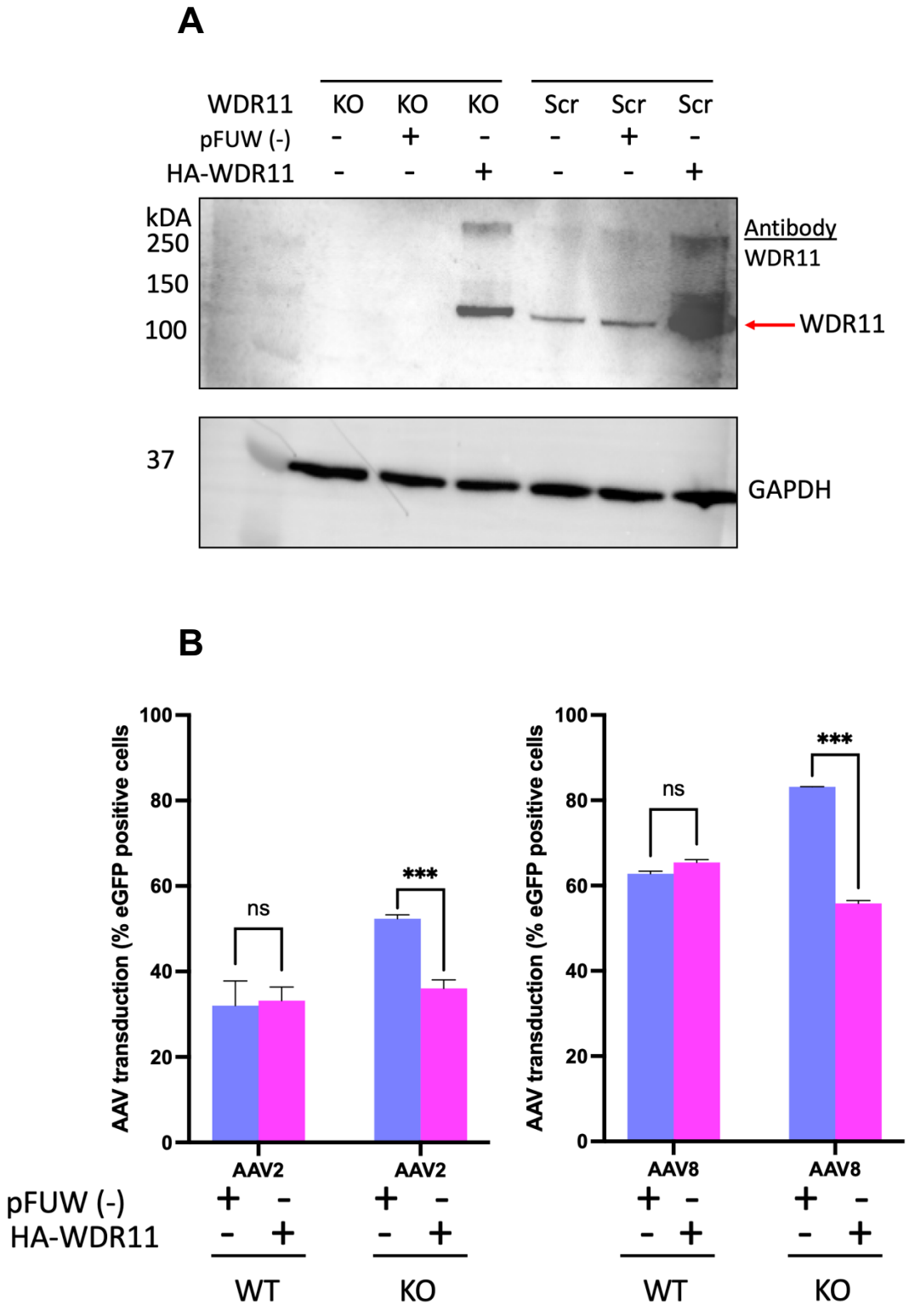

Figure 3.9: Exogenous overexpression of WDR11 via pFUW vector rescues WDR11 KO phenotype in HeLa cells. (A) Western blot showing WDR11 expression in WT (Scr) and WDR11 KO HeLa cells after transduction with lentiviral constructs. (B) WDR11 KO in HeLa cells increases AAV transduction for AAV2 and AAV8. Rescue with overexpression of WDR11 in WDR11 KO returns AAV transduction to WT levels. Data represents mean ± SD of 3 independent experiments performed in triplicate. (*p<0.05,**p<0.01, ***p<0.001,****p<0.0001)

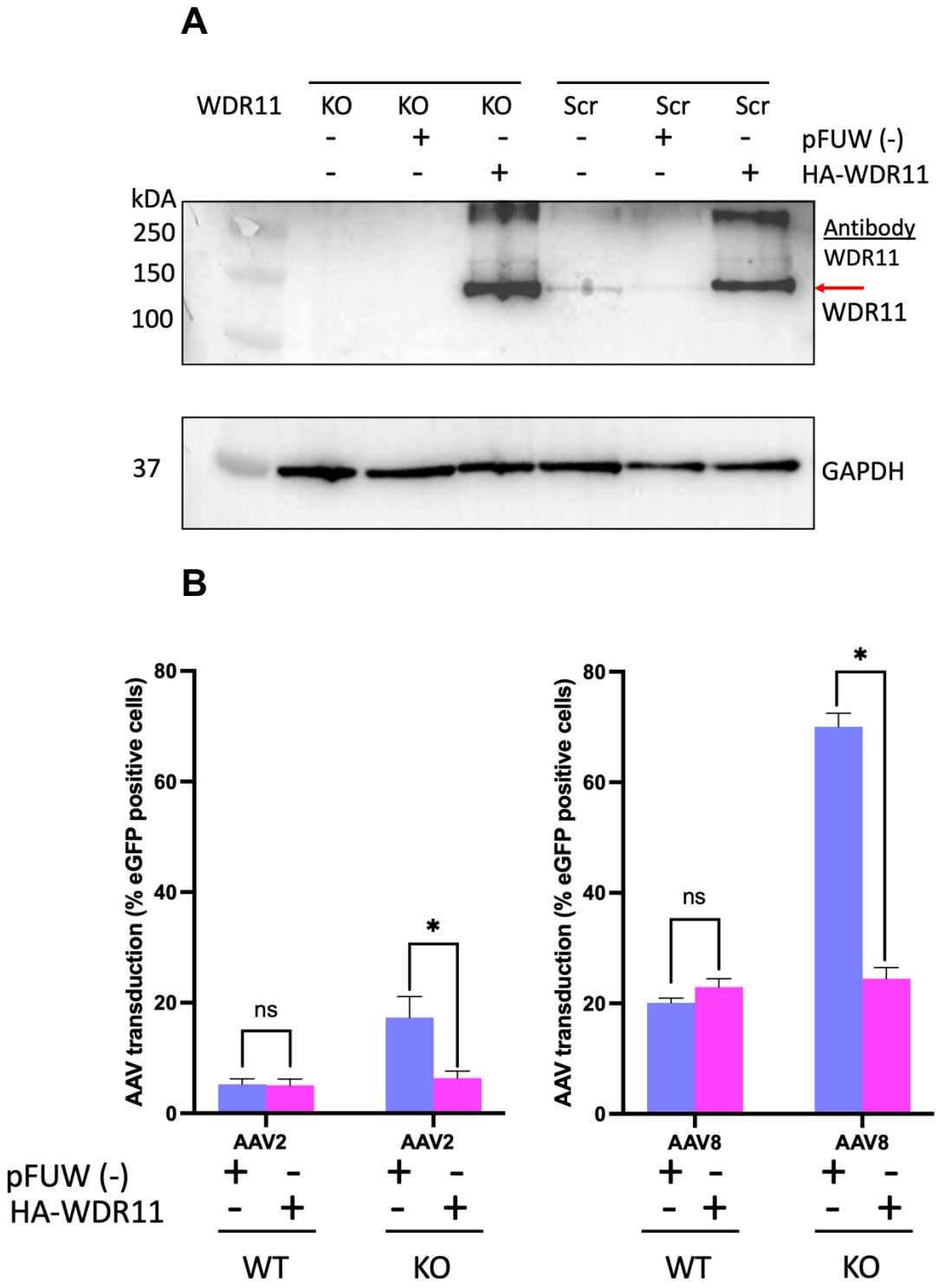

Figure 3.10: Exogenous overexpression of WDR11 via pFUW vector rescues WDR11 KO phenotype in eHAP cells. (A) Western blot showing WDR11 expression in WT (Scr) and WDR11 KO eHAP cells after transduction with lentiviral constructs. (B) WDR11 KO in eHAP cells increases AAV transduction for AAV2 and AAV8. Rescue with overexpression of WDR11 in WDR11 KO returns AAV transduction to WT levels. Data represents mean ± SD of 3 independent experiments performed in triplicate. (*p<0.05,**p<0.01, ***p<0.001,****p<0.0001)

3.5 WDR11 increases cell membrane localisation of AAVR

A study by Navarro Negredo et al. in 2018 showed increased expression of acidic cluster containing proteins at the surface of WDR11 KO cells, which includes AAVR. This finding was the basis of our interest in studying WDR11 as a potential host factor for AAV transduction. The levels of AAVR expression were first assessed in eHAP WT and WDR11 KO cells using western blot, which showed no difference in overall levels of AAVR in the cells. To observe plasma membrane levels of AAVR, eHAP WT and WDR11 KO cells were collected and probed with an anti-AAVR antibody for 1 hour, followed by a secondary incubation with an Alexa Fluor 488 (AF 488)- fluorophore conjugated secondary antibody for 30 minutes. AAVR KO cells and WT cells transduced with a lentivirus overexpressing AAVR were used as a negative and positive control respectively. Following this cell surface staining, all cells were analyzed for AF 488 expression using flow cytometry. eHAP WT cells showed a 9.7% mean AF 488 positive percentage, while WDR11 KO cells showed 80.1%, an 8-fold increase in the % AF 488 positive cells. WDR11 KO cells showed percentages similar to the AAVR-overexpressing cells (positive control) which showed 83.7% AF 488 positive cells (Fig. 3.11).

Looking at cell scatter plots and histogram plots for WT and WDR11 KO cells, we can observe a clear rightward shift in the cell distribution, which indicates a significant increase in AF 488 expression. The experiment also included a secondary antibody-only control, along with an unstained control to make sure that increased AF 488 expression correlates well with increased plasma membrane localisation of AAVR.

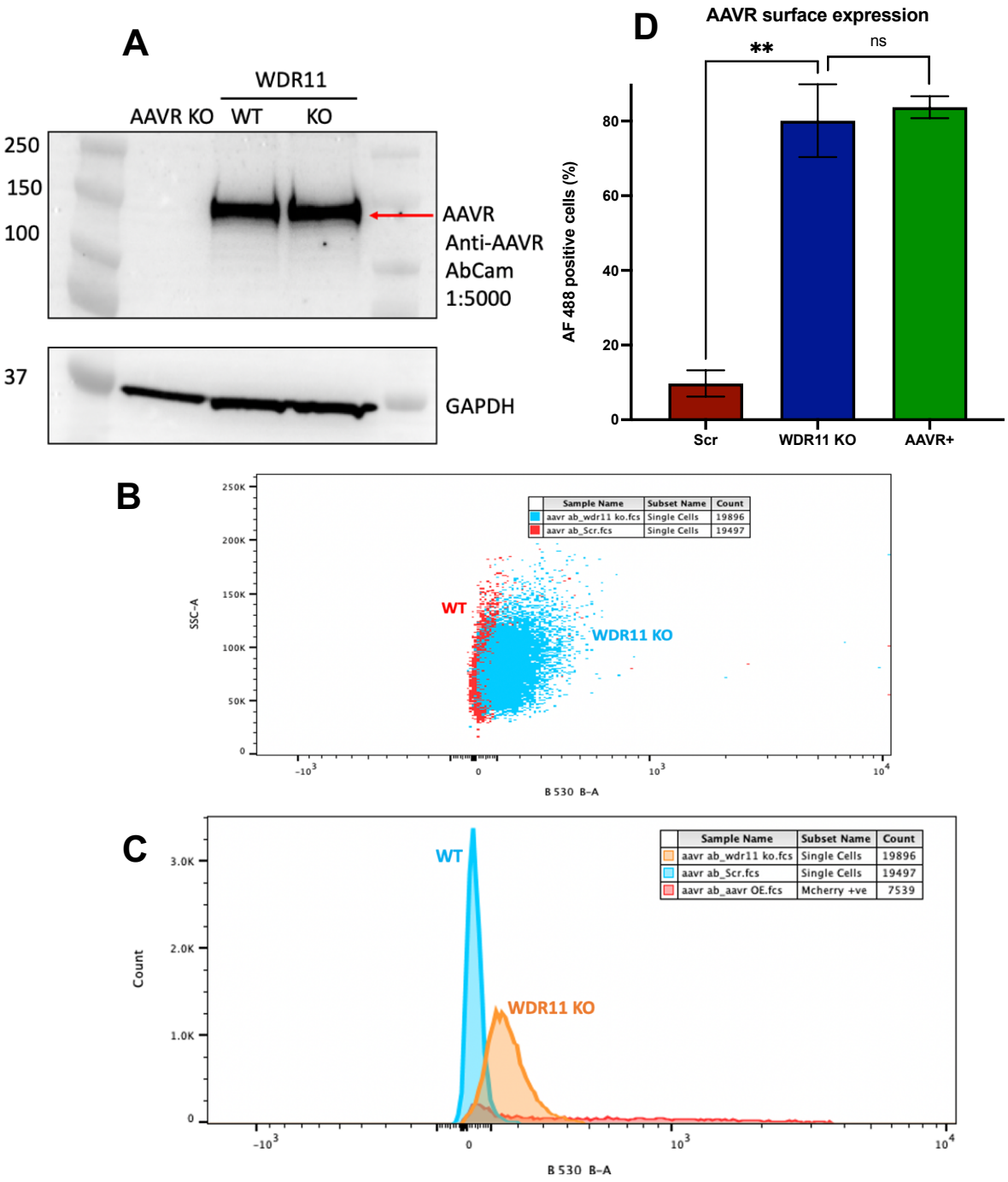

Figure 3.11: WDR11 KO eHAP cells exhibit increased expression of AAVR at the cell surface. Cell surface levels of AAVR were analysed using staining and flow cytometry. Percentage of cells showing greater Alexa Fluor 488 expression than the control (stained AAVR KO cells) were analyzed. WDR11 KO cells show a significantly higher level of AAVR expression at the cell surface as evidenced by (A) western blot showing no change in overall levels of AAVR in WT vs WDR11 KO eHAP cells, (B) cell scatter plot showing the rightward scatter of WDR11 KO cells stained for AAVR showing higher level of cell surface expression, (C) Histogram showing change in the distribution of AF-488 fluorescence and (D) Higher percentage of AF-488 positive WDR11 KO cells. Data represents mean ± SD of 3 independent experiments performed in triplicate (*p<0.05,**p<0.01, ***p<0.001,****p<0.0001).

3.5 Subcellular localisation studies using immunofluorescence

To visualise WDR11 expression and subcellular localisation a blue fluorescent protein (BFP)-tagged WDR11 protein was made. The construction of a BFP-WDR11 fusion protein was carried out using restriction cloning as described in section 2.5.6. The construct was cloned into the pcDNA3.1 vector and the correct insertion was verified using restriction digestion and sequencing (Fig. 3.12). However due to restrictions on time, this construct could not be transfected into cells to perform imaging studies.

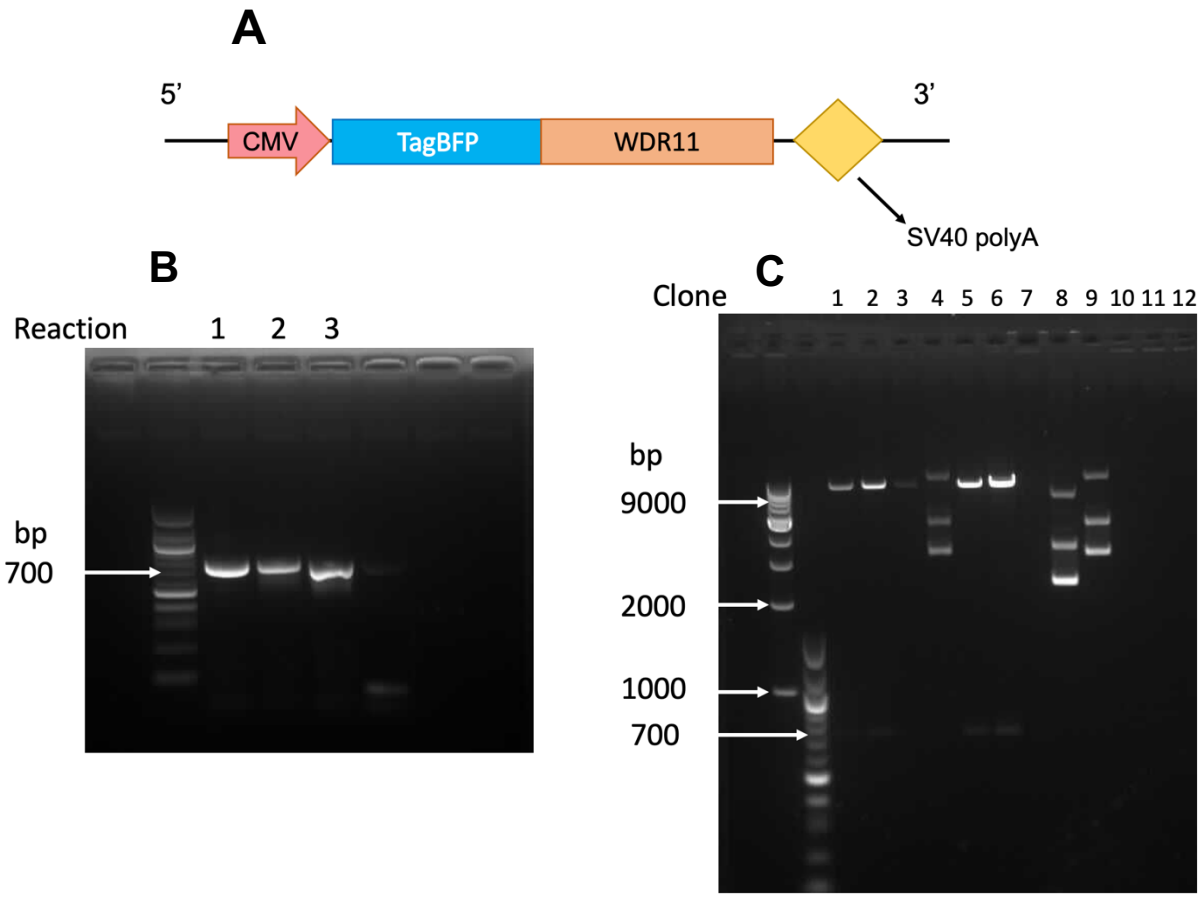

Figure 3.12: Construction of the BFP-WDR11 fusion protein. (A) a linear map of the pcDNA3.1 vector containing the fused construct. (B) Agarose gel electrophoresis images showing correct amplification of the tagBFP ORF (open reading frame). (C) Agarose gel image showing confirmatory restriction digestion of the cloned construct using *XbaI* and *AgeI.*

To visualise the subcellular localisation of AAVR in eHAP cells, both WT and WDR11 KO cells were fixed, permeabilised and stained with anti-AAVR primary antibody and an AlexaFluor 488 (AF 488) conjugated secondary antibody. Cells were also stained

with DAPI to visualise the nucleus and with an anti-GM-130 primary antibody and AlexaFluor 594 (AF 594) conjugated secondary antibody to mark the Golgi network. AAVR is shown to colocalize with the Golgi marker (Fig. 3.13) in both WT as well as WDR11 KO cells, and there does not appear to be an obvious difference in AAVR localisation upon knockout of WDR11. Since the cells have been permeabilised, differences or evidence of AAVR at the plasma membrane cannot be observed or distinguished. Further experiments using a plasma membrane marker and non-permeabilised cells might give us more information about changes in the plasma membrane levels of AAVR.

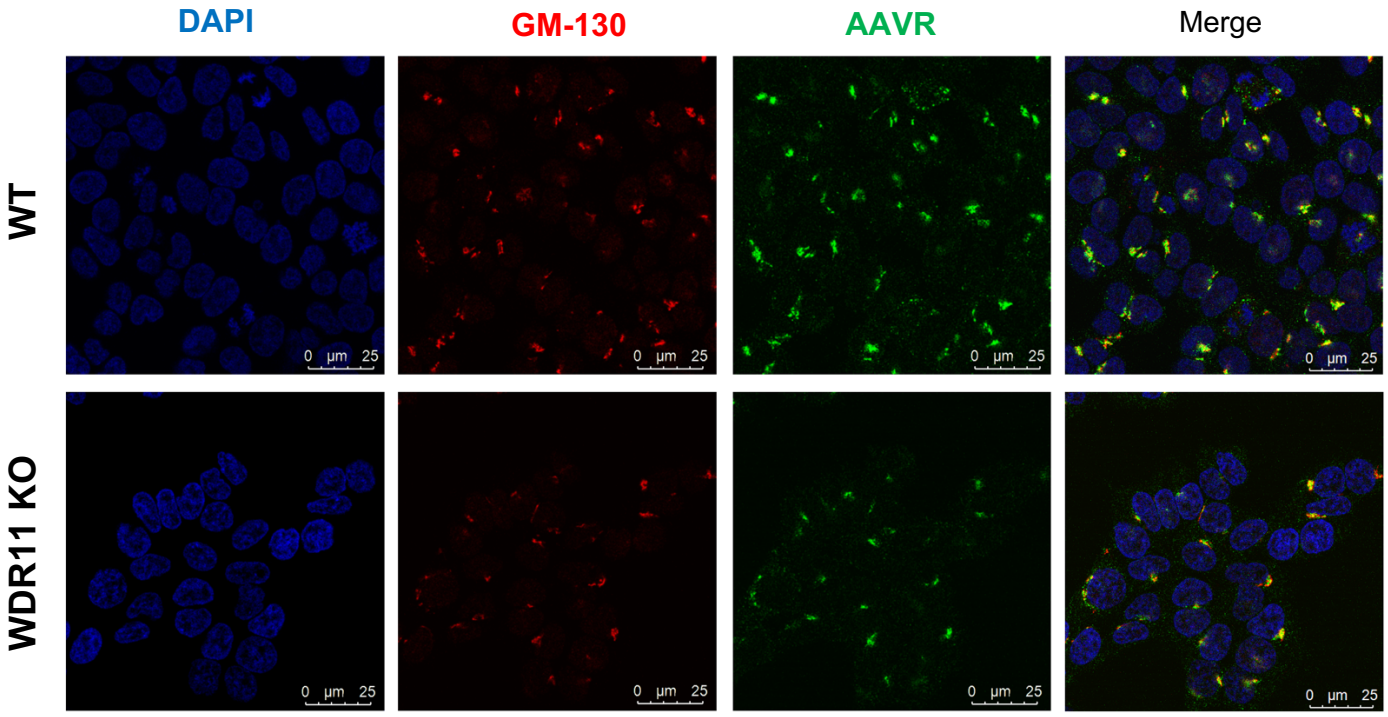

Figure 3.13: Subcellular localisation of AAVR in eHAP WT and WDR11 KO cells. Cells were stained for AAVR (AF 488, green), GM-130 to mark the Golgi (AF 594, red) and DAPI (blue). Images were acquired using the Leica TCS SP8 laser scanning confocal microscope using the 93x magnification glycerol immersion lens. Scale bar = 25 µm.

4. Discussion

4.1 The role of WDR11 in intracellular trafficking

WDR11 influences the sorting of AP-1-dependent vesicles in a stable complex with two other poorly characterised proteins- FAM91A1 and C17orf75. In the study by Navarro Negredo et al. WDR11 was also found to be more tightly constricted in its juxtanuclear location than AP-1, which points to its role being in the tail-end of the capture of vesicles targeted to the Golgi (Navarro Negredo *et al.*, 2018). WDR11 and FAM91A1 were shown to be essential for correct localisation of the WDR11 complex through knockout studies. The WDR11 complex has also been shown to associate with Golgin-245 to mediate vesicle capture. Golgin-245 also associates with multiple AP-1-dependent cargo proteins (Navarro Negredo *et al.*, 2017), which provides evidence for a model where the WDR11 complex is recruited onto AP-1-derived vesicles, which are then recruited by Golgin-245 at the TGN. The WDR11 protein contains two seven-bladed beta-propeller structures, which can aid protein-protein interactions. This provides more evidence for its role as a tethering protein. Experiments performed by Navarro Negredo et al. and Shin et al (Shin *et al.*, 2017) showed that a protein called TBC1D23 might be responsible for linking the WDR11 complex to the Golgi-tethered Golgin-245, acting as a bridge that mediates the recognition of Golgi-targeted endosomal cargo. It is possible that the WDR11 complex recognises certain rab/SNARE-like components put in place when the AP-1 vesicle is formed, which leads to its correct recruitment, or a “coincidence detection” like model (Navarro Negredo *et al.*, 2018). In such a model, wherein the association of multiple small proteins like GTPases, coat components and specific cargo proteins form a recruitment signal composed of multiple weaker interactions. With the growing evidence of WDR11’s role in guiding endosome-to-TGN trafficking, the sections after will explore its connection to AAV uptake in more detail.

4.2 WDR11 knockout leads to an increase in AAV transduction

As described in section 3.4, WDR11 knockout *in vitro* led to an increase in AAV2 and AAV8 transduction in the two different cell-types tested, eHAP and HeLa. However, the fold-change in transduction differed in both cell lines. The tissue specificity and tropism exhibited by various AAV serotypes is well documented, and the effect of WDR11 knockout needs to be observed in multiple cell types, and for all clinically relevant and commonly studied serotypes of AAV. It is noteworthy that gene knockouts involving non-essential proteins created for loss-of-function studies using CRISPR/Cas9 undergo genetic compensation that not only causes clonal variation, but may also cause discrepancies in phenotype when observed over longer periods of time (Salanga and Salanga, 2021). It would hence be beneficial to observe acute and time-dependent depletion in WDR11 protein levels and observe the change in AAVR surface expression and susceptibility to AAV transduction real time. This could be achieved by using auxin-inducible degron (AID) tags to rapidly deplete WDR11.

4.3 The increase in AAV transduction and its dependence on changes in AAVR localisation

The results obtained in this study demonstrated that WDR11 KO leads to increase in the levels of AAVR on the cell surface as measured by flow cytometry, but not as a result of increased total AAVR protein expression. The Navarro Negredo et al. study observed the same phenomenon with AAVR localization but determined using immunofluorescence. Other acidic cluster-containing proteins like CIMPR and Furin also showed increased cell-surface expression. Though the localisation of these proteins to the plasma membrane can be explained by defective sorting to the Golgi, it would be interesting to see how the intracellular localisation of these proteins is affected by knocking out other components of the WDR11 complex. Results of intracellular labelling experiments were preliminary, and there was no discernible change in the subcellular localisation of AAVR. However, using carefully designed microscopy experiments, we could potentially label different compartments of the endosomal machinery and examine any alterations in AAVR localisation in a quantifiable manner. When mis-sorting of cargo in WDR11-deficient cells was

observed in part at the electron microscope level using an antibody uptake assay (Navarro Negredo *et al.*, 2018), a significantly increased presence of acidic cluster proteins was observed in multi-vesicular bodies instead of the tubular/TGN compartments. A possible explanation for this phenomenon would be that in the absence of the WDR11 complex, endosome-to-TGN retrograde trafficking is hampered, leading to the increased localisation of AP-1 dependent cargo in late endosomes.

Through the experiments conducted as described in section 3.5, we have demonstrated that there is an increase in AAVR levels at the cell surface. However, these results have only been shown in one cell line (eHAP). Notably the fold-change in AAV transduction in eHAP cells upon genetic knockout of WDR11 was much higher than the effect seen in HeLa cells. With the background of known differences in cell-type serotype interactions and differences in host cell trafficking machinery, it would be worthwhile to perform a similar experiment in different cell types. Since we have constructed a lentiviral system to overexpress WDR11, it would also be interesting to perform a rescue experiment similar to the transduction rescue study described in section 3.4 to observe whether AAVR levels at the plasma membrane in functionally compensated WDR11 KO cells return to normal WT levels.

There are no crystal structures or structural models currently available for WDR11, and analysis of its sequence does not predict any AAV-binding domains. This might mean that the interactions between AAV and WDR11 are mediated by other proteins like AAVR. The trafficking mechanism proposed for WDR11 detailed in chapter 1 and section 4.1 involves other acidic cluster-containing proteins, as well as other AP-1-dependent cargo. Hence, global trafficking processes within the cell are likely to be affected by WDR11 KO. It would be interesting to see if other viruses or toxins that utilize retrograde trafficking pathways and are transported from the endosomes to the Golgi, like ricin (Bassik *et al.*, 2013).

Dudek et al. have identified a lineage of AAV capsids which use a AAVR-independent mechanisms for internalisation, including AAV4 and AAVrh32.33 (Dudek *et al.*, 2018). This offers the possibility of alternate trafficking and internalisation pathways for AAV, also perhaps raising the question of whether the effect seen on WDR11 knockout is limited to being mediated through increasing the plasma membrane availability of AAVR. However, AAV4 does not transduce any of the normally studied cell types,

including the ones used in this study (HEK293T, HeLa, eHAP). AAV4 does transduce the African green monkey-origin cell line Cos7, however the WDR11-ortholog expressed in these cells was not detected on western blot with the anti-human WDR11 antibody used in this study. Hence, the dependence of the observable WDR11 KO phenotype on AAVR-dependent serotypes of AAV could not be verified.

4.4 The paradoxical role of WDR11 as a restriction factor

Through various studies, WDR11 has been proven to associate with AP-1-derived vesicles and guide their sorting to the TGN with other Golgi-associated proteins. As discussed in Chapter 1, AAV trafficking to the Golgi has been proven to be an important step in its journey to the nucleus. If the retrograde trafficking pathway that is responsible for taking AAVR-bound AAV to the Golgi is hampered by the knockout of WDR11, it seems counterintuitive that this would lead to increased AAV transduction. However, in light of multiple lines of evidence which prove that WDR11 leads to an increase in the plasma membrane localisation of AAVR, the primary factor that drives the internalization of many AAV serotypes, the picture starts to become a bit clearer.

A similarly puzzling finding was reported in the case of the toxin ricin. Ricin trafficking follows the endosome-TGN-endoplasmic reticulum (ER) route, and knocking down WDR11 *in vitro* led to an increased sensitivity of the cells to ricin (Bassik *et al.*, 2013). The genetic interaction map generated also showed that knockdown of other cellular trafficking machinery, including AP-1 and retromers increased the sensitivity of cells to ricin. However, the fact that the same phenomenon was observed with other cargo that is shuttled from endosomes to the TGN adds a degree of confidence to our findings.

The aforementioned effects seen on the knockout and knockdown of trafficking proteins might be explained by compensatory transport pathways or creation of bottlenecks that cause the cargo to resort to non-canonical routes. The GARP complex has been shown to tether endosomal vesicles carrying certain cargo to the TGN (Conibear and Stevens, 2000). When components of this complex were deleted in yeast, cargo proteins were mislocalised to the vacuole. A possible hypothesis that could explain this phenomenon is that the cell maintains a homeostasis between vesicle budding and vesicle fusion, and obstructing the sorting and fusion of certain vesicles containing targeted cargo can cause the cell to compensate by allowing fusion

with non-target compartments, for example the fusion of vesicles containing acidic cluster-containing proteins with plasma membrane and endosomes instead of the target Golgi. As increased availability of surface AAVR would enable more AAV binding, the presence of compensatory pathways could allow for an increase in overall transduction, which are discussed below.

4.5 Alternative transport pathways for AAV

AAVR knockdown and rescue studies by Pillay et al. (Pillay *et al.*, 2016) have shown that the C-terminal motif that directs AAVR localisation to the TGN is essential for AAV transduction. However, deletion of this motif and compensation with other motifs that are responsible for trafficking of cargo to other non-TGN endocytic compartments were also shown to rescue AAV transduction, albeit to a lesser extent than other TGN-targeting motifs. This speaks to the possibility that AAV may be trafficked to other cellular compartments and might be able to mediate cytosolic escape through these compartments.

Though it has been shown that AAV uses an AP-1-dependent mechanism (essential for clathrin-mediated endocytosis) to undergo endocytosis, it is not known whether this is a strict requirement. As detailed in Chapter 1, alternative pathways of endocytosis include caveolin-dependent endocytosis, clathrin and caveolin-independent micropinocytosis, and clathrin-independent carriers (CLICs) in glycosylphosphatidylinositol (GPI)-anchored protein-enriched endosomal compartments (GEECs) (Duan *et al.*, 1999; Riyad and Weber, 2021). A study in ricin trafficking that expressed a dominant-negative version of dynamin (an essential component of clathrin-mediated endocytosis) in Cos7 cells did not affect ricin trafficking (Spooner and Lord, 2015). This finding could not be replicated in HeLa cells, which points to cell-type specific mechanisms for intracellular cargo processing.

Another theory that could explain the compensation of the loss of WDR11 by other pathways is that trafficking to the Golgi remains conserved, but it is carried out by other endosome-to-TGN retrograde trafficking pathways that do not utilize the WDR11 complex. Cells use a carefully regulated but redundant mechanism involving sorting proteins, tethering proteins, and SNARE proteins to efficiently direct cargo to the correct destination once they have been enclosed in vesicles. SNARE proteins facilitate fusion with the target membrane after the transport vesicles have been

captured by the tethering proteins. There are various proteins that carry out these functions, and they depend on the cargo being transported and the source and destination compartments. This diversity and redundancy of function to some extent provides various pathways for AAV to reach the Golgi. Alternative trafficking pathways to the Golgi including syntaxin 5 (Section 1.3.2) and the GARP complex mentioned in Section 4.4 are possible candidates for mediators of such an alternate transport pathway. Even in the absence of WDR11, CIMPR-containing cargo targeted to the Golgi are still captured by Golgin-245 (Navarro Negredo *et al.*, 2018). When the tethering protein TBC1D23 is knocked out in cells, all cargo capture by the Golgi is lost. This points to the possibility that TBC1D23 is a more essential component with multiple populations of endosome-to-TGN carriers acting in tandem. This theory that is supported by the binding of TBC1D23 to the WASH complex, which has been shown to act at membranes to help actin polymerisation for sorting of cellular cargo.

4.6 Conclusion

Through this study, it has been demonstrated that WDR11 plays an important role in the trafficking of adeno-associated virus (AAV), possibly through increasing the availability of uptake receptors on the cell surface. Due to the critical nature of the global endosome-to-TGN route, it is likely that there are many candidate proteins and complexes working to facilitate this transport. Experiments like RNA-seq, interaction and the use of acute depletion of WDR11 using degrons will definitively answer whether and how WDR11 contributes to AAV trafficking. In line with our original goal of dissecting cell-type specific mechanisms and pathways affecting AAV transduction, the study of WDR11 is hence a promising line of investigation that will improve our understanding of AP-1-dependant vesicle trafficking and AAV biology.

5. References

Aghajanian, H, Rurik, JG, and Epstein, JA (2022). CAR-based therapies: opportunities for immuno-medicine beyond cancer. Nat Metab 4, 163–169.

Alhakamy, NA, Curiel, DT, and Berkland, CJ (2021). The era of gene therapy: From preclinical development to clinical application. Drug Discov Today 26, 1602–1619.

Anguela, XM, and High, KA (2019). Entering the Modern Era of Gene Therapy. Annu Rev Med 70, 273–288.

Arabi, F, Mansouri, V, and Ahmadbeigi, N (2022). Gene therapy clinical trials, where do we go? An overview. Biomed Pharmacother 153, 113324.

Bassik, MC et al. (2013). A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. Cell 152, 909–922.

Bulcha, JT, Wang, Y, Ma, H, Tai, PWL, and Gao, G (2021). Viral vector platforms within the gene therapy landscape. Signal Transduct Target Ther 6, 1–24.

Chen, Q, Shen, Y, and Zheng, J (2021). A review of cystic fibrosis: Basic and clinical aspects. Anim Models Exp Med 4, 220–232.

Chernova, OB, Hunyadi, A, Malaj, E, Pan, H, Crooks, C, Roe, B, and Cowell, JK (2001). A novel member of the WD-repeat gene family, WDR11, maps to the 10q26 region and is disrupted by a chromosome translocation in human glioblastoma cells. Oncogene 20, 5378–5392.

Conibear, E, and Stevens, TH (2000). Vps52p, Vps53p, and Vps54p Form a Novel Multisubunit Complex Required for Protein Sorting at the Yeast Late Golgi. Mol Biol Cell 11, 305–323.

Cox, DBT, Platt, RJ, and Zhang, F (2015). Therapeutic genome editing: prospects and challenges. Nat Med 21, 121–131.

Cring, MR, and Sheffield, VC (2022). Gene therapy and gene correction: targets, progress, and challenges for treating human diseases. Gene Ther 29, 3–12.

Dhungel, BP, Bailey, CG, and Rasko, JEJ (2021). Journey to the Center of the Cell: Tracing the Path of AAV Transduction. Trends Mol Med 27, 172–184.

Duan, D, Li, Q, Kao, AW, Yue, Y, Pessin, JE, and Engelhardt, JF (1999). Dynamin is required for recombinant adeno-associated virus type 2 infection. J Virol 73, 10371–10376.

Dudek, AM, Pillay, S, Puschnik, AS, Nagamine, CM, Cheng, F, Qiu, J, Carette, JE, and Vandenberghe, LH (2018). An Alternate Route for Adeno-associated Virus (AAV) Entry Independent of AAV Receptor. J Virol 92, e02213-17.

Dulak, J (2021). Gene therapy. The legacy of Wacław Szybalski. Acta Biochim Pol 68, 359–375.

Ertl, HCJ (2022). Immunogenicity and toxicity of AAV gene therapy. Front Immunol 13, 975803.

George, LA et al. (2020). Long-Term Follow-Up of the First in Human Intravascular Delivery of AAV for Gene Transfer: AAV2-hFIX16 for Severe Hemophilia B. Mol Ther 28, 2073–2082.

Hoeijmakers, JH (2001). Genome maintenance mechanisms for preventing cancer. Nature 411, 366–374.

den Hollander, AI, Black, A, Bennett, J, and Cremers, FPM (2010). Lighting a candle in the dark: advances in genetics and gene therapy of recessive retinal dystrophies. J Clin Invest 120, 3042–3053.

Ibraheem, D, Elaissari, A, and Fessi, H (2014). Gene therapy and DNA delivery systems. Int J Pharm 459, 70–83.

Jain, A et al. (2017). CRISPR-Cas9-based treatment of myocilin-associated glaucoma. Proc Natl Acad Sci U S A 114, 11199–11204.

Jinek, M, Chylinski, K, Fonfara, I, Hauer, M, Doudna, JA, and Charpentier, E (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–821.

Kim, H-G et al. (2010). WDR11, a WD Protein that Interacts with Transcription Factor EMX1, Is Mutated in Idiopathic Hypogonadotropic Hypogonadism and Kallmann Syndrome. Am J Hum Genet 87, 465–479.

Kim, Y et al. (2018). WDR11-mediated Hedgehog signalling defects underlie a new ciliopathy related to Kallmann syndrome. EMBO Rep 19, 269–289.

Kimmelman, J (2003). Protection at the cutting edge: the case for central review of human gene transfer research. CMAJ Can Med Assoc J 169, 781–782.

Kumar, SR, Markusic, DM, Biswas, M, High, KA, and Herzog, RW (2016). Clinical development of gene therapy: results and lessons from recent successes. Mol Ther Methods Clin Dev 3, 16034.

Li, C, and Samulski, RJ (2020). Engineering adeno-associated virus vectors for gene therapy. Nat Rev Genet 21, 255–272.

Liu, Y, Joo, K-I, and Wang, P (2013). Endocytic processing of adeno-associated virus type 8 vectors for transduction of target cells. Gene Ther 20, 308–317.

Luo, Y (2019). Refining CRISPR-based genome and epigenome editing off-targets. Cell Biol Toxicol 35, 281–283.

Mallik, S, Bailey, CG, and Rasko, JEJ (2022). Approved gene therapies in Australia: coming to a store near you. Intern Med J 52, 1313–1321.

Martínez-Jiménez, F et al. (2020). A compendium of mutational cancer driver genes. Nat Rev Cancer 20, 555–572.

Maus, MV, Grupp, SA, Porter, DL, and June, CH (2014). Antibody-modified T cells: CARs take the front seat for hematologic malignancies. Blood 123, 2625–2635.

Mohammadinejad, R et al. (2020). In vivo gene delivery mediated by non-viral vectors for cancer therapy. J Control Release Off J Control Release Soc 325, 249–275.

Naslavsky, N, and Caplan, S (2018). The enigmatic endosome – sorting the ins and outs of endocytic trafficking. J Cell Sci 131, jcs216499.

Navarro Negredo, P, Edgar, JR, Manna, PT, Antrobus, R, and Robinson, MS (2018). The WDR11 complex facilitates the tethering of AP-1-derived vesicles. Nat Commun 9, 596.

Navarro Negredo, P, Edgar, JR, Wrobel, AG, Zaccai, NR, Antrobus, R, Owen, DJ, and Robinson, MS (2017). Contribution of the clathrin adaptor AP-1 subunit μ1 to acidic cluster protein sorting. J Cell Biol 216, 2927–2943.

Nonnenmacher, ME, Cintrat, J-C, Gillet, D, and Weber, T (2015). Syntaxin 5-Dependent Retrograde Transport to the trans-Golgi Network Is Required for Adeno-Associated Virus Transduction. J Virol 89, 1673–1687.

Park, H et al. (2019). In vivo neuronal gene editing via CRISPR–Cas9 amphiphilic nanocomplexes alleviates deficits in mouse models of Alzheimer's disease. Nat Neurosci 22, 524–528.

Pillay, S et al. (2016). An essential receptor for adeno-associated virus infection. Nature 530, 108–112.

Pillay, S et al. (2017). Adeno-associated Virus (AAV) Serotypes Have Distinctive Interactions with Domains of the Cellular AAV Receptor. J Virol 91, e00391-17.

Pipe, S, Leebeek, FWG, Ferreira, V, Sawyer, EK, and Pasi, J (2019). Clinical Considerations for Capsid Choice in the Development of Liver-Targeted AAV-Based Gene Transfer. Mol Ther - Methods Clin Dev 15, 170–178.

Pupo, A, Fernández, A, Low, SH, François, A, Suárez-Amarán, L, and Samulski, RJ (2022). AAV vectors: The Rubik's cube of human gene therapy. Mol Ther 30, 3515–3541.

Ran, FA, Hsu, PD, Wright, J, Agarwala, V, Scott, DA, and Zhang, F (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8, 2281–2308.

Reyon, D, Tsai, SQ, Khayter, C, Foden, JA, Sander, JD, and Joung, JK (2012). FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30, 460–465.

Rivière, I, and Sadelain, M (2017). Chimeric Antigen Receptors: A Cell and Gene Therapy Perspective. Mol Ther J Am Soc Gene Ther 25, 1117–1124.

Riyad, JM, and Weber, T (2021). Intracellular trafficking of adeno-associated virus (AAV) vectors: challenges and future directions. Gene Ther 28, 683–696.

Salanga, CM, and Salanga, MC (2021). Genotype to Phenotype: CRISPR Gene Editing Reveals Genetic Compensation as a Mechanism for Phenotypic Disjunction of Morphants and Mutants. Int J Mol Sci 22, 3472.

Sayed, N, Allawadhi, P, Khurana, A, Singh, V, Navik, U, Pasumarthi, SK, Khurana, I, Banothu, AK, Weiskirchen, R, and Bharani, KK (2022). Gene therapy: Comprehensive overview and therapeutic applications. Life Sci 294, 120375.

Seisenberger, G, Ried, MU, Endreß, T, Büning, H, Hallek, M, and Bräuchle, C (2001). Real-Time Single-Molecule Imaging of the Infection Pathway of an Adeno-Associated Virus. Science 294, 1929–1932.

Seo, S, Mullins, RF, Dumitrescu, AV, Bhattarai, S, Gratie, D, Wang, K, Stone, EM, Sheffield, V, and Drack, AV (2013). Subretinal Gene Therapy of Mice With Bardet-Biedl Syndrome Type 1. Invest Ophthalmol Vis Sci 54, 6118–6132.

Shin, JJH, Gillingham, AK, Begum, F, Chadwick, J, and Munro, S (2017). TBC1D23 is a bridging factor for endosomal vesicle capture by golgins at the trans-Golgi. Nat Cell Biol 19, 1424–1432.

Spooner, RA, and Lord, JM (2015). Ricin trafficking in cells. Toxins 7, 49–65.

Summerford, C, and Samulski, RJ (1998). Membrane-Associated Heparan Sulfate Proteoglycan Is a Receptor for Adeno-Associated Virus Type 2 Virions. J Virol 72, 1438–1445.

Taylor, KE, and Mossman, KL (2015). Cellular Protein WDR11 Interacts with Specific Herpes Simplex Virus Proteins at the trans-Golgi Network To Promote Virus Replication. J Virol 89, 9841–9852.

Verdera, HC, Kuranda, K, and Mingozzi, F (2020). AAV Vector Immunogenicity in Humans: A Long Journey to Successful Gene Transfer. Mol Ther J Am Soc Gene Ther 28, 723–746.

Vojta, A, Dobrinić, P, Tadić, V, Bočkor, L, Korać, P, Julg, B, Klasić, M, and Zoldoš, V (2016). Repurposing the CRISPR-Cas9 system for targeted DNA methylation. Nucleic Acids Res 44, 5615–5628.

Wang, D, Tai, PWL, and Gao, G (2019). Adeno-associated virus vector as a platform for gene therapy delivery. Nat Rev Drug Discov 18, 358–378.

Winkle, M, El-Daly, SM, Fabbri, M, and Calin, GA (2021). Noncoding RNA therapeutics — challenges and potential solutions. Nat Rev Drug Discov 20, 629–651.

Wirth, T, Parker, N, and Ylä-Herttuala, S (2013). History of gene therapy. Gene 525, 162–169.

Wu, Z, Asokan, A, and Samulski, RJ (2006). Adeno-associated Virus Serotypes: Vector Toolkit for Human Gene Therapy. Mol Ther 14, 316–327.

Xiong, W, Wu, DM, Xue, Y, Wang, SK, Chung, MJ, Ji, X, Rana, P, Zhao, SR, Mai, S, and Cepko, CL (2019). AAV *cis* -regulatory sequences are correlated with ocular toxicity. Proc Natl Acad Sci 116, 5785–5794.

Yin, H, Kanasty, RL, Eltoukhy, AA, Vegas, AJ, Dorkin, JR, and Anderson, DG (2014). Non-viral vectors for gene-based therapy. Nat Rev Genet 15, 541–555.

Zengel, J, and Carette, JE (2020). Structural and cellular biology of adeno-associated virus attachment and entry. In: Advances in Virus Research, Elsevier, 39–84.

Zhao, Z, Anselmo, AC, and Mitragotri, S (2022). Viral VECTOR-BASED gene therapies in the clinic. Bioeng Transl Med 7.

Zolotukhin, S, and Vandenberghe, LH (2022). AAV capsid design: A Goldilocks challenge. Trends Mol Med 28, 183–193.