

A CRISPR screen using BV2 cells identifies sialic acid synthesis genes as required for reovirus attachment and infection

by

Kelly Lynn Urbanek

BS Animal Biology, Delaware Valley College, 2011

Submitted to the Graduate Faculty of

Department of Human Genetics

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH
GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Kelly Lynn Urbanek

It was defended on

May 8, 2019

and approved by

Committee Member:

F. Yesim Demirci, MD
Associate Professor, Human Genetics, Graduate School of Public Health
University of Pittsburgh

Committee Member:

Amy L Hartman, PhD
Assistant Professor, Infectious Disease and Microbiology, Graduate School of Public Health
Assistant Professor, Microbiology and Molecular Genetics, School of Medicine
University of Pittsburgh

Committee Member:

John Williams, MD
Chief of the Division of Pediatric Infectious Diseases, UPMC Children's Hospital of Pittsburgh
Professor, Department of Pediatrics, School of Medicine
University of Pittsburgh

Thesis Director:

Terence Dermody, MD
Vira I. Heinz Professor and Chair, Department of Pediatrics
Professor, Microbiology and Molecular Genetics, School of Medicine, University of Pittsburgh
Physician-in-Chief and Scientific Director, UPMC Children's Hospital of Pittsburgh

Copyright © by Kelly Lynn Urbanek

2019

A CRISPR screen using BV2 cells identifies sialic acid synthesis genes as required for reovirus attachment and infection

Kelly Lynn Urbanek, MS

University of Pittsburgh, 2019

Abstract

Engagement of cell-surface receptors by viruses is a critical determinant of tropism and disease. The reovirus attachment protein, $\sigma 1$, first binds sialylated and proteinaceous receptors to mediate infection, but the specific requirements on different cell types are not entirely known. To identify host factors required for reovirus replication and subsequent cell killing, we conducted a CRISPR-knockout screen targeting over 20,000 genes in murine microglial BV2 cells. Candidate genes identified as required for reovirus-induced cell death were highly enriched for sialic acid synthesis and transport. Two of the top candidates identified, cytidine monophosphate N-acetylneuraminic acid synthetase (*Cmas*) and solute carrier family 35 member A1 (*Slc35a1*), promote sialic acid expression on the cell surface. Two reovirus strains that differ in the capacity to bind sialic acid, T3SA⁺ and T3SA⁻, were used to evaluate *Cmas* and *Slc35a1* as potential host genes required for reovirus infection. Following CRISPR-Cas9 disruption of either gene, cell-surface expression of sialic acid was diminished. These results correlated with decreased binding of T3SA⁺, a strain known to engage sialic acid, and no change in the low-level binding of T3SA⁻, a strain that does not engage sialic acid. Furthermore, infectivity of T3SA⁺ was diminished to levels of T3SA⁻ in CRISPR-modified cells. Following exogenous expression of *Cmas* and *Slc35a1* into the respective null cells, sialic acid expression was restored. These results demonstrate that *Cmas* and *Slc35a1*, which are required for cell-surface expression of sialic acid, enhance reovirus attachment. Moreover, these findings provide additional evidence that sialic

acid, which is expressed on most cells, serves as an attachment factor for reovirus. While reovirus is currently not a major public health concern, reoviruses are a highly tractable experimental model for the study of viral pathogenesis. These findings shed light on general principles of virus-receptor interactions which are important determinants of dissemination and tropism of many viruses affecting public health today.

Table of Contents

Acknowledgements	xii
1.0 Introduction.....	xii
1.1 Genetic Screening.....	2
1.1.1 RNAi.....	2
1.1.2 Haploid genetic screens with retroviral gene trapping.....	4
1.1.3 CRISPR-Cas screens.....	5
1.2 Sialic Acid (SA).....	7
1.2.1 Activation and transport of SA.....	8
1.3 Reovirus.....	10
1.3.1 <i>Mammalian orthoreoviruses</i> (reoviruses) background.....	10
1.3.2 Reovirus lab strains	13
1.3.3 SA as an attachment factor in reovirus infection.....	15
2.0 Thesis Goal	16
3.0 Significance of Research.....	17
4.0 Methods.....	18
4.1 Cell culture, viruses, and antibodies	18
4.2 CRISPR screen	19
4.3 CRISPR screen sequencing and analysis	20
4.4 STRING analysis of <i>Cmas</i> and <i>Slc35a1</i>	21
4.5 Production of <i>Cmas</i> and <i>Slc35a1</i> knockout cells	22
4.6 Flow cytometry	23

4.6.1 Preparation and staining of cells for flow cytometry.....	23
4.6.2 Bulk and single cell sorting by flow cytometry.....	23
4.7 Presto blue cell viability assay	24
4.8 Fluorescent focus unit assessment of reovirus infectivity	25
4.9 Cmas and Slc35a1 protein detection by immunoblot.....	26
4.10 <i>Cmas</i> and <i>Slc35a1</i> cDNA transfection of Δ <i>Cmas</i> and Δ <i>Slc35a1</i>	27
4.11 Statistical analysis.....	27
5.0 Results	29
5.1 CRISPR screen identifies host factors required for reovirus replication	29
5.2 STRING analysis	32
5.3 Generation of cell populations with CRISPR-targeted <i>Slc35a1</i> or <i>Cmas</i>	34
5.4 Characterization of CRISPR cell line gene disruption	34
5.5 Analysis of cell viability during reovirus infection.....	38
5.6 Quantification of reovirus infectivity of WT and BV2 CRISPR clones	39
5.7 Evaluation of JAM-A and NgR1 expression by immunoblot.....	41
5.8 Evaluation of reovirus binding to CRISPR clones	42
5.9 Complementation of <i>Cmas</i> and <i>Slc35a1</i>	42
6.0 Discussion.....	44
7.0 Summary and Future Directions.....	49
Bibliography	52

List of Tables

Table 1: Several genetic disorders result from disruption of the SA synthesis pathway.....	10
Table 2: Whole-genome CRISPR screen identifies SA synthesis genes as required for reovirus induced death in BV2 cells.	32

List of Figures

Figure 1: CRISPR-Cas mediated adaptive immunity in prokaryotes.	6
Figure 2: Schematic of the sialic acid (SA) synthesis pathway.	9
Figure 3: Electrophoretic profile of reovirus T1L, T3D, T3SA+, and T3SA- gene segments.	11
Figure 4: Reovirus structure and replication.	12
Figure 5: SA binding site on the reovirus attachment protein $\sigma 1$	14
Figure 6: A CRISPR-Cas genetic screen was conducted to identify host factors required for reovirus infection.	31
Figure 7: Mapping the interaction networks of <i>Slc35a1</i> and <i>Cmas</i>	33
Figure 8: SA expression is altered on CRISPR clones after single cell sorting.	36
Figure 9: Sialic acid expression is diminished after CRISPR knockout and restored with transient transfection.	37
Figure 10: BV2 CRISPR clones are protected from reovirus-induced cell death.	39
Figure 11: <i>Cmas</i> and <i>Slc35a1</i> are required for reovirus infection.	40
Figure 12: JAM-A and NgR1 expression is not detected on WT, $\Delta Cmas$, and $\Delta Slc35a1$ BV2 cells.	41
Figure 13: <i>Cmas</i> and <i>Slc35a1</i> are required for reovirus binding to cells.	43

List of Abbreviations

BSA	Bovine serum albumin
<i>Cmas</i>	Cytidine Monophosphate N-Acetylneuraminic Acid Synthetase
CNS	Central nervous system
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR guide RNA
DAPI	4',6-diamidino-2-phenylindole
DMEM	Dulbecco's modified eagle medium
dsRNA	Double-stranded RNA
FBS	Fetal bovine serum
FDR	False discovery rate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIBM	Hereditary inclusion body myopathy
ISVPs	Infectious subvirion particles
JAM-A	Junctional adhesion molecule A
MEL	Murine erythroleukemia
mRNA	messenger RNA
MOI	Multiplicity of infection
NgR1	Nogo-receptor 1
PBS	Phosphate-buffered Saline
PCR	Polymerase chain reaction
PFU	Plaque forming unit

RNAi	RNA interference
RT	Room temperature
SA	Sialic acid
sgRNA	short-guide RNA
shRNA	short hairpin RNA
siRNA	small interfering RNA
<i>Slc35a1</i>	Solute Carrier Family 35 Member A1
T1	Serotype 1 reovirus
T1L	Type 1 Lang
T3	Serotype 3 reovirus
T3D	Type 3 Dearing
WGA	Wheat germ agglutinin
WT	Wildtype

Acknowledgements

I am forever grateful to Dr. Terry Dermody for welcoming me into his lab as a research technician and granting me the opportunity to complete my master's degree in parallel. Terry has been unwaveringly supportive of my scientific development through my research, as well as my overall career goals. Terry is a terrific mentor. I will forever cherish the knowledge and experiences he has shared with me.

I am thankful to Dermody lab members, both past and present. The overarching kind support and caring nature of all of the lab members was greatly appreciated. I would like to thank Dr. Pavithra Aravamudhan, Pam Brigleb, Adam Brynes, Pengcheng Shang, Chris Lee, Anthony Lentscher, Jen Melvin, Nicole McAllister, Dr. Laurie Silva, and past lab members Judy Brown and Camila Guzman. Pavi for her willingness to provide helpful and thoughtful discussion; Adam for reliably being equipped with a joke to put a smile on my face; Pengcheng for always greeting me with a positive and cheerful attitude; Chris for his ability to make me laugh and aptitude to challenge me with thought-provoking questions; Anthony for his reliable support, encouragement, and senior graduate student advice; Jen for her wonderful skill as an administrator and for always providing me with the support and kind words of a caring colleague; Nicole for being a steady support system to not only my research but also my career goals; Laurie for being a dependable voice of reason and a willingness to provide me with advice and inspirational and motivational words; Judy for being such a caring and supportive individual; and Camila for her willingness to share and brainstorm interesting data. I would like to also thank Dr. Krishnan Raghunathan for his help in constructing the STRING networks presented in this thesis and his continued support and interest in my research.

I will remain forever grateful to those Dermody lab members who provided critical reviews and much valued feedback during my thesis writing process. I am truly thankful for the support and encouragement of Pam Bringleb, Adaeze Izuogu, Danica Sutherland, and Dr. Gwen Taylor. Pam for being an individual always offering positive and encouraging words; Ada for providing support, feedback, and scientific advice in such a professional and friendly manner; Danica for being my toughest critic and holding me to the highest standard all the while pushing and motivating me to keep working hard; Dr. Gwen Taylor for thinking of me in such a light as to be successful in the Dermody lab and always supporting me in pursuing my master's degree. Gwen has been a wonderful colleague and scientific mentor.

I am thankful to Rob Orchard and Craig Wilen of Skip Virgin's laboratory for their hard work in conducting the CRISPR screens and CRISPR analysis discussed in this thesis. Without their hard work, my research would not have been possible. I am also grateful to the staff of the Flow Cytometry Core at the Rangos Research Center. Josh, Mike, Alex, and Claire were instrumental in teaching me the techniques of flow cytometry as well as helping me in interpreting and troubleshooting experiments.

I would like to thank Dr. Yesim Demirci, Dr. Amy Hartman, and Dr. John Williams whom as members of my thesis committee have provided guidance, ideas, and suggestions that have helped to improve the quality of my thesis work. I have appreciated your mentorship and am so grateful to the time you were willing to dedicate to me during my graduate career.

I am thankful to the members of the Human Genetics department in the Graduate School of Public Health and the department of Pediatric Infectious Disease at the University of Pittsburgh.

I am grateful to my undergraduate scientific mentor, Dr. Kathryn Ponnock. Thank you for sparking my interest in scientific research and continuing to motivate me in pursuing my ultimate

career goal of teaching. Without the experiences and knowledge I gained from you, I would not be completing this journey. I respect you so much, cherish our friendship, and can only hope to someday be half the mentor and educator you are.

Lastly, I would like to thank my family whose commitment to my happiness and success is limitless. Without your support, guidance, and endless love I would not have completed this journey. To my mom, Rita, for always pushing me to do the best that I can, to reach for the highest goals, for teaching me to never doubt my personal strength, and instilling in me the driven and persevering attitude that I have today. To my father, Tim, for teaching me how important it is to value and stand-up for yourself, for showing me how to learn from my mistakes, how to succeed with a hard-working attitude, and for reminding me to always remember where I came from. To my step-father, Bob, and step-mother, Patty, for being supportive parental extensions that I have come to sincerely appreciate and treasure. Finally, to my husband, Ian, who is undoubtedly my biggest supporter and cheerleader. You have been next to me throughout this entire journey, always cheering me on, and I will remain eternally grateful to you. I admire the strength, love, and respect you continue to shower our relationship with. I love you so much and look forward to the road ahead as we continue our incredible journey together.

1.0 Introduction

Genome-wide screening is a technique used to identify host genes required for successful viral infection. Recently, clustered regularly interspaced short palindromic repeats (CRISPR)-Cas screening has been implemented to identify these genes. The use of this groundbreaking technique makes CRISPR-Cas screening a powerful tool to further our understanding of virus-host interactions and shed new light onto cellular pathways required for viral infection. This new knowledge has greatly expanded an understanding of infectious diseases, which aides in the development of new therapeutics and vaccines targeting these pathogenic microorganisms.

Viruses are infectious agents that replicate only inside the living cells of an organism. The first interaction between virus and cell occurs at the cell membrane, where the virus binds to attachment factors or receptors to initiate an infectious cycle. Many viruses use sialic acid (SA) as an attachment factor. Mammalian orthoreoviruses (reoviruses) engage SA using the $\sigma 1$ viral attachment protein in an initial, low-affinity binding event before engaging in high-affinity binding with specific receptors. Understanding the host genes involved in the synthesis and expression of SA, as well as how SA enhances these initial binding steps, is important to further our knowledge of the relationship between virus and host.

In this thesis, I first introduce genetic screening strategies used to investigate virus-host interactions. I discuss the role of SA, a monosaccharide found on the outer surface of the cell, in viral infection and review the current understanding of reovirus and $\sigma 1$. Next, I describe the results of experiments I conducted to evaluate the function of cytidine monophosphate N-acetylneuraminic acid synthetase (*Cmas*) and solute carrier family 35 member A1 (*Slc35a1*), two genes required for cell-surface expression of SA, in promoting efficient reovirus attachment and

infection. Collectively, these studies validate two genes from a genome-wide screen and provide additional evidence that SA serves as an attachment factor for reovirus, enhancing viral attachment and infectivity. Lastly, I discuss new questions raised by these discoveries, and explore future research directions.

1.1 Genetic Screening

Interactions between virus and host are essential for the establishment of a productive infection. Therefore, understanding the host response is equally important to understanding the pathogen when studying infection mechanisms. Numerous strategies have been used to investigate virus-host interactions. Genetic screening techniques can identify gene and protein networks that are required for successful viral replication and have proven to be valuable tools to examine virus-host interactions. Three screening technologies, RNA interference (RNAi), haploid genetic screening, and CRISPR-Cas screening, use the repression or loss-of-function of a gene as the basis for identifying host factors.

1.1.1 RNAi

Small interfering RNAs (siRNA) and short hairpin RNAs (shRNA) are short RNA molecules targeting messenger RNA (mRNA) to repress specific gene expression by induction of RNAi and translational inhibition. Many siRNA and shRNA libraries offering near-complete genome coverage are commercially available. In this screening technique, siRNAs or shRNAs are delivered into a chosen cell line. Transfection of siRNAs results in transient gene expression

knockdown, while cDNAs encoding shRNAs integrate into the genome and result in stable knockdown. Following gene expression knockdown, cells are inoculated with a virus and incubated to allow for infection. Statistical analysis of viral infection efficiency generates a list of genes that are required for viral infection.

RNAi screens have elucidated host genes required for a variety of viruses to infect cells. For example, the de novo pyrimidine synthesis pathway was identified as a host pathway required for Ebola virus genome replication and transcription using a genome-wide siRNA [5]. Additionally, using the siRNA screening technique, Nogo receptor 1 (NgR1) [6] and the chaperonin T-complex protein-1 ring complex [7] were identified as host factors required for cell entry and assembly of reovirus, respectively.

An advantage to using RNAi as a screening method is the capacity to easily target a specific gene. By identifying many infection-related host genes simultaneously, this screening technique identifies key components of protein interaction networks and signaling pathways. Additionally, this screening method can identify proviral and antiviral factors. However, incomplete knockdown of the targeted gene may lead to haploinsufficiency resulting in an incomplete and unobservable phenotype. Another drawback of the RNAi screening approach is the likelihood of off-target effects that arise from partial complementarity of the sense or antisense RNA strands to an unintended target. While flawed, using RNAi in a genome-wide screen opened the door to large-scale screening techniques. This system provided a forward-genetic approach in which specific genes could be targeted for mutagenesis to identify host genes important to viral infection.

1.1.2 Haploid genetic screens with retroviral gene trapping

Haploid genetic screening employs haploid cells and thus requires the inactivation of a single allele to elucidate loss-of-function phenotypes. In this approach, haploid cells are subjected to retroviral gene trapping in which a trap vector is designed to target a host gene. Insertional mutagenesis renders the target gene nonfunctional. After viral infection, the virus-resistant cells are assessed by deep sequencing to identify host genes necessary for this process.

The design of this screening technique most often identifies genes required for viral entry and the early steps of viral replication. For example, using KBM-7 cells, near-haploid cells with a myeloid lineage, protein proteolipid protein 2 was identified as a host factor required for Kaposi's sarcoma-associated herpesvirus infection [8]. Similarly, NPC intracellular cholesterol transporter 1 was identified as required for Ebola virus replication [9].

An advantage of using a haploid genetic screening technique is the capacity to identify both dominant and recessive genetic conditions. Mutations introduced into the haploid genome directly display the corresponding phenotype, while effects may be masked in diploid or polyploid cells due to incomplete genetic ablation. While the use of haploid cells results in the complete disruption of the targeted gene, this particular genotype limits the choice of cell types that can be used for this screening technique. As with any genetic screening technique, gene candidates arising from a haploid genetic screen must be further validated. Haploid genetic screening has provided a high-throughput genetic screening tool that can identify host factors used by pathogens.

1.1.3 CRISPR-Cas screens

CRISPR and associated proteins (Cas) has greatly expanded the toolbox used to explore virus-host interactions by recreating the prokaryotic CRISPR-Cas adaptive immune system to efficiently induce mutations in eukaryotic cells. Following prokaryote infection by a bacteriophage, foreign DNA is recognized by a set of Cas endonucleases, and short segments are integrated into the bacterial genome (Figure 1). Transcription of this archived sequence generates a CRISPR guide RNA (crRNA) that is subsequently bound by a Cas endonuclease. This crRNA-Cas complex is directed to the foreign DNA, which forms perfect complementary base-pairings with the crRNA. Upon this pairing, Cas endonuclease activity cleaves and degrades the foreign, double-stranded DNA.

The adaptation of this bacterial immune system to selectively disrupt target gene function has established CRISPR-Cas as a screening technique. To target the entire genome, Cas9-expressing cells are transduced with a library of lentiviruses encoding a multitude of short guide RNAs (sgRNAs). Each sgRNA has a specific sequence which targets a specific gene. To overcome off-target effects, or nonspecific and unintended genetic modifications, sgRNA libraries often include multiple sgRNAs targeting each gene of interest [10]. After lentiviral transduction, cells are placed under antibiotic selection and propagated to allow phenotypic maturation and decay of the target protein. The newly established library of cells are infected with a virus. Positive or negative selection approaches are used to identify gene-edited cells with a desired phenotype or that have dropped out of the population, respectively [11]. Deep sequencing and bioinformatic analyses are conducted to quantify sgRNA enrichment, which identifies host genes necessary for viral replication.

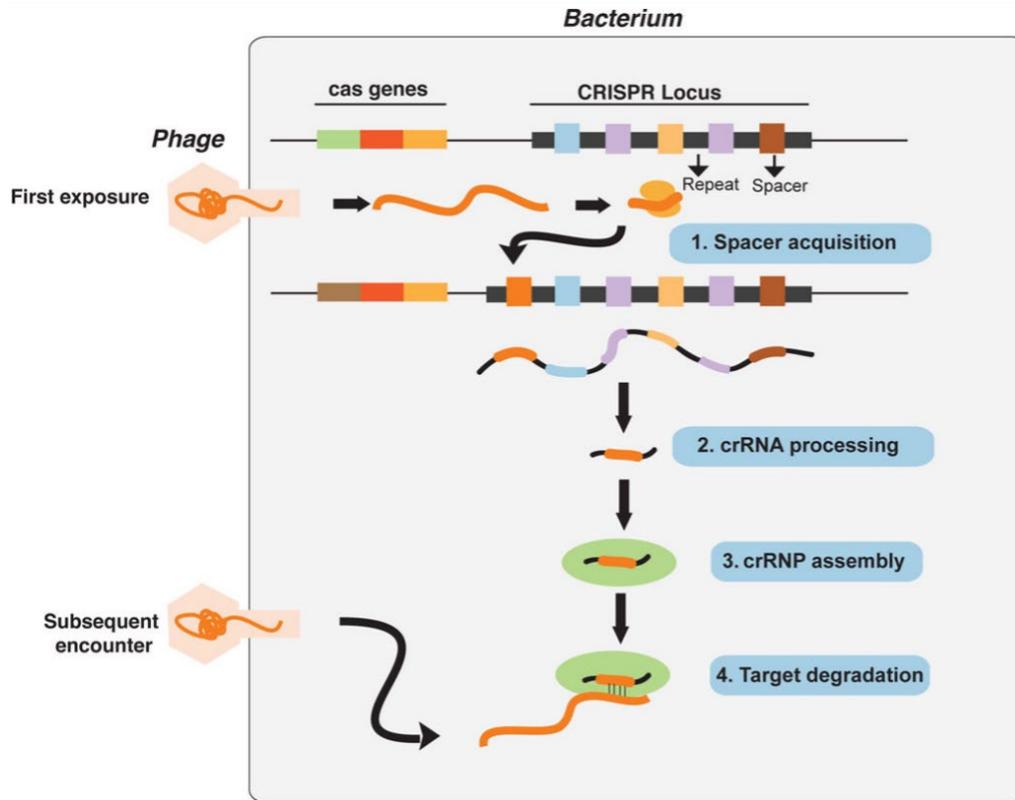


Figure 1: CRISPR-Cas mediated adaptive immunity in prokaryotes.

Following initial infection by a bacteriophage, a bacterial CRISPR locus is established within the genome. Transcription of this archived sequence produces CRISPR guide RNAs (crRNA). During a subsequent encounter, crRNA is complemented perfectly by foreign DNA which signals to Cas endonucleases to employ site-specific cleavage of the dsDNA. Figure reproduced from Kannan and Ventura [3].

A cell-adhesion molecule, *Mxra8*, was identified as an entry mediator for alphaviruses using a CRISPR screen [12]. Reduced infectivity levels of chikungunya, Ross River, Mayaro, and O'nyong nyong viruses were observed following gene editing of *Mxra8*. Further validation studies revealed the binding of *Mxra8* to surface-exposed regions of chikungunya virus, which provided a potential new target to mitigate infection and disease caused by arthritogenic alphaviruses. A cell surface glycoprotein, *CD300lf*, was discovered to be required for murine norovirus infection [13]. After further validation of the CRISPR screen results, *CD300lf* was confirmed to be a receptor for norovirus. Several ER proteins were identified as host factors required for multiple viruses of the

Flaviviridae family displaying the power of this genetic screening tool. These data also illuminated a potentially conserved mechanism in the replication of all flaviviruses.

An advantage to using the CRISPR-Cas screening technique is the high confidence in the data gathered by the screen. Introducing multiple sgRNAs targeting the same gene, in a redundant fashion, allows for the independent identification of a single gene multiple times and reduces the risk of off-target effects. Additionally, the knockout of alleles by the CRISPR-Cas mechanism often results in striking phenotypes, a greater signal-to-noise ratio, and fewer false-positives [14]. While this screening technique has proven to be powerful, off-target effects remain a drawback of CRISPR-Cas screening. In this instance, the nucleotides guiding the sgRNA-Cas complex will incorrectly bind to non-complementary DNA, resulting in disruption of an unintended gene. Interestingly, the off-target effects observed when using the CRISPR-Cas screening method are considerably less than those observed with RNAi [15]. The reproducibility of CRISPR-Cas screening and reduced risk of off-target effects makes this strategy a powerful approach to probe virus-host interactions to identify important host genes.

1.2 Sialic Acid (SA)

SAs are a highly diverse family of acidic sugars with a nine-carbon backbone found in higher vertebrates. SAs are most commonly appended to the terminating branches of underlying sugar chains bound to proteins or lipids, in which the entire molecule is referred to as a glycoprotein or glycolipid. SAs function to stabilize membranes, facilitate interactions with the environment, enhance cell-cell adhesion and signaling, and regulate receptor affinity interactions [16].

1.2.1 Activation and transport of SA

The primary functions of the SA synthesis pathway are biosynthesis of SA, sialylation of glycoproteins and glycolipids, and transport of these glycoconjugates to the cell surface. Following biosynthesis of SA in the cytosol, SA must be activated and transported to the Golgi before addition to growing glycan chains (Figure 2). The activation of SA to CMP-SA, a nucleotide donor compound, takes place in the nucleus and is catalyzed by CMP-SA synthase, an enzyme encoded by *Cmas*. Following this conversion, CMP-SA returns to the cytosol where it is delivered into the lumen of the Golgi by the CMP-SA transporter and subsequently glycosylated. This antiporter protein is encoded by *Slc35a1*. SA is often α -linked to carbohydrate chains found on glycoproteins and glycolipids. Once forming a bond with SA, these molecules are referred to as sialylated glycoproteins and glycolipids. The final product, a glycoconjugate complex, is transported to the cell membrane where it interacts with extracellular entities. Disruptions introduced into either *Cmas* or *Slc35a1* result in incomplete expression of SA on the cell surface.

Genetic disorders resulting from disruptions in the SA biosynthetic pathway are rare but lead to many complications and a shortened life expectancy (Table 1). Many of these disorders disrupt neuromuscular processes. Hereditary inclusion body myopathy (HIBM) and sialuria are disorders caused by mutations of the *GNE* gene, which encodes an enzyme involved in SA synthesis [17]. Young adults with HIBM display symptoms typical of a neuromuscular disorder, such as muscle weakness and atrophy [18]. Infants affected with sialuria are born with jaundice, hepatosplenomegaly, and microcytic anemia [19]. Salla disease, also referred to as SA storage disease, results from a disruption of the *Slc17a5* gene and primarily affects the nervous system.

Infants born with Salla disease have poor muscle tone and progressive neurological deterioration in the first year of life [20].

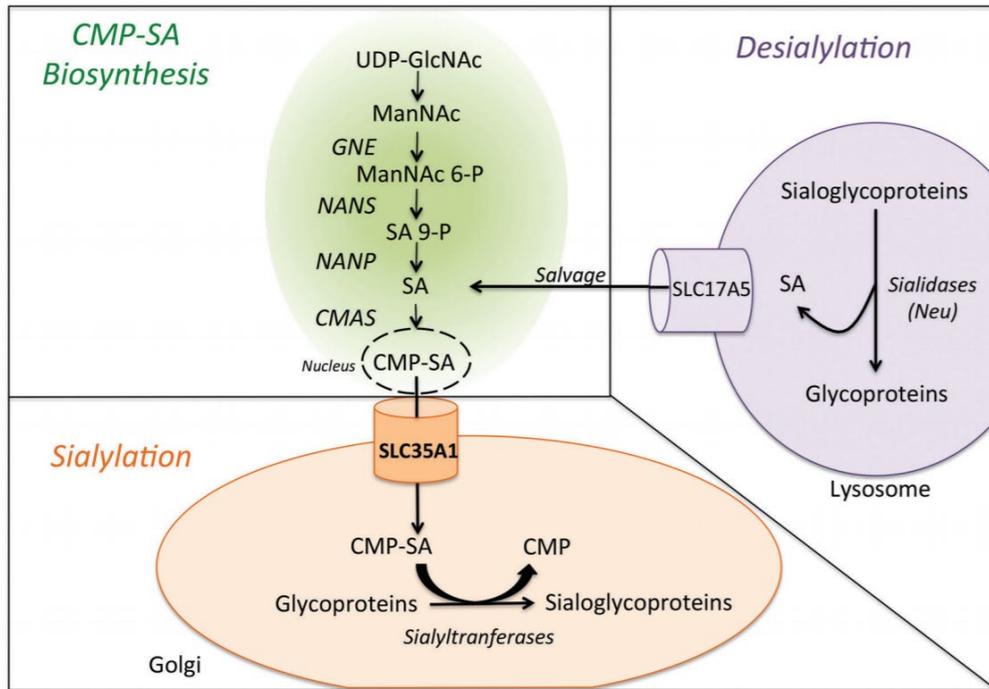


Figure 2: Schematic of the sialic acid (SA) synthesis pathway.

The majority of the initial steps in the SA synthesis pathway take place in the cytoplasm. The conversion of SA to CMP-SA, which is catalyzed by *CMAS*, occurs in the nucleus. CMP-SA enters the Golgi through the active transporter *SLC35a1*. The final step in sialylation is the addition of activated SA to growing side chains which is catalyzed by specific sialyltransferases. Figure reproduced from Kauskot et al. [1].

Currently, there is no cure and no way to prevent the progression of HIBM, sialuria, or Salla disease. While patients with these rare genetic disorders are encouraged to be frequently evaluated by a multidisciplinary team of healthcare professionals (including genetic counselors, neurologists, and therapists) lifelong mobility and neurological challenges often result from each of these disorders. In an effort to increase our knowledge about SA and related genetic disorders, it is important to continue studying these diverse molecules that are essential for vertebrate life.

Table 1: Several genetic disorders result from disruption of the SA synthesis pathway.

Genetic disorder	Gene affected	General classification	OMIM identification number
Salla disease	SLC17A5	Free sialic acid storage disorder	604369
Intermediate salla disease	SLC17A5	Free sialic acid storage disorder	269920
Infantile free sialic acid storage disease (ISSD)	SLC17A5	Free sialic acid storage disorder	269920
Sialuria	GNE	Elevated free sialic acid	269921
Sialidosis	NEU1	Progressive storage of sialic acid	256550
Congenital disorder of glycosylation, type IA (CDG1A)	PMM2	Sialic acid deficiency	212065
Congenital disorder of glycosylation, type II f (CDG2F)	SLC35A1	Abnormal glycosylation	603585

1.3 Reovirus

1.3.1 Mammalian orthoreoviruses (reoviruses) background

Reoviruses are nonenveloped, double-stranded RNA (dsRNA) viruses. The reovirus genome is comprised of ten segments of dsRNA which are encapsulated within two concentric protein shells, the outer capsid and inner core. The gene segments are grouped by size into three large (L), three medium (M), and four small (S) gene segments (Figure 3). The majority of the outer capsid is composed of hexameric complexes of $\mu 1$ and $\sigma 3$, which is disrupted by a hollow

cylindrical structure composed of five $\lambda 2$ molecules. The filamentous viral attachment protein, $\sigma 1$, embeds within this structure. (Figure 4A,B). The reovirus S1 gene encodes the $\sigma 1$ protein, which is comprised of three structurally different domains, a tail domain that is anchored to the virion, a central body domain, and a head domain.

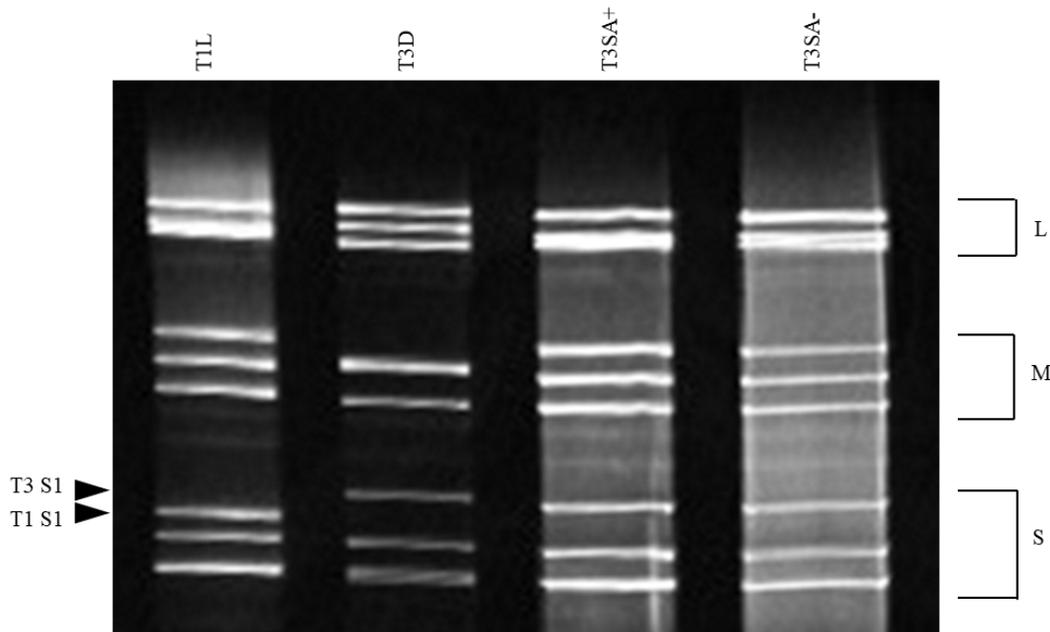


Figure 3: Electrophoretic profile of reovirus T1L, T3D, T3SA+, and T3SA- gene segments. Purified virions were electrophoresed in an SDS-polyacrylamide gel, followed by ethidium bromide staining to visualize viral gene segments. Size classes of gene segments (L, M, S) are indicated. T3SA- and T3SA+ viruses contain nine gene segments from T1L and the S1 gene segment from T3 reoviruses, T3C44 or T3C44-MA, respectively.

Reovirus infection is initiated by the attachment of $\sigma 1$ to cell-surface sialylated carbohydrates and to the proteinaceous receptors junctional adhesion molecule-A (JAM-A) [21] or NgR1 [6] (Figure 4C). Following receptor-mediated endocytosis, outer-capsid proteins $\sigma 3$ and $\mu 1$ are cleaved to yield infectious subvirion particles (ISVPs) [22, 23]. Degradation of these proteins facilitates endocytosis and release of transcriptionally active viral cores into the cytoplasm [24-26]. These cores synthesize message-sense RNAs, which serve as templates for translation

and minus-strand synthesis to generate genomic dsRNA [27-30]. After new particle assembly, reovirus progeny are released from infected cells to complete the viral lifecycle.

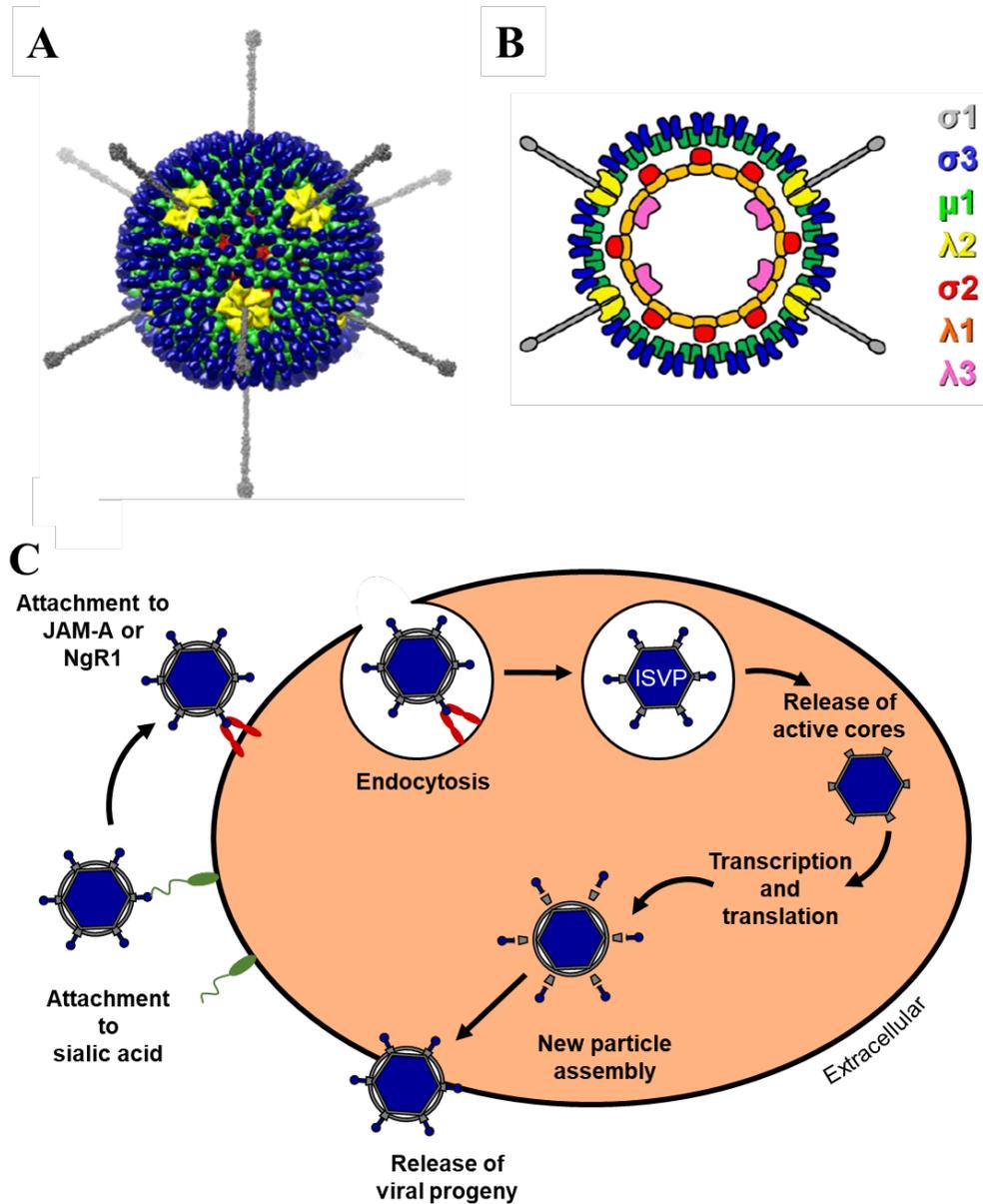


Figure 4: Reovirus structure and replication.

(A) The reovirus virion shown to scale with a model of the extended conformed of $\sigma 1$ (grey) or shown as a schematic (B). (C) A simplified schematic of reovirus replication. Reovirus infection begins with virus attachment to cell-surface sialic acids and the proteinaceous receptors JAM-A or NgR1. Following receptor-mediated endocytosis, the outer capsid proteins are removed, forming infectious subviral particles (ISVPs). Transcriptionally active viral cores are released into the cytoplasm where transcription and translation occur. After new particle assembly, the newly synthesized reovirus progeny are released from the infected cell. (A) and (B) are reproduced from Functions of the Viral Attachment Protein in Reovirus Neurovirulence, Dissertation, Danica M. Sutherland [2]. Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311[4].

1.3.2 Reovirus lab strains

There are three reovirus serotypes, serotype 1 (T1), serotype 2 (T2), and serotype 3 (T3), classified based on neutralizing antibody responses and hemagglutination-inhibition activity [31-34]. Prototypic strains T1 Lang (T1L), T2 Jones (T2J), and T3 Dearing (T3D) were isolated from the stools of children during an active outbreak of gastroenteritis [35]. The majority of reovirus research has focused on T1 and T3 strains, as these are genetically more divergent from one another [36, 37]. As demonstrated by the crystal structures of T1L and T3D $\sigma 1$ proteins, the SA-binding domain of these two viruses differ [38]. The SA-binding site is located in the $\sigma 1$ head domain of T1L and the $\sigma 1$ tail domain of T3D (Figure 5A,B), resulting in differential binding to SAs [39]. T1L reovirus binds to a specific $\alpha 2,3$ -linked *N*-acetyl-SA [38] while T3D reovirus binds terminal $\alpha 2,3$ -, $\alpha 2,6$ -, or $\alpha 2,8$ -linked SAs [40].

To elucidate the function of SA during reovirus infection, T3SA⁺ and T3SA⁻, two viruses that differ in the capacity to engage SA, were recovered and characterized [41]. Murine erythroleukemia (MEL) cells are normally not very susceptible to reovirus infection. MEL cells were infected with T3C44, a T3 reovirus field-isolate strain incapable of binding SA. During serial passage, T3C44 mutants were selected that had gained the capacity to bind SA and infect MEL cells. This SA binding variant (T3C44-MA) was sequenced and a single point mutation was identified (Leu²⁰⁴→Pro). To circumvent the potential problem of additional mutations in other genes that may have been selected during serial passage, cells were coinfecting with these viruses and T1L to generate reassortants. The resulting viruses, containing nine gene segments of T1L and

the S1 gene segment of either T3DC44 or T3DC44-MA, were named T3SA- and T3SA+, respectively (Figure 5C, D).

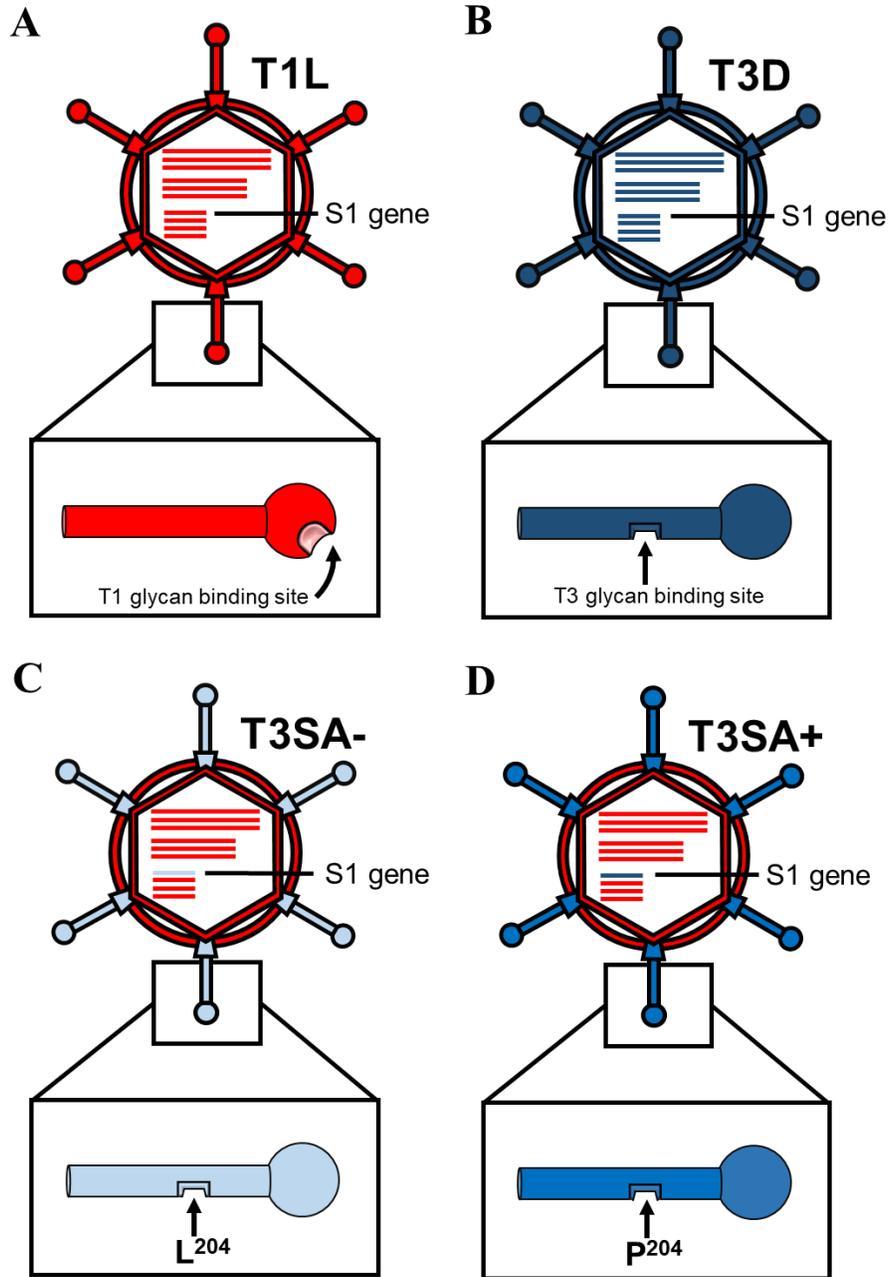


Figure 5: SA binding site on the reovirus attachment protein $\sigma 1$.

The S1 gene encodes $\sigma 1$. (A) The reovirus field isolate, T1L. The SA binding site of T1L is located in the head domain of $\sigma 1$. (B) The reovirus field isolate, T3D. The SA binding site of T3D is located in the tail domain of $\sigma 1$. (C) The reovirus laboratory strain, T3SA-. This virus contains nine gene segments from T1L (red) and the S1 gene segment (blue) from a T3 field isolate, T3C44, which is incapable of binding sialic acid (L²⁰⁴). (D) The reovirus laboratory strain, T3SA+. This virus contains nine gene segments from T1L (red) and the S1 gene segment (blue) from sialic acid binding variant T3C44-MA (P²⁰⁴).

A plasmid based reverse genetics system [42] was used to recover T3SA+ and T3SA- [43]. In comparison to a forward genetic approach, which relies on pre-existing genome alterations to identify genes responsible for a phenotype, reverse genetics takes a more direct and efficient approach to identify the function of mutations. Using targeted mutagenesis, changes to the viral genome are introduced at specific sites creating a tailored virus with engineered variations. The viruses isolated using this system are often less likely to acquire additional mutations during passaging and are therefore used to study functions of specific reovirus gene products.

1.3.3 SA as an attachment factor in reovirus infection

Glycans do not directly mediate reovirus entry. Instead, viruses transiently bind SA with low affinity until a higher-affinity receptor, like JAM-A or NgR1, is encountered and cell entry is initiated [41]. In this way, SA strengthens cell adhesion. *In vitro*, the interaction between reovirus and SA is required for efficient reovirus binding and infection of many cell types. Only glycan-binding reovirus strains are able to infect murine embryonic fibroblasts (MEFs) [38, 44], while SA is dispensable for reovirus binding and infection of L929 fibroblasts [38, 40, 44]. *In vivo*, reovirus pathogenesis is enhanced by SA engagement [45]. Glycan-binding T1 and T3 reovirus strains induce more severe disease in the central nervous system (CNS) than their glycan-blind counterpart virus strains [46] [43]. Additionally, oily hair syndrome results from T3SA+ infection of the bile duct epithelium but is not observed following infection with T3SA- [45]. Viral dissemination from the intestine to sites of secondary replication is also enhanced by reovirus strains capable of efficiently binding glycans [45]. Defining host genes in the SA synthesis pathway will further our understanding of the function of SA engagement in reovirus attachment and infection of different cell types.

2.0 Thesis Goal

The focus of my thesis work is to identify and validate host genes required for reovirus replication. Two genes, *Cmas* and *Slc35a1*, were identified in a CRISPR screen in mouse microglial cells as required for reovirus-induced cell death. During my thesis research, I sought to understand the function of these two genes during reovirus infection. The first specific aim of my project was to evaluate SA expression on the surface of $\Delta Cmas$ and $\Delta Slc35a1$ CRISPR clones and complemented cell lines. The second specific aim of my project was to determine whether *Cmas* and *Slc35a1* are required for reovirus binding and infection.

3.0 Significance of Research

Viral receptors and attachment factors are important determinants of dissemination and tropism during reovirus-induced disease. Work presented in this thesis focuses on two genes, *Cmas* and *Slc35a1*, whose functions contribute to SA expression on the cell surface. The expression of SA mediates enhanced reovirus infection of microglial cells. Understanding the molecular interactions that occur between reovirus attachment protein $\sigma 1$ and cell-surface moieties, such as SA, will advance knowledge of reovirus pathogenesis and contribute to general principles of pathogen-SA engagement. My research elucidates two host genes that make microglial cells susceptible to reovirus infection and expands the current understanding of the receptors on microglial cells that are targeted by reovirus. These studies could help to identify cell-surface molecules that are mediating infection of neural cells by T3 reoviruses, thus expanding the knowledge of reovirus neurotropism.

4.0 Methods

4.1 Cell culture, viruses, and antibodies

BV2 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented to contain 10% fetal bovine serum (FBS), 1% HEPES, and 1% Penicillin/Streptomycin (referred to as BV2 Maintenance Medium). Puromycin (2.5 μ g/mL; Sigma Aldrich) and blasticidin (4 μ g/mL; ThermoFisher Scientific) were added to the medium as appropriate (see below). When both puromycin and blasticidin were added, the medium is referred to as BV2 Selection Medium.

Parental (WT), CRISPR-edited parental ($\Delta Cmas$ Parent and $\Delta Slc35a1$ Parent), CRISPR-edited bulk-sorted ($\Delta Cmas$ Bulk and $\Delta Slc35a1$ Bulk), and CRISPR-edited single-cell sorted ($\Delta Cmas$ and $\Delta Slc35a1$) BV2 cells (where the Δ signifies disruption of either the *Slc35a1* or *Cmas* gene) were cultured in BV2 Selection Medium unless otherwise noted.

Reovirus strains, T3SA⁺ and T3SA⁻, were recovered using plasmid-based reverse genetics as previously described [43]. T3SA⁻ differs from strain T3SA⁺ by a single point mutation in the S1 gene (encodes Leu204 in T3SA⁻ σ 1 and Pro204 in T3SA⁺ σ 1). Virus was purified from infected L929 cells by cesium chloride gradient centrifugation [47], and viral plaque forming unit (PFU) titers were determined by plaque assay [48]. Monolayers of spinner-adapted L929 fibroblast cells were adsorbed in duplicate with serial 10-fold dilutions of virus and incubated at room temperature (RT) for 1 hour (h). Cells were immediately overlaid and fed (3 days post-inoculation) with a 1:1 (vol/vol) mixture of 2% Bacto-Agar (Fisher Scientific) dissolved in deionized water and 2X199 medium (Caisson Labs) supplemented to contain 10% FBS, 1% penicillin, 1% streptomycin, and 1% L-glutamine. A final overlay supplemented to contain 0.04% neutral red (Fisher Scientific)

was added 6 days post-inoculation and incubated overnight. Plaques were enumerated, and viral titers are reported as plaque-forming units per mL of original sample.

For virus binding assays, virus particle number was estimated by spectral absorbance at 260 nm ($1 \text{ OD}_{260} = 2.1 \times 10^{12}$ particles/mL). Reovirus virions were labeled with succinimidyl-ester Alexa Fluor™ 488 (A30005; ThermoFisher Scientific) to produce fluoresceintated particles [49].

Goat polyclonal JAM-A-specific antibody (AF1077, R&D Systems) and goat polyclonal NgR1-specific antibody (AF1208, R&D Systems) were used at 1:1,000 dilutions in immunoblot assays. Reovirus polyclonal serum collected from rabbits immunized and boosted with reovirus strain T1L or T3D were mixed 1:1 (vol:vol) and cross-adsorbed on WT BV2 cells to deplete non-specific antibodies, and used at a 1:1,000 dilution for FFU assays. All primary and secondary antibodies were diluted in 1X TBS supplemented to contain 0.01% Tween-20 (TBS-T; Sigma-Aldrich).

4.2 CRISPR screen

The CRISPR screen, sequencing, and analysis was conducted by Rob Orchard and Craig Wilen in the laboratory of Skip Virgin at Washington University School of Medicine. BV2 cells were transduced with pXPR_101 lentivirus encoding Cas9 (Addgene; 52962) and propagated for 11 days with BV2 Maintenance Medium supplemented to contain blasticidin. These parental BV2 or BV2-Cas9 cells were transduced for 2 days with pXPR_011 expressing eGFP (Addgene; 59702) and a short guide RNA (sgRNA) targeting eGFP at a multiplicity of infection (MOI) of less than 1 PFU/cell. Cells were selected for 5 days with BV2 Selection Medium. The frequency of eGFP-expressing cells was quantified by flow cytometry.

The murine Asiago sgRNA CRISPR library contains six independent genome-wide pools, in which each pool contains unique sgRNAs targeting 20,077 mouse genes. Four pools of the Asiago library were transduced into 5×10^7 BV2 cells at an MOI of 0.2 PFU/cell to establish four BV2 libraries. Two days post transduction, cells were transferred to BV2 Selection Medium and propagated for 5 additional days. For each experimental condition, 10^7 BV2 library cells expressing Cas9 and sgRNAs, were seeded in duplicate into T175 tissue culture flasks (Greiner Bio-One). Cells were inoculated with Opti-MEM supplemented to contain PBS^{-/-} (mock) or reovirus strains T3SA+ or T3SA- at an MOI of 100 PFU/cell. Cells were incubated at RT for 1 h, followed by the addition of 20 mL of DMEM supplemented to contain 10% FBS, 1% penicillin/streptomycin, 1% sodium pyruvate, and 1% sodium bicarbonate. After 9 (T3SA+ or T3SA- conditions) or 2 days post-inoculation (mock condition), cells were harvested and genomic DNA (gDNA) was isolated from surviving cells using a QIAmp DNA Mini Kit (QIAGEN) according to the manufacturer's instructions.

4.3 CRISPR screen sequencing and analysis

Illumina sequencing and STARS analyses were conducted as previously described [50]. The gDNA was aliquoted into a 96-well plate (Greiner Bio-One) with up to 10 µg gDNA in a 50 µL total volume per well. A polymerase chain reaction (PCR) master mix containing ExTaq DNA polymerase (Clontech), ExTaq buffer (Clontech), dNTPs, P5 stagger primer, and water was prepared. PCR master mix (40 µL) and 10 µL of a barcoded primer were added to each well containing gDNA. Samples were amplified using the following protocol: 95°C for 1 min, followed by 28 cycles of 94°C for 50 s, 52.5°C for 30 s, and 72°C for 30 s, and ending with a final 72°C

extension for 10 min. PCR product was purified using Agencourt AMPure XP SPRI beads (Beckman Coulter) according to the manufacturer's instructions. Samples were sequenced using a HiSeq 2000 (Illumina).

Following deconvolution of the barcodes in the P7 primer, sgRNA sequences were mapped to a reference file of sgRNAs from the Asiago library. To account for the varying number of reads per condition, read counts per sgRNA were normalized to 10^7 total reads per sample. Normalized values were then log-2 transformed. sgRNAs that were not detected were arbitrarily assigned a read count of 1. sgRNA frequencies were analyzed using STARS software to produce a rank-ordered score for each gene. This score correlated with the sgRNA candidates that were above 10% of the total sequenced sgRNAs. Genes scoring above this threshold in either of the two independent subpools and in at least two of the four independent genome-wide pools were assigned a STAR score. In addition to the STAR score, screen results were compared using False Discovery Rates (FDRs) analyses to monitor gene-specific signal versus background noise. Statistical values of independent replicates were averaged.

4.4 STRING analysis of *Cmas* and *Slc35a1*

The STRING database was used to construct a gene interaction network for *Cmas* and *Slc35a1* based on *Mus musculus* as the model organism. All active interaction sources were used to produce a list of genes. Gene interaction scores greater than 0.4 and thresholds of ten primary interactions and five secondary interactions were used to generate the network. The interaction networks generated were functionally enriched for genes in the SA metabolic pathway. The q-values, which correspond to the likelihood that each identified gene is required for reovirus

infection, were obtained from the analysis of the CRISPR screen using T3SA+. Not all genes in the interaction network were identified in the CRISPR screen. Genes with q-values of less than 0.01 were considered to be required for T3SA+ infection. Genes that were identified in the CRISPR screen and had a q-value greater than 0.01 were considered to be unnecessary for T3SA+ infection.

4.5 Production of *Cmas* and *Slc35a1* knockout cells

CRISPR-edited parental WT, $\Delta Cmas$ B11, $\Delta Cmas$ B12, $\Delta Slc35a1$ B7, $\Delta Slc35a1$ B8 and $\Delta Slc35a1$ B9 BV2 cells were established by members of Skip Virgin's laboratory. BV2-WT cells expressing Cas9 were transduced with one of five lentiviruses expressing different sgRNAs:

Cmas sgRNA B11: 5'-CACCGGCAACTTTCTGGAGGTCAGT-3'

Cmas sgRNA B12: 5'-CACCGGCGCTGGTGCTGGCCCGCGG-3'

Slc35a1 sgRNA B7: 5'-CACCGCCTTGTGTATCTTAAAGCTA-3'

Slc35a1 sgRNA B8: 5'-CACCGTATCACTTCTGTGATACACA-3'

Slc35a1 sgRNA B9: 5'-CACCGGTATGCTGTGCAGGAACAACA-3'

Cells with disrupted *Cmas* or *Slc35a1* genes were preferentially selected and subsequently maintained in BV2 Selection Medium.

4.6 Flow cytometry

4.6.1 Preparation and staining of cells for flow cytometry

BV2 cells were detached from tissue-culture plates using CellStripper Dissociation Reagent (Corning) and quenched with double the volume of BV2 Selection Medium. Cells were pelleted after quenching and all subsequent steps at 1500 rpm at 4°C for 5 min. Cells were washed twice with PBS^{-/-} and re-pelleted.

For lectin-binding studies, cells were resuspended in PBS^{-/-} supplemented to contain 0.005mg/mL fluorescein labeled wheat germ agglutinin (Vector Laboratories; WGA). For reovirus-binding studies, cells were resuspended in PBS^{-/-} supplemented to contain 10⁵ particles/cell of fluoresceintated T3SA+ or T3SA-. Cells treated with WGA, reovirus, or PBS^{-/-} were incubated while rotating at 4°C for 1 h. Cells were re-pelleted, washed 2X with PBS^{-/-} to remove any unbound lectin or virus, and fixed in PBS^{-/-} supplemented to contain 1% paraformaldehyde. Propidium iodide (1 µL/sample) was added to all samples except the unstained control. Cells were analyzed using a LSRII flow cytometer (BD Biosciences) and sorted using a FACS Aria flow cytometer (BD Biosciences). Results were quantified using FlowJo V10 software.

4.6.2 Bulk and single cell sorting by flow cytometry

$\Delta CmasB11$, $\Delta CmasB12$, $\Delta Slc35a1B7$, $\Delta Slc35a1B8$, and $\Delta Slc35a1B9$ BV2 cells were detached, washed, incubated with WGA, and prepared for analysis following the steps previously described (4.6.1). Parental cell populations binding the least amount of lectin, $\Delta CmasB11$ ($\Delta CmasParentB11$) and $\Delta Slc35a1B8$ ($\Delta Slc35a1ParentB8$), were selected for further experiments.

$\Delta Cmas$ ParentB11 and $\Delta Slc35a1$ ParentB8 cells were detached, washed, incubated with WGA, and prepared for analysis following the steps described (4.6.1). Cells were isolated using a FACSAria flow cytometer. A total of 10,000 $\Delta Cmas$ ParentB11 or $\Delta Slc35a1$ ParentB8 BV2 cells displaying the least amount of lectin binding were selected and seeded into wells of a 6-well tissue culture plate containing 2 mL of BV2 Maintenance Medium. These bulk sorted populations were referred to as $\Delta Cmas$ Bulk mixed populations 1-3 and $\Delta Slc35a1$ Bulk mixed populations 1-4. Following propagation of the bulk-sorted populations, lectin binding was assessed as described (4.6.1).

Single cells from $\Delta Cmas$ Bulk mixed population 3 and $\Delta Slc35a1$ Bulk mixed population 3 BV2 cells were sorted into wells of a 96-well tissue culture plate containing 100 μ L of BV2 Maintenance Medium and propagated to establish clonal populations referred to as $\Delta Cmas$ Bulk clones 1-3 and $\Delta Slc35a1$ Bulk clones 1-4. Lectin binding was assessed via flow cytometry as described (4.6.1). BV2- $\Delta Cmas$ Bulk clone 2 and BV2- $\Delta Slc35a1$ Bulk clone 3, the clones exhibiting the least amount of lectin binding, were used for all further experiments and are referred to as $\Delta Cmas$ and $\Delta Slc35a1$ for simplicity throughout the remainder of this thesis.

4.7 Presto blue cell viability assay

WT, $\Delta Cmas$, $\Delta Slc35a1$ cells were plated at a density of 10^4 cells/well in 96-well tissue culture plates and incubated at 37°C overnight. Cells were adsorbed with reovirus at an MOI of 100 PFU/cell and incubated at RT for 1 h. The virus inoculum was removed and replaced with 100 μ L of BV2 Selection Medium. At various intervals post-inoculation (0, 24, 48h), 10 μ L of PrestoBlue Cell Viability Reagent (ThermoFisher Scientific) was added to wells, plates were

incubated at 37°C for 10 min, and total well fluorescence was quantified at 570 using a Synergy H1 microplate reader (BioTek). Three wells per condition were assessed in two independent experiments.

4.8 Fluorescent focus unit assessment of reovirus infectivity

WT, $\Delta Cmas$, and $\Delta Slc35a1$ cells were plated at a density of 10^4 cells/well in 96-well tissue culture plates and incubated at 37°C overnight. Cells were adsorbed with reovirus at an MOI of 100 PFU/cell and incubated at RT for 1 h. The virus inoculum was removed, and 100 μ L of BV2 Selection Medium was added to the monolayer of cells. Cells were incubated at 37°C for 24 h, washed 1X with PBS^{-/-}, and fixed with 100 μ L of ice-cold methanol at -20°C for at least 30 min. Fixed cells were washed 2X with PBS^{-/-}, blocked with 1% BSA for 30 min, and incubated with reovirus antiserum diluted 1:1000 in PBS^{-/-}, 0.5% Triton X-100 at RT for 1 h. Cells were washed two times with PBS^{-/-} and incubated with anti-rabbit Alexa488 secondary antibody (Thermo Fisher) at a dilution of 1:10,000 at RT for 1 h. Cells were washed two times with PBS^{-/-}, and nuclei were counterstained with 4',6-diamidino-2-phenylindole (DAPI, ThermoFisher Scientific) at 1:2,000 dilution at RT for 5 min. Cells were imaged for reovirus antigen and DAPI using a Lionheart FX automated microscope (BioTek) equipped with a 20X air objective. The percentage of cells infected with reovirus (number of reovirus infected cells/total number of cells) was quantified using Gen5+ software (BioTek).

4.9 Cmas and Slc35a1 protein detection by immunoblot

WT, $\Delta Cmas$, and $\Delta Slc35a1$ BV2 cells were plated at 0.5×10^6 cells/well in 6-well tissue culture plates and incubated at 37°C overnight. Cells were washed 2X with PBS^{-/-}, lifted in 1 mL of PBS^{-/-} by cell scraping, and collected into Eppendorf tubes. Cells were pelleted at 1000 rpm at 4°C for 5 min, resuspended in 150 μL of RIPA buffer (Sigma Aldrich) supplemented to contain 10% protease inhibitors (Thermo Fisher), and incubated on ice for 5 min. Lysed cells were centrifuged at 10,000 rpm at 4°C for 10 min. The soluble fraction was transferred to a 1.5 mL Eppendorf tube. Total protein in the soluble fraction was quantified by Lowry assay using the DC Protein Assay Kit II (Bio-Rad; 5000112) following the manufacturer's instructions. Samples were diluted 1:1 (vol:vol) (30 μg for WT, $\Delta Cmas$, and $\Delta Slc35a1$ cell lysates; 5 μg for positive controls) in 2X Laemmli sample buffer with 2-Mercaptoethanol and incubated at 95°C for 10 min. Samples were loaded into wells of 4-20% Bis-Tris gels (Bio-Rad; 4561096) and electrophoresed at 100 V for 90 min. Following electrophoresis, proteins were transferred to nitrocellulose membranes for immunoblotting. The nitrocellulose membranes were blocked with Licor Blocking Buffer (Licor) at RT for 30 mins and incubated with JAM-A- or NgR1-specific antibodies at 4°C overnight. Cells were washed three times TBS-T at RT for a total of 30 mins and incubated with an IRDye anti-goat 680 secondary antibody (Li-COR Biosciences) at a dilution of 1:10,000 at RT for 1 h. Gels were scanned using an Odyssey CLx imaging system (Li-COR).

4.10 *Cmas* and *Slc35a1* cDNA transfection of $\Delta Cmas$ and $\Delta Slc35a1$

Plasmids containing the mouse *Cmas* (Accession No. NM_009908.2) and *Slc35a1* (Accession No. NM_011895.3) cDNAs in pcDNA3.1+/C-(K)-DYK or pcDNA3.1(+)-N-DYK vectors, respectively, were obtained from Genescript. DH5 α competent cells (Biopioneer; GACC50) were transformed with plasmids, plated on agar supplemented to contain 100 μ g/mL ampicillin, and propagated at 37°C overnight. Individual colonies were picked and propagated at 37°C overnight in 250 mL LB Broth supplemented to contain 100 μ g/mL ampicillin. DNA was purified using a HiSpeed Plasmid Midi Kit (QIAGEN; 12643) according to the manufacturer's instructions.

One day prior to transfection, $\Delta Cmas$ and $\Delta Slc35a1$ cells were plated at a density of 0.5×10^6 cells per well in 6-well tissue culture plates. Approximately 1.0 μ g of *Cmas* or *Slc35a1* DNA was combined with FuGene 6 (Promega) in Opti-MEM (Gibco) and incubated at RT for approximately 40 min. Opti-MEM mixture was added dropwise to plated cells, and cells were incubated at 37°C for 24 h post-transfection before SA expression or reovirus-binding was assessed using flow cytometry, as described in section 4.6.1. Transfected $\Delta Cmas$ and $\Delta Slc35a1$ are denoted as $\Delta Cmas+Cmas$ and $\Delta Slc35a1+Slc35a1$, respectively.

4.11 Statistical analysis

All statistical tests were conducted using PRISM 7 (GraphPad Software). *P* values less than 0.05 were considered to be statistically significant. Descriptions of the specific tests used are

found in the figure legends. Differences are compared to WT BV2 cell values unless otherwise noted.

5.0 Results

5.1 CRISPR screen identifies host factors required for reovirus replication

We hypothesized that host genes required for reovirus replication could be identified using a CRISPR-Cas survival screen. Using such a strategy, we selectively disrupted target gene function across the entire mouse genome to identify host factors required for reovirus-induced cell death (Figure 6). Four reovirus strains (T1L, T3D, T3SA+, or T3SA-) were inoculated into a BV2 mouse microglial CRISPR cell library containing gene disruptions targeting over 20,000 genes. Nine days post inoculation, gDNA was isolated from the surviving cells and deep sequenced. A STARS analysis was conducted to identify enriched CRISPR gRNAs within the surviving cell population. These four prototype reovirus strains were chosen because of differences in their capacity to bind to SA. T1L reovirus binds to a specific α 2,3-linked *N*-acetyl-SA [38] while T3D reovirus binds terminal α 2,3-, α 2,6-, or α 2,8-linked SAs [40]. T3SA+ and T3SA- viruses differ by a single amino acid polymorphism in the σ 1 body domain (Figure 5) [41]. This polymorphism allows T3SA+ to engage cell-surface SA, while T3SA- does not.

Following infection with T3SA+, sequencing and analysis of the surviving cells revealed an enrichment of sgRNAs targeting genes encoding components of the 40S and 60S ribosomal subunits (Table 2). After initial entry steps, transcriptionally active reovirus core particles are released into the cytoplasm [51]. Similar to all other viruses, reoviruses use host ribosomes to translate viral mRNAs. Cells lacking *Rplp2*, *Rps8*, *Rpl27*, or *Rpl7a* are hypothesized to have nonfunctional ribosomes incapable to translating viral mRNAs, thus blocking reovirus replication. Cells lacking these important ribosomal genes did not succumb to reovirus-induced cell death.

These results suggest that translation of viral mRNAs is required for reovirus-induced cell death, as previously shown [28-30, 52], and serve as validation for the CRISPR screening method.

Following infection with T1L, T3D, and T3SA+, surviving cells displayed an enrichment of sgRNAs targeting genes that promote SA expression on the cell surface. This trend was not observed subsequent to infection with T3SA- reovirus (Table 2). Three SA synthesis genes, *St3gal4*, *Slc35a1*, and *Cmas*, were identified following inoculation with T1L, T3D, or T3SA+. *St3gal4* encodes a member of the sialyltransferase 29 protein family. This group of enzymes function in the glycosylation and production of α 2,3-linked sialoglycoconjugates. The identification of *St3gal4* in T1L, T3D, and T3SA+ CRISPR screens was consistent with our understanding of the type of SA engaged by both T1 and T3 reoviruses. Both of these reovirus serotypes can engage α 2,3-linked-SAs.

Two other identified genes, *Cmas* and *Slc35a1*, are involved in critical steps of the synthesis pathway for all SAs including α 2,3-, α 2,6-, and α 2,8-linked sialoglycoconjugates (Figure 2). Mutations introduced into either *Cmas* or *Slc35a1* disrupt the synthesis of all SAs and result in an intracellular accumulation of free SAs [53]. Interestingly, genes required for SA synthesis were not identified following T3SA- infection. The enrichment of genes required for SA synthesis, in addition to our prior understanding of SA as an attachment factor for reovirus, made these genes attractive candidates for further validation studies. My research concentrated on validation and characterization of *Slc35a1* and *Cmas*, since these two genes were identified in the CRISPR screen following T1L, T3D, and T3SA+ infection and are critical to the synthesis of all SAs.

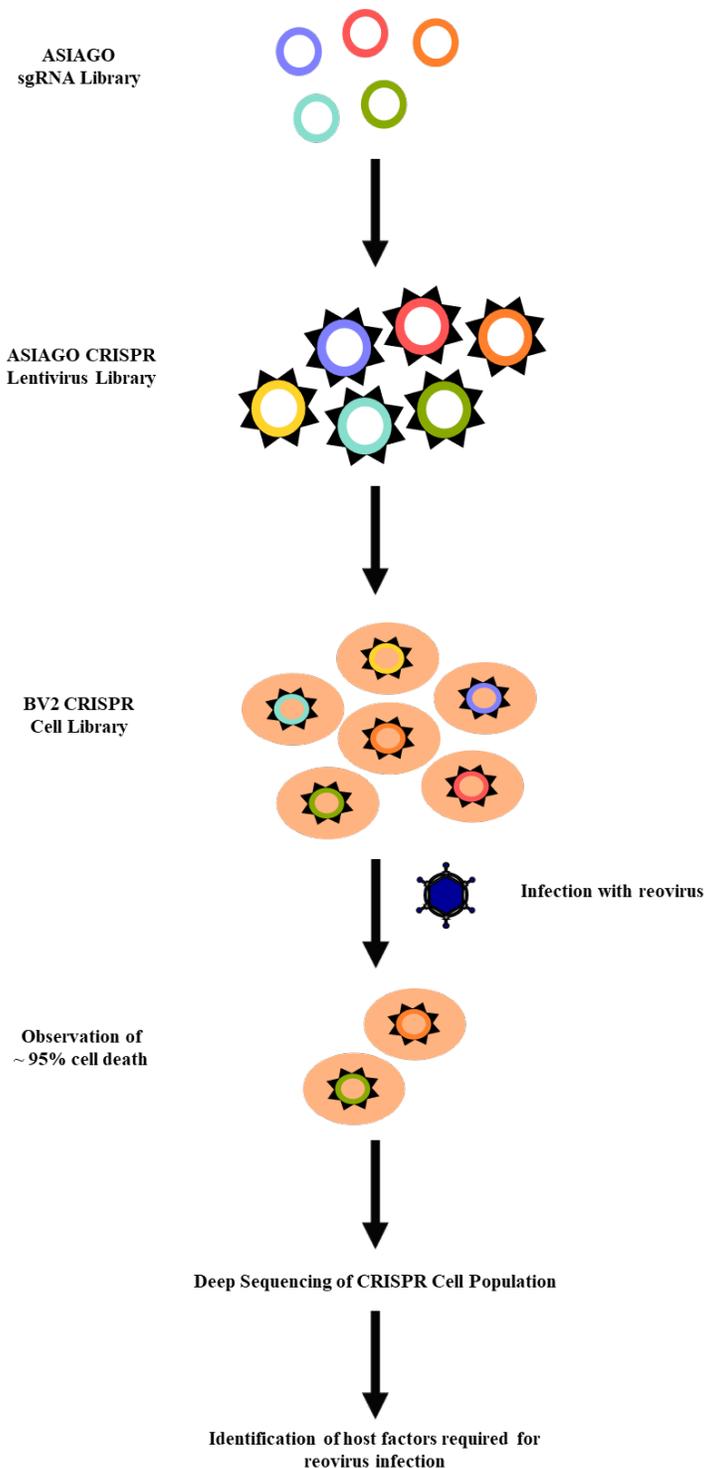
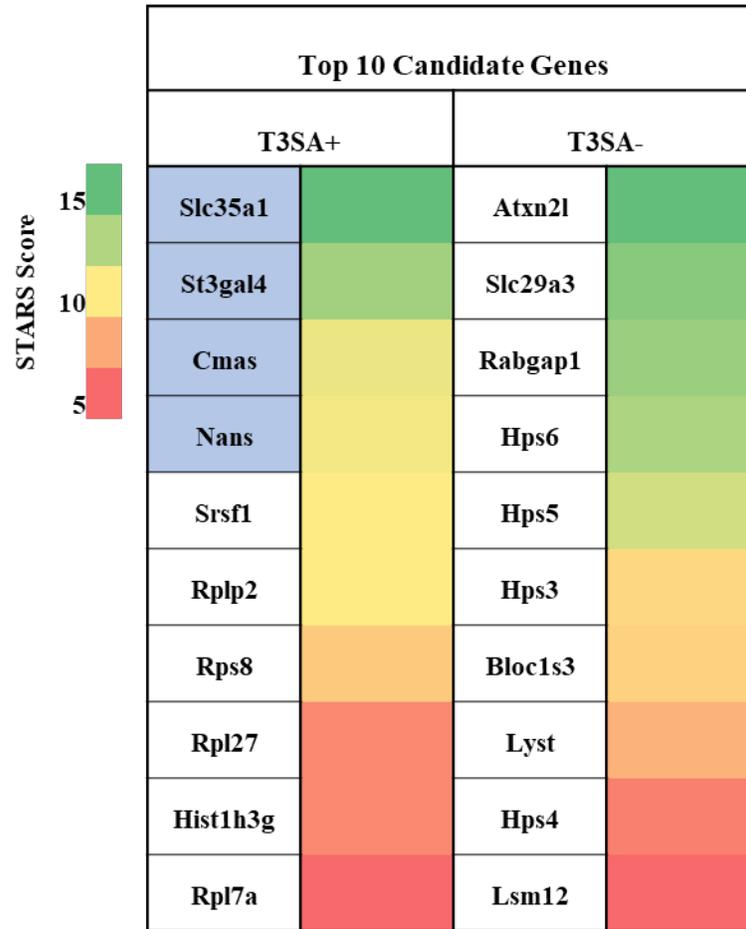


Figure 6: A CRISPR-Cas genetic screen was conducted to identify host factors required for reovirus infection.

The mouse Asiago sgRNA CRISPR library was delivered by lentiviral transduction into BV2 microglial cells. Cells were adsorbed with reovirus strains T1L, T3D, T3SA+, or T3SA- at an MOI of 100 PFU/cell and monitored for cell death. Upon observation of ~95% cell death, surviving cells were deep sequenced, and genomic DNA was analyzed.

Table 2: Whole-genome CRISPR screen identifies SA synthesis genes as required for reovirus induced death in BV2 cells.



The top 10 candidates of the CRISPR screens for T3SA+ and T3SA- ranked by STAR analyses. Heat map indicates STAR value. Genes required for SA synthesis are indicated by blue shading.

5.2 STRING analysis

The STRING network database was used to determine gene-interaction networks for *Cmas* and *Slc35a1* (Figure 7). Four genes, *Cmas*, *Slc35a1*, *St3gal4*, and *Nans*, were identified in the CRISPR screen with a significant q-value of less than 0.01. These genes are indicated in red. Eight proteins in the networks, determined not to be required for T3SA+ infection based on the q-values

in the CRISPR screen, are indicated in blue. Those genes identified as interactors in the STRING network but not identified in the CRISPR screen are indicated in gray. The interaction networks were functionally enriched for genes in the SA metabolic pathway (GO: 0006054; $FDR_{Cmas} = 2.12 \times 10^{-19}$; $FDR_{Slc35a1} = 1.45 \times 10^{-13}$). Since the genes identified as required for T3SA+ infection in the CRISPR screen reside in hubs of the interaction networks, the overall SA metabolic pathway may be required for T3SA+ infection.

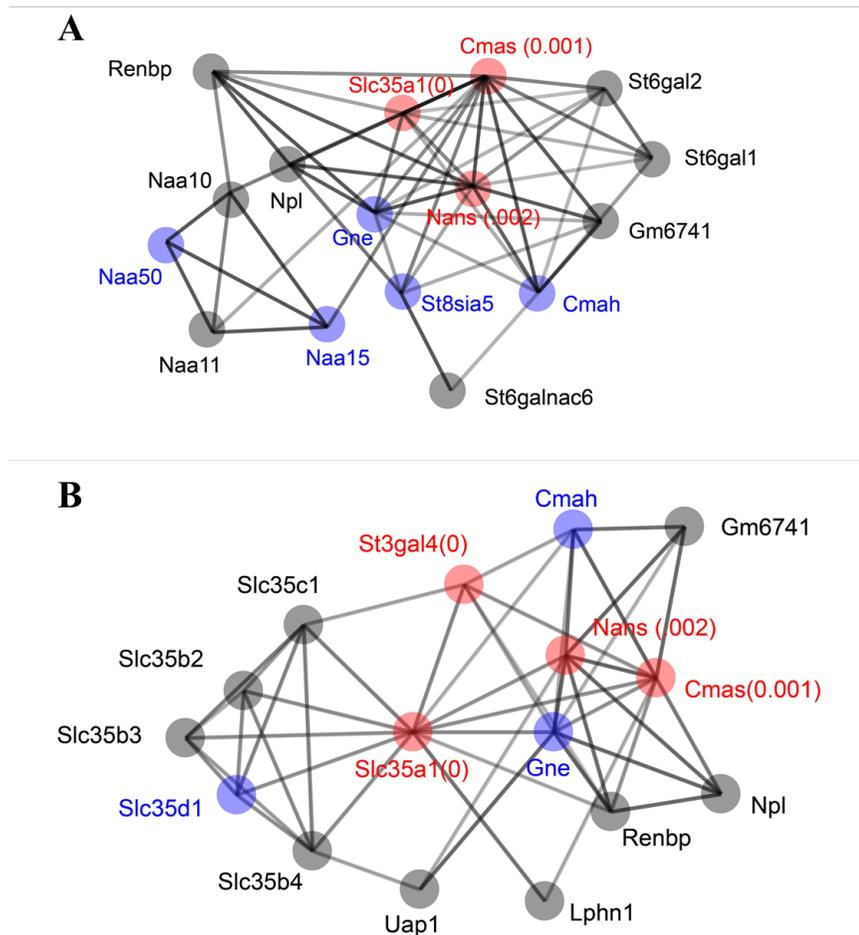


Figure 7: Mapping the interaction networks of *Slc35a1* and *Cmas*. Using the STRING database, interaction networks were prepared for (A) *Cmas* and (B) *Slc35a1* based on *Mus musculus* as the model organism. Gene interaction scores greater than 0.4 and thresholds of ten primary interactions and five secondary interactions were used to establish the network. The q-values were obtained from the analysis of the CRISPR screen using T3SA+. Genes with q-values less than 0.01 were considered to be required for T3SA+ infection and are indicated in red. Genes with q-values greater than 0.01 were not considered to be required for T3SA+ infection and are indicated in blue. Genes that were identified in the interaction network but not identified in the CRISPR screen are indicated in gray.

5.3 Generation of cell populations with CRISPR-targeted *Slc35a1* or *Cmas*

Using our understanding of the importance of *Cmas* and *Slc35a1* in the SA synthesis pathway, we sought to evaluate the importance of these two genes during reovirus infection. By generating cell lines lacking the *Cmas* or *Slc35a1* gene by CRISPR knockout, we were able to test the hypothesis that these genes are required for reovirus infection. Three independent BV2- Δ *Slc35a1* cell populations and two independent BV2- Δ *Cmas* cell populations were engineered. Each population was designed to incorporate different sgRNAs targeting either *Cmas* or *Slc35a1*. The resulting cell populations were mixed populations, in which approximately 80% of cells contained gene ablations for *Cmas* or *Slc35a1*, and approximately 20% of cells did not incorporate the sgRNA therefore the targeted gene remained functional.

5.4 Characterization of CRISPR cell line gene disruption

To assess the expression of SA on the cell surface, we used a fluorescein-labeled lectin. The lectin WGA, derived from *Triticum vulgare*, binds numerous sialoglycoconjugates that terminate in α 2,3-, α 2,6-, and α 2,8-linked SA residues [54]. Upon assessment of lectin binding by flow cytometry, Δ *Cmas* parental population B11 (Δ *Cmas*ParentB11) and Δ *Slc35a1* parental population B8 (Δ *Slc35a1*ParentB8) displayed the lowest amount of SA expression on the cell surface as assessed by lectin binding (data not shown).

To establish a cell line with a single population expressing a reduced amount of SA, I conducted bulk cell sorting of Δ *Cmas*ParentB11 and Δ *Slc35a1*ParentB8 BV2 cells. Following collection of 10,000 low-lectin binding cells from these two populations, I propagated these cells

(called “Bulk”) for several passages. I assessed lectin binding by flow cytometry for four $\Delta Cmas$ Bulk and three $\Delta Slc35a1$ Bulk mixed populations. The bulk populations that expressed the lowest level of SA were $\Delta Cmas$ Bulk mixed population 3 and $\Delta Slc35a1$ Bulk mixed population 3 (data not shown).

To establish a truly clonal cell line expressing a reduced amount of SA, I conducted single-cell sorting of $\Delta Cmas$ Bulk population 3 and $\Delta Slc35a1$ Bulk population 3. Following single-cell sorting and propagation, I assessed lectin binding by flow cytometry for three $\Delta Cmas$ Bulk population 3 clones and four $\Delta Slc35a1$ Bulk population 3 clones (Figure 8). The second clone of $\Delta Cmas$ Bulk population 3, referred to as $\Delta Cmas$ for the remainder of this thesis, and the third clone of $\Delta Slc35a1$ Bulk population 3, referred to as $\Delta Slc35a1$ for the remainder of this thesis, were the clonal populations that expressed the lowest levels of SA. I used these clones for all further experiments.

To compare SA expression on the cell surface of WT, $\Delta Cmas$, and $\Delta Slc35a1$ cells, I conducted lectin-binding assays and quantified cell fluorescence using flow cytometry. The level of lectin binding was significantly decreased on $\Delta Cmas$ and $\Delta Slc35a1$ cells relative to WT cells (Figure 9). These data demonstrate that disruption of *Cmas* or *Slc35a1*, two genes involved in the SA synthesis pathway, efficiently impairs SA expression on the cell surface.

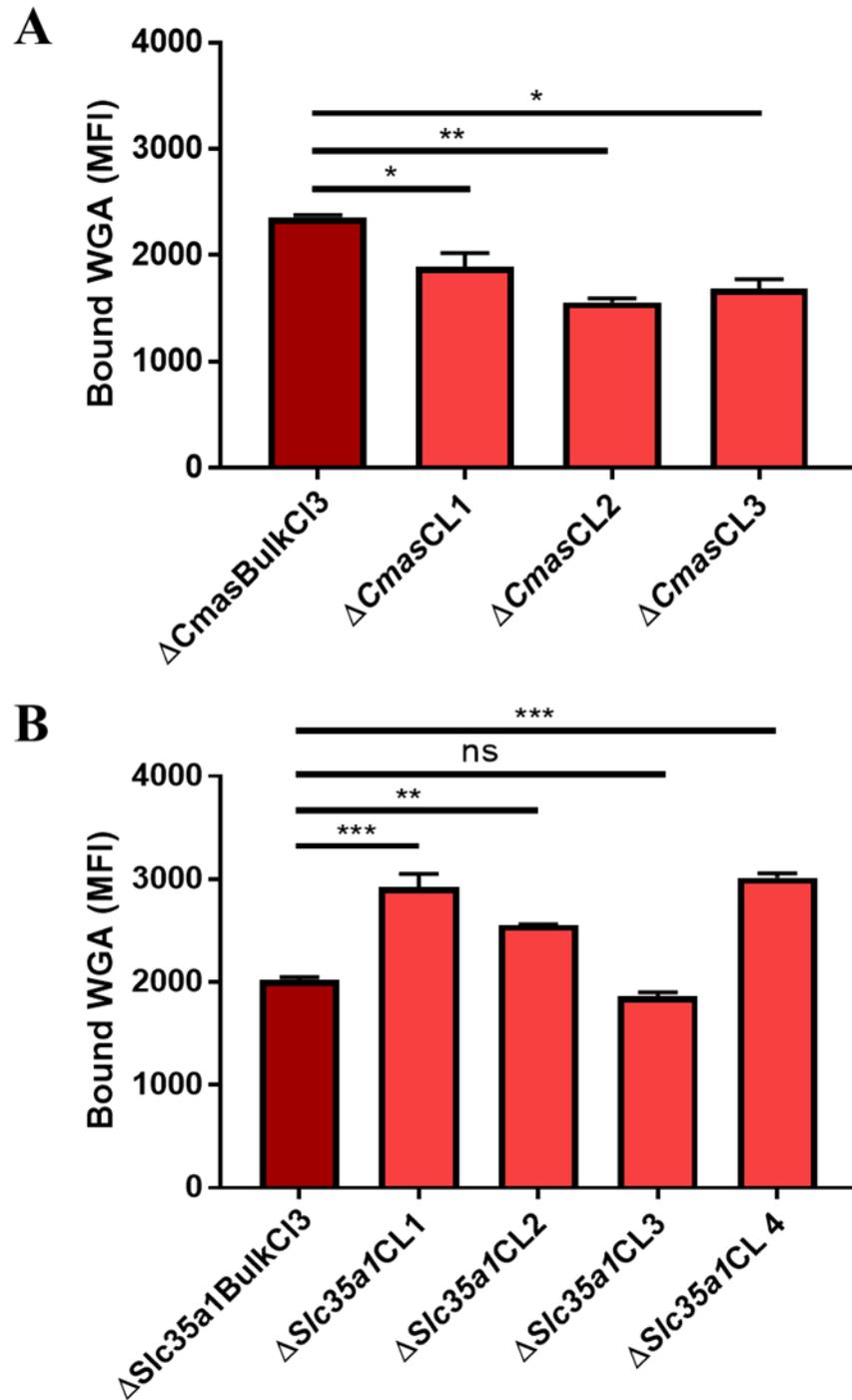


Figure 8: SA expression is altered on CRISPR clones after single cell sorting. $\Delta Cmas$ (A) and $\Delta Slc35a1$ (B) cells were incubated with 0.005mg/mL fluorescein labeled wheat germ agglutinin (WGA) to assess cell-surface sialic acid expression. Fluorescence was detected by flow cytometry and the median fluorescence intensity (MFI) was quantified. Data represent one experiment with duplicated samples. Error bars represent standard error of the mean (SEM). Values that differ significantly from WT by one-way ANOVA with with Dunnett’s multiple comparisons test are indicated (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$).

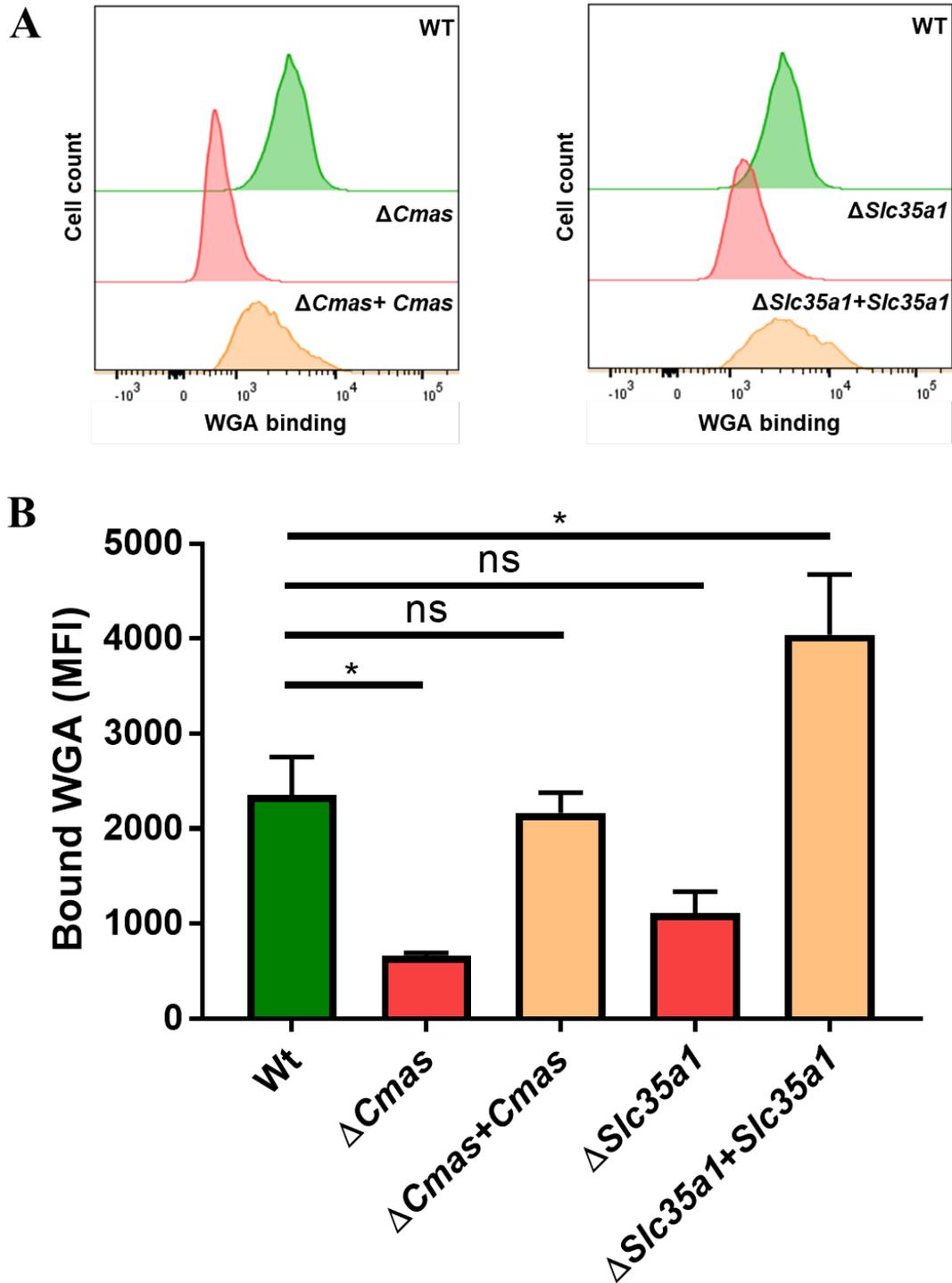


Figure 9: Sialic acid expression is diminished after CRISPR knockout and restored with transient transfection.

Cells were incubated with fluorescein-labeled WGA to assess cell-surface sialic acid expression. Cell fluorescence was detected by flow cytometry. (A) Representative flow cytometry plots of fluorescein WGA binding to WT, CRISPR KO cells, and transfected cells. (B) The MFI was quantified. Data represent two independent experiments each with duplicated samples. Error bars represent SEM. Values that differ significantly from WT by one-way ANOVA with Dunnett's multiple comparisons test are indicated (*, $P < 0.05$).

5.5 Analysis of cell viability during reovirus infection

Cell viability, or assessment of the metabolic activity of cells in a population, was used to validate the host genes, *Cmas* and *Slc35a1*, identified in the CRISPR screen. To identify the capacity of $\Delta Cmas$ and $\Delta Slc35a1$ cells to remain resistant to reovirus infection and determine the effect of *Cmas* or *Slc35a1* gene disruption on reovirus-induced cell death, I inoculated WT, $\Delta Cmas$, and $\Delta Slc35a1$ BV2 cells with T3SA+ or T3SA- (Figure 10). T3SA+ and T3SA- were used for validation studies because of the difference in the capacity of these strains to bind SA. Using these viruses, phenotypic changes caused by a disruption in *Cmas* and *Slc35a1* could be directly correlated with SA engagement.

At 0, 24, and 48 h post-inoculation (hpi) with T3SA-, cell viability was comparable between all three cell lines (Figure 10A). At 24 and 48 hpi with T3SA+, $\Delta Cmas$ and $\Delta Slc35a1$ cell lines displayed significantly enhanced viability when compared to WT cells (Figure 10B). The viability exhibited by $\Delta Cmas$ and $\Delta Slc35a1$ after inoculation with T3SA+ was comparable to the viability observed after inoculation with T3SA-, a virus incapable of engaging SA. These data are consistent with the results of the CRISPR screen and indicate that genetic disruption of *Cmas* and *Slc35a1* protects murine microglial cells from reovirus-induced cell death.

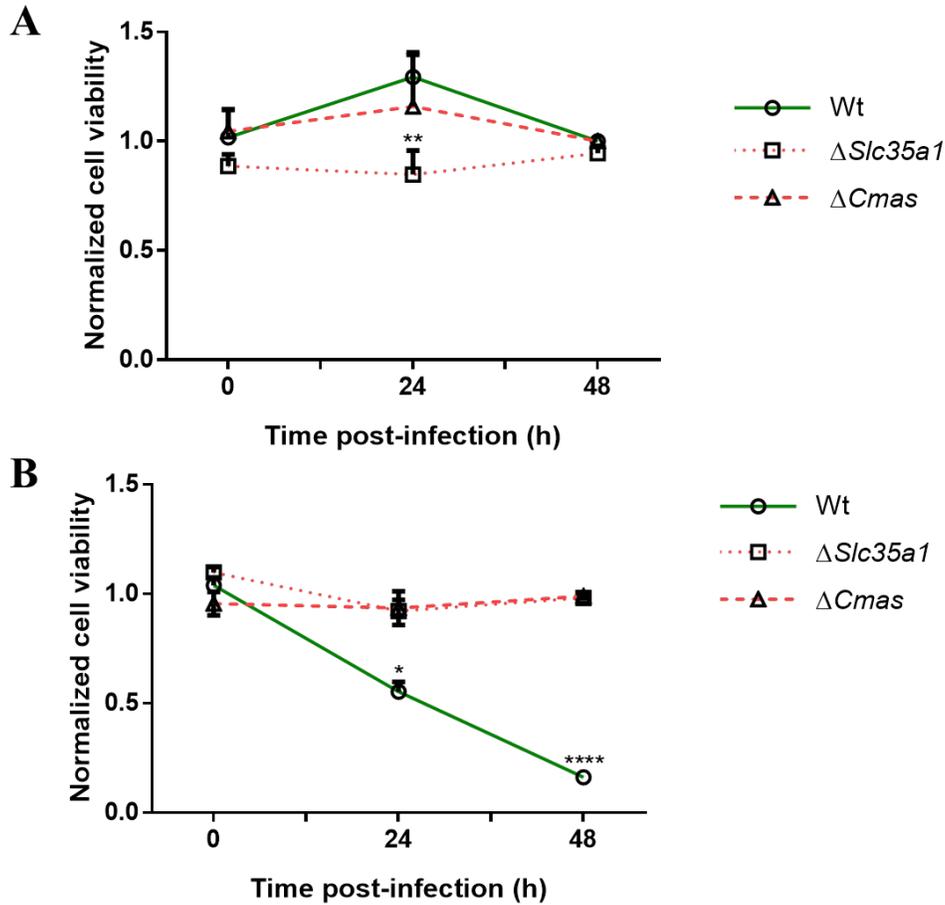


Figure 10: BV2 CRISPR clones are protected from reovirus-induced cell death. The cell lines shown were adsorbed with (A) T3SA- or (B) T3SA+ and cell viability was quantified using the PrestoBlue® assay at 0, 24, 48 hpi. Data presented are normalized to the uninfected condition of the respective cell line. Data represent two independent experiments, each with duplicate samples. Error bars represent SEM. Values that differ significantly from WT by one-way ANOVA with with Tukey's multiple comparisons test are indicated (*, $P < 0.05$; **, $P < 0.01$; ****, $P < 0.0001$).

5.6 Quantification of reovirus infectivity of WT and BV2 CRISPR clones

To test our hypothesis that *Cmas* and *Slc35a1* encode host factors required for productive reovirus infection, I tested the capacity of reovirus to infect cells with *Cmas* or *Slc35a1* gene disruptions. I inoculated $\Delta Cmas$ and $\Delta Slc35a1$ cells with T3SA+ or T3SA- at an MOI of 100 PFU/cell and monitored infectivity at 24 hpi by FFU assay. $\Delta Cmas$ and $\Delta Slc35a1$ cells exhibited

significantly decreased T3SA+ infectivity compared with WT cells (Figure 11), which is consistent with results obtained from the viability studies. The level of T3SA+ infectivity following inoculation of $\Delta Cmas$ and $\Delta Slc35a1$ cells was comparable to that observed following T3SA- inoculation. These data suggest that reovirus is incapable of efficiently infecting murine microglial cells lacking *Cmas* and *Slc35a1*. Interestingly, WT BV2 cells had generally poor infectivity with less than 60% of cells scoring positive for infection despite the high MOI used.

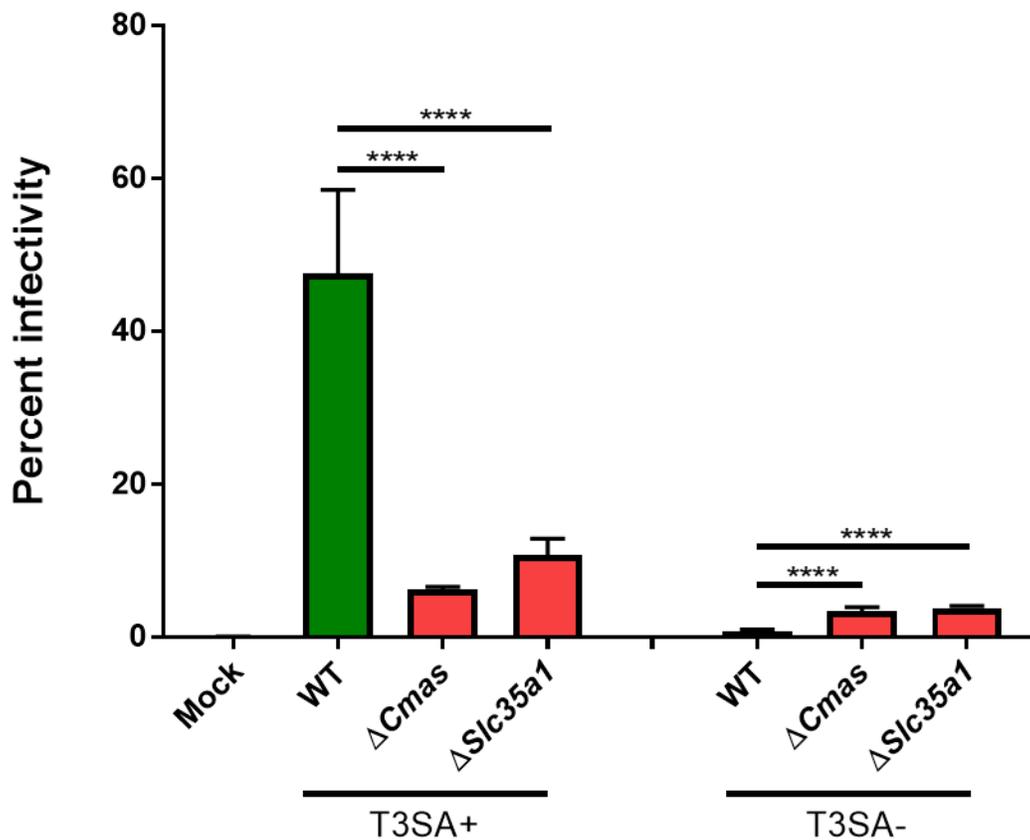


Figure 11: *Cmas* and *Slc35a1* are required for reovirus infection.

Cells were adsorbed with T3SA+ or T3SA- at an MOI of 100 PFU/cell. The percentage of infected cells was determined by enumeration of reovirus-positive cells 24 hpi from immunofluorescence images. Data represent two independent experiments each containing two replicates. Error bars represent SEM. Values that differ significantly from WT by one-way ANOVA with Dunnett's multiple comparisons test are indicated (****, $P < 0.001$).

5.7 Evaluation of JAM-A and NgR1 expression by immunoblot

To further understand the low infectivity levels observed in WT BV2 cells, the expression of JAM-A and NgR1, two known reovirus receptors, on WT, $\Delta Cmas$, and $\Delta Slc35a1$ cells was quantified by immunoblotting. I lysed WT, $\Delta Cmas$, and $\Delta Slc35a1$ cells, isolated the soluble protein fraction, and quantified the total protein in the cell lysates by Lowry assay. The soluble protein fraction of cell lysates were resolved by SDS-PAGE and probed for protein expression by immunoblotting using receptor-specific polyclonal antibodies. No JAM-A or NgR1 protein was visualized (Figure 12). Nonetheless, the poor susceptibility of BV2 cells to reovirus infection may result from diminished expression of reovirus receptors by these cells.

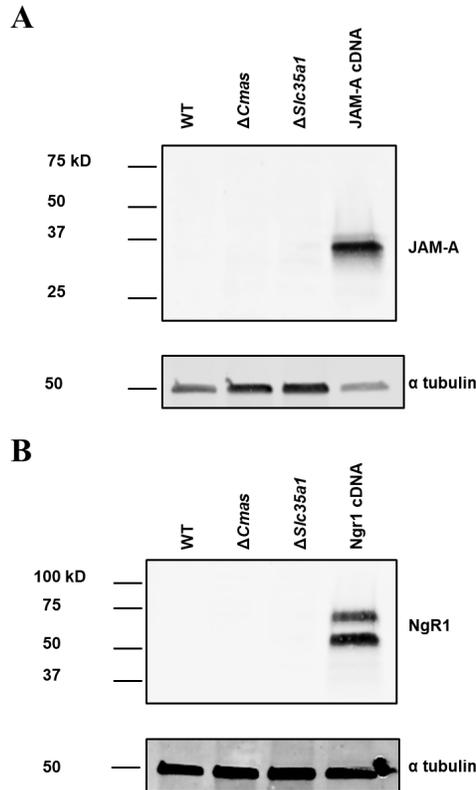


Figure 12: JAM-A and NgR1 expression is not detected on WT, $\Delta Cmas$, and $\Delta Slc35a1$ BV2 cells.

Cells were lysed in RIPA buffer containing protease inhibitors. Soluble protein was analyzed by SDS-PAGE and visualized by immunoblotting with either (A) JAM-A- or (B) NgR1-specific antibodies and an α -tubulin-specific monoclonal antibody as a loading control.

5.8 Evaluation of reovirus binding to CRISPR clones

To further explore the initial interactions between reovirus and cells with *Cmas* or *Slc35a1* gene disruptions, I studied the initial reovirus binding step. I hypothesized that (1) T3SA+, a virus capable of engaging SA would bind to WT cells more efficiently than to $\Delta Cmas$ and $\Delta Slc35a1$ cells and that (2) T3SA-, a virus incapable of engaging SA would not bind to WT, $\Delta Cmas$, or $\Delta Slc35a1$ cells.

To examine the capacity of reovirus to bind to the surface of $\Delta Cmas$ and $\Delta Slc35a1$ cells, I incubated cells with fluoresceintated T3SA+ or T3SA- and quantified bound virus by flow cytometry (Figure 13A,B). The quantity of T3SA+ virus bound to the surface of $\Delta Cmas$ and $\Delta Slc35a1$ cells was significantly decreased relative to that of WT cells (Figure 13C). Moreover, the quantity of T3SA+ virus bound to $\Delta Cmas$ and $\Delta Slc35a1$ cells was similar to that observed following incubation with T3SA- virus. These results indicate that murine microglial cells with *Cmas* and *Slc35a1* gene disruptions are incapable of efficiently engaging reovirus.

5.9 Complementation of *Cmas* and *Slc35a1*

To determine whether the lectin- and virus-binding phenotypes observed in studies of $\Delta Cmas$ and $\Delta Slc35a1$ cells were due to the targeted gene disruptions, I used a gain-of-function approach to complement the CRISPR knockout clones. I introduced cDNA encoding *Cmas* or *Slc35a1* into $\Delta Cmas$ or $\Delta Slc35a1$ cells using transient transfection and allowed 24 h for protein expression before assessing WGA or virus binding. For both CRISPR clones, lectin binding was restored to levels comparable to or exceeding that of WT BV2 cells (Figure 9). The increased

levels of lectin binding indicate increased SA expression on the surface of transfected cells. These data support the hypothesis that the observed decrease in SA expression in $\Delta Cmas$ and $\Delta Slc35a1$ cells is due to *Cmas* and *Slc35a1* gene knockout, respectively. In preliminary experiments, virus binding was not restored to levels comparable to WT BV2 cells, for both CRISPR clones (Figure 13C). These experiments are being repeated.

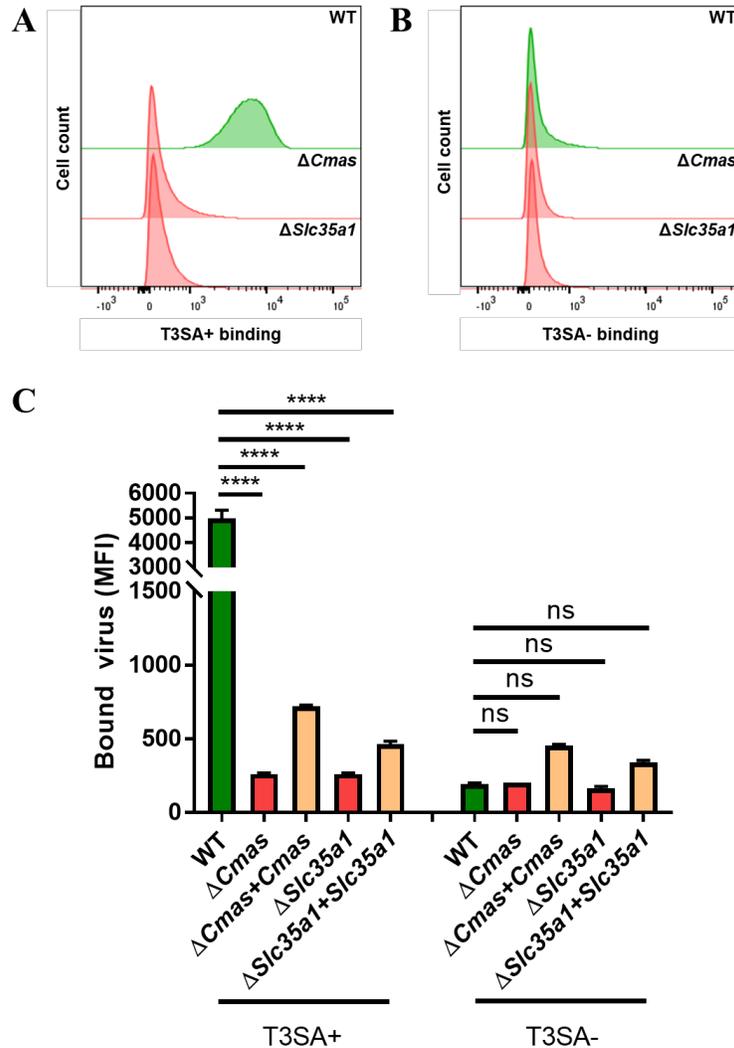


Figure 13: *Cmas* and *Slc35a1* are required for reovirus binding to cells.

Cells were adsorbed with fluorescently labeled (A) T3SA+ or (B) T3SA- to assess virus binding. Fluorescence was detected by flow cytometry. Representative flow cytometry plots are shown. (C) The MFI of Alexa-488 was quantified. Data represent two experiments with duplicated samples. Error bars represent SEM. Values that differ significantly from WT by one-way ANOVA with Dunnett's multiple comparisons test are indicated (*, $P < 0.05$; **, $P < 0.01$).

6.0 Discussion

This work has identified two host genes, *Cmas* and *Slc35a1*, as required for early entry steps in reovirus infection of microglial cells and provided additional insight into how genes involved in the synthesis and expression of SA enhance these initial entry steps. It is important to further our understanding of engagement of cell-surface receptors by viruses, as the attachment step in viral replication is a critical determinant of tropism and disease. In particular, understanding host genes that regulate successful reovirus infection of microglial cells, an important component of the innate immune response in the CNS, could identify potential targets to limit neuropathogenesis.

In this study, we used CRISPR cells-survival screen to identify host genes required for reovirus infection. Following infection with T3SA+, top candidates based on STARS scores revealed an enrichment of genes involved in the SA synthesis pathway (Table 2). At the cell surface, reovirus interacts with SA during the initial steps of infection. Proteinaceous receptors have been previously identified for reovirus. However, a complete understanding of host genes required for efficient reovirus binding remain unknown making this subject an area of interest and ongoing research. Three top candidates identified in the CRISPR screen following T1L, T3D, and T3SA+ inoculation were *St3gal4*, *Slc351*, and *Cmas*. *St3gal4* is a gene involved in the sialylation of α -2,3-linked glycoconjugates [55]. Since both T1 and T3 viruses engage α 2,3-linked SAs, the identification of *St3gal4* as a top candidate in the CRISPR screen provided validation of the specificity of the screen results. *Slc35a1* and *Cmas* are involved in early steps of the SA pathway and are essential for the production of all types of SAs. Interestingly, *Slc35a1* was recently identified as an essential host factor for influenza virus cell entry [56]. It is important to understand

the role of these two host genes during the initial interaction between virus and host, since many viruses engage SA as an attachment factor.

To define the importance of *Cmas* and *Slc35a1* in SA synthesis, a STRING network was constructed for each gene (Figure 7A,B). The resultant networks display the predicted gene interactions of *Cmas* or *Slc35a1*. Interestingly, STRING analysis revealed that not all genes involved in the SA synthesis pathway were identified in the CRISPR screen. Essential genes are often not identified using this genetic screening technique, as some disruptions caused by sgRNAs result in cell death or block proliferation causing a loss from the cell library [11]. A lack of identification of some SA biosynthesis genes could be attributed to their requirement for cell survival, as anticipated for *Naa10* and *Naa11*, or the capacity for reovirus to still induce cell death in BV2 microglial cells with these gene ablations, as anticipated for *Npl* and *Renb*. *Naa10* and *Naa11*, two genes identified by the STRING network to interact with *Cmas*, are required for post-translational protein acetylation and thus essential for normal cell function (Figure 7A) [57]. I hypothesize that *Naa10* and *Naa11* are essential for cell survival; therefore, cells with these gene ablations would not be viable and thus not identified by the CRISPR screen. Similarly, *Npl* and *Renbp*, genes identified by the STRING networks to interact with *Cmas* and *Slc35a1*, are involved in the degradation, desialylation, and salvage pathways of SA synthesis but are not required for the expression of SA on the cell surface (Figure 7A,B) [58]. I hypothesize that T3SA⁺ can efficiently infect BV2 cells with *Npl* and *Renbp* ablations and induce cell death. Interestingly, several genes involved in the synthesis of α 2,6-linked SAs (e.g. *St6galnac6*, *St6gal1*, and *St6gal2*) are depicted in the *Slc35a1* STRING network but were not identified using the CRISPR screen as candidates required for T3SA⁺ infection (Figure 7B). Since T3 reovirus strains have the capacity to bind three different types of SAs (α 2,3-, α 2,6-, or α 2,8-linked SAs), it is possible that T3SA⁺

and T3D can initiate viral entry by engaging α 2,3- or α 2,8-linked SAs to infect these cells and ultimately cause cell death. T1L and T3SA- are not thought to engage α 2,6-linked SAs and, therefore, would not be predicted to select cells with ablation of these genes. Due to the identification of *Cmas* and *Slc35a1* in the CRISPR screen following infection with T1L, T3D or T3SA+ and their requirement in the global production of SAs, a previously defined attachment factor for reovirus, we decided to validate the function of these two genes during reovirus infection.

Attachment factors are cellular moieties involved in the very first binding event between virus and cell by providing low-affinity but high-avidity interactions to tether the virus to the cell membrane. Viral entry receptors are cell-surface molecules that provide specific functions required for productive viral infection [59]. Receptors binding often directly triggers signaling cascades causing molecular changes leading to viral entry and subsequent infection. Understanding the strength and valency of these interactions improves knowledge of their function. Affinity refers to the strength of a single binding interaction, whereas avidity quantifies the strength of multiple binding interactions occurring simultaneously. Binding strengths of viral attachment proteins with glycans, found abundantly on the cell surface, are often assessed using avidity since viral attachment proteins, like reovirus σ 1, are present in high copy number on the viral capsid, and each copy often displays several identical binding sites. T3SA+ reovirus is thought to bind SA with low affinity ($K_D \approx 180$ - 1700 nM using a bivalent analyte model) but high overall avidity ($K_D \approx 5$ nM) [41]. In contrast, the affinity of T3 reovirus for JAM-A is significantly higher ($K_D \approx 60$ nM) [21]. These data suggest that SA serves as an attachment factor for reovirus, while JAM-A serves as a specific receptor. It is likely that types and distributions of SAs and other receptors on the cell surface influence virus avidity to cells [41]. Increased avidity between virus and cell as a

consequence of SA binding would be especially important for infecting cell with little cell-surface JAM-A expression.

I demonstrated that JAM-A and NgR1, two previously identified reovirus receptors are not detectable on WT, $\Delta Cmas$, and $\Delta Slc35a1$ BV2 cells by immunoblotting (Figure 12). The low-level expression of both of these receptors on BV2 cells may provide an explanation for the overall low levels of infectivity I observed following reovirus inoculation of WT BV2 cells (Figure 11). Despite hypothesized low levels or absence of JAM-A and NgR1, reovirus is capable of infecting WT BV2 cells. These findings are consistent with the hypothesis that SA can function as an entry receptor on BV2 microglial cells and are similar to results observed by Barton, et al. [21]. Prior to infection with T3SA+, HeLa cells were incubated with an anti-human JAM-A antibody. Replication of T3SA+ was minimally reduced suggesting that receptors other than JAM-A, perhaps SA, function in reovirus entry.

T3 reovirus strains are thought to infect microglial cells. *In vitro*, reovirus strain T3 Abney can infect amoeboid microglial cells, a subpopulation of activated microglia [60]. Infection levels of amoeboid microglia greater than 40% were observed by immunofluorescence of primary microglial cultures. Interestingly, the levels of infectivity observed during my research were comparable to those obtained by Goody, et al. Microglial cells are an essential component of the innate immune response in the CNS. These cells comprise the resident phagocytic cell population of the myeloid lineage and function to remove damaged neurons and maintain homeostasis in the CNS. Macropinocytosis, a type of endocytosis that involves the nonspecific uptake of extracellular material, is often exhibited by macrophage cells of the innate immune system including microglia [61]. Perhaps reovirus targets SA as a receptor on microglial cells to induce macropinocytosis and viral entry, bypassing the need to engage JAM-A or NgR1. Furthermore, the capacity of reovirus

to infect microglial cells could be an evolutionary advantage. Infecting these cells could allow reovirus to propagate robustly throughout the CNS with little detection due to avoidance of innate immune detection.

Virus-receptor interactions often regulate tissue tropism and viral pathogenesis. *In vivo*, the host determinants of reovirus dissemination are not fully understood. Reovirus transmission occurs primarily through fecal-oral routes. Following peroral inoculation, T3SA⁺ and T3SA⁻ infect the intestine, which serves as a primary site of replication. However, T3SA⁺ is detected earlier in sites of secondary replication, such as the brain and liver, indicating an enhanced dissemination efficiency [45]. T3 reoviruses spread by neural and hematogenous routes to infect neurons of the CNS, where they cause apoptosis and lethal encephalitis. Following direct intracranial inoculation, T3SA⁺ is more neurovirulent, reaches higher viral titers, and causes more apoptosis, when compared to T3SA⁻ [45] [62]. However, T3SA⁺ and T3SA⁻ exhibit comparable tropism [62]. Combined, these data suggest SAs function as attachment factors *in vivo* than definitive receptors. It is possible that the dissemination of T3 reoviruses capable of engaging SA through the CNS could be altered by tissue-specific gene disruptions targeting *Cmas* or *Slc35a1*.

The work presented in this thesis identified two genes, *Cmas* and *Slc35a1*, which are required for the synthesis and expression of SA on the cell surface, as necessary for reovirus infection of microglial cells. It is important to understand interactions between reovirus attachment protein $\sigma 1$ and cell-surface moieties, such as SA, to expand the current understanding of the receptors on microglial cells that are targeted by reovirus. Overall, my research elucidates two host genes involved in the initial virus-host interface that make microglial cells susceptible to reovirus infection.

7.0 Summary and Future Directions

Additional experiments must be completed to fully understand the contribution of *Cmas* and *Slc35a1* in the initial entry steps of reovirus infection. Most cell lines that display a glycan dependence for infection are more susceptible to T3 reovirus strains, likely due to the capacity of T3 reovirus to bind multiple different types of SA (α 2,3-, α 2,6-, or α 2,8-linked SA) [40]. I have shown that T3SA⁺ does not efficiently infect *Cmas* and *Slc35a1* CRISPR knockout cells, suggesting that both host genes are required for reovirus infection of murine microglial cells (Figure 11). To further understand the entry events in the reovirus infection cycle in which *Cmas* and *Slc35a1* contribute, I conducted a reovirus binding assay. In this assay, I observed a reduction in the binding of reovirus to *Cmas* and *Slc35a1* CRISPR knockout cells relative to the binding of WT cells (Figure 13), indicating that these host genes serve an important function in the initial interaction between reovirus and microglial cells.

To verify that the observed reduction in reovirus binding to and infectivity of BV2 cells is attributed to *Cmas* and *Slc35a1* disruptions, a gain-of-function approach to reintroduce this genetic information into knockout cells should be used. In preliminary experiments, I transiently transfected cDNAs encoding either *Cmas* or *Slc35a1* into Δ *Cmas* or Δ *Slc35a1* cells. After allowing 24 h for protein expression, I evaluated virus binding to these cells (Figure 13). Although lectin binding was restored (Figure 9), virus binding was not restored to levels comparable to WT BV2 cells. While I continue to optimize and troubleshoot the conditions for these experiments, some conclusions can be drawn from these data. It is possible that the introduction of *Cmas* or *Slc35a1* cDNA does not fully restore functionality to the SA synthesis pathway; therefore, the SAs that are expressed on the cell surface of these complemented cells are not structurally sufficient or

functional. While the lectin binding experiments are not entirely consistent with this conclusion, it could be that WGA displays less specificity for SA binding than T3 reoviruses. This hypothesis could be tested by introducing cDNAs encoding either *Cmas* or *Slc35a1* into $\Delta Cmas$ or $\Delta Slc35a1$ cells, allow 24 h for protein and SA expression, and conduct a binding assay using T1L, a reovirus strain capable of binding only $\alpha 2,3$ -linked *N*-acetyl-SA. I would anticipate observing results mirroring those observed previously for T3SA+ in which virus binding was not restored. However, if a restoration of virus binding is observed, such a result would indicate that not all types of SAs are being restored and expressed on the cell surface. Lastly, establishing stable cell lines of $\Delta Cmas+Cmas$ or $\Delta Slc35a1+Slc35a1$ is an important next step to ensure a homogenous population of cells with restored *Cmas* or *Slc35a1* expression levels.

The level of SA detected by WGA binding on the cell surface of $\Delta Cmas$ or $\Delta Slc35a1$ cells is reduced when compared to WT cells. However, the mRNA and protein levels of *Cmas* and *Slc35a1* in the knockout cells should be quantified. Real-time quantitative PCR is a technique used to quantify a specific RNA by monitoring cDNA amplification using fluorescence. Using antibodies specific for *Cmas* or *Slc35a1*, protein levels can be quantified by flow cytometry or immunoblotting assays. These techniques would provide a more precise and quantitative assessment of the levels of *Cmas* and *Slc35a1* expressed by WT, CRISPR knockout, and complemented BV2 cells.

JAM-A and NgR1 expression on the surface of BV2 cells was undetectable by immunoblotting. To further explore the presence of these two known reovirus receptors, the mRNA and protein levels of *JAM-A* and *NgR1* should be quantified. Using antibodies specific for JAM-A or NgR1, protein levels on WT BV2 cells could be detected and quantified by flow cytometry. I anticipate that JAM-A and NgR1 protein levels will be undetectable on BV2 cells,

supporting the hypothesis that SA serves as a reovirus receptor on microglial cells. However, if JAM-A and NgR1 expression is observed using flow cytometry, then both receptors are indeed expressed on the surface of BV-2 cells, perhaps in low enough amounts to be undetected by immunoblotting. An interesting follow-up experiment would be to incubate WT BV2 cells with JAM-A- and NgR1-specific antibodies prior to a virus binding assay using T3SA+. I would anticipate observing no change in the level of virus binding, which would support the hypothesis that SA acts as a reovirus receptor on cells with limited cell-surface expression of JAM-A or NgR1. However, if a decrease in binding is observed, then I would conclude that JAM-A and NgR1 are expressed on the surface of BV2 cells, perhaps in low abundance, and SA is serving as an attachment factor rather than a receptor.

The work presented in this thesis has identified, evaluated, and begun to validate *Cmas* and *Slc35a1* as two important host genes required for reovirus infection of microglial cells. These genes and their protein products play an important role in the early steps of reovirus infection. Disruption of these genes in microglial cells results in a decreased abundance of SA at the cell surface and decreased reovirus binding and infectivity relative to WT microglial cells. These studies advance current knowledge of reovirus replication and identify two host genes that mediate infection of neural cells by T3 reoviruses, potentially providing therapeutic targets to limit reovirus neuropathogenesis.

Bibliography

1. Kauskot, A., et al., *A mutation in the gene coding for the sialic acid transporter SLC35A1 is required for platelet life span but not proplatelet formation*. *Haematologica*, 2018. **103**(12): p. e613-e617.
2. Sutherland, D.M., *Functions of the Viral Attachment Protein in Reovirus Neurovirulence*, in *Microbiology and Immunology*. 2018, Vanderbilt University: Nashville, TN. p. 120.
3. Kannan, R. and A. Ventura, *The CRISPR revolution and its impact on cancer research*. *Swiss Med Wkly*, 2015. **145**: p. w14230.
4. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis*. *J Comput Chem*, 2004. **25**(13): p. 1605-12.
5. Martin, S., et al., *A genome-wide siRNA screen identifies a druggable host pathway essential for the Ebola virus life cycle*. *Genome Med*, 2018. **10**(1): p. 58.
6. Konopka-Anstadt, J.L., et al., *The Nogo receptor NgR1 mediates infection by mammalian reovirus*. *Cell Host Microbe*, 2014. **15**(6): p. 681-91.
7. Knowlton, J.J., et al., *The TRiC chaperonin controls reovirus replication through outer-capsid folding*. *Nat Microbiol*, 2018. **3**(4): p. 481-493.
8. Timms, R.T., et al., *Haploid genetic screens identify an essential role for PLP2 in the downregulation of novel plasma membrane targets by viral E3 ubiquitin ligases*. *PLoS Pathog*, 2013. **9**(11): p. e1003772.
9. Carette, J.E., et al., *Ebola virus entry requires the cholesterol transporter Niemann-Pick C1*. *Nature*, 2011. **477**(7364): p. 340-3.
10. Doench, J.G., *Am I ready for CRISPR? A user's guide to genetic screens*. *Nat Rev Genet*, 2018. **19**(2): p. 67-80.
11. Gebre, M., J.L. Nomburg, and B.E. Gewurz, *CRISPR-Cas9 genetic analysis of virus-host interactions*. *Viruses*, 2018. **10**(2).
12. Zhang, R., et al., *Mxra8 is a receptor for multiple arthritogenic alphaviruses*. *Nature*, 2018.
13. Orchard, R.C., et al., *Discovery of a proteinaceous cellular receptor for a norovirus*. *Science*, 2016. **353**(6302): p. 933-6.
14. Evers, B., et al., *CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes*. *Nat Biotechnol*, 2016. **34**(6): p. 631-3.
15. Ran, F.A., et al., *In vivo genome editing using Staphylococcus aureus Cas9*. *Nature*, 2015. **520**(7546): p. 186-91.
16. Varki, A. and R. Schauer, *Sialic Acids*, in *Essentials of Glycobiology*, A. Varki, et al., Editors. 2009, Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY).
17. No, D., et al., *Novel GNE mutations in autosomal recessive hereditary inclusion body myopathy patients*. *Genet Test Mol Biomarkers*, 2013. **17**(5): p. 376-82.
18. Carrillo, N., M.C. Malicdan, and M. Huizing, *GNE Myopathy: Etiology, diagnosis, and therapeutic challenges*. *Neurotherapeutics*, 2018. **15**(4): p. 900-914.
19. Martinez, N.N., et al., *Sialuria: Ninth patient described has a novel mutation in GNE*. *JIMD Rep*, 2019. **44**: p. 17-21.
20. Matsuura, R., et al., *First patient with Salla disease confirmed by genomic analysis in Japan*. *Pediatr Neurol*, 2018. **81**: p. 52-53.

21. Barton, E.S., et al., *Junction adhesion molecule is a receptor for reovirus*. Cell, 2001. **104**(3): p. 441-451.
22. Sturzenbecker, L.J., et al., *Intracellular digestion of reovirus particles requires a low pH and is an essential step in the viral infectious cycle*. J Virol 1987. **61**(8): p. 2351-2361.
23. Ebert, D.H., et al., *Cathepsin L and cathepsin B mediate reovirus disassembly in murine fibroblast cells*. J Biol Chem, 2002. **277**(27): p. 24609-17.
24. Chandran, K. and M.L. Nibert, *Animal cell invasion by a large nonenveloped virus: reovirus delivers the goods*. Trends Microbiol, 2003. **11**(8): p. 374-82.
25. Coombs, K.M., *Stoichiometry of reovirus structural proteins in virus, ISVP, and core particles*. Virology, 1998. **243**(1): p. 218-28.
26. Silverstein, S.C., et al., *The mechanism of reovirus uncoating and gene activation in vivo*. Virology, 1972. **47**(3): p. 797-806.
27. Banerjee, A.K. and A.J. Shatkin, *Transcription in vitro by reovirus-associated ribonucleic acid-dependent polymerase*. J Virol, 1970. **6**(1): p. 1-11.
28. Chang, C.T. and H.J. Zweerink, *Fate of parental reovirus in infected cell*. Virology, 1971. **46**(3): p. 544-55.
29. Li, J.K.-K., et al., *The plus strand of reovirus gene S2 is identical with its in vitro transcript*. Virology, 1980. **105**: p. 282-286.
30. Shatkin, A.J. and M. Kozak, *Biochemical aspects of reovirus transcription and translation*, in *The Reoviridae*, W.J. Joklik, Editor. 1983, Plenum Press: New York. p. 79-106.
31. Rosen, L., *Serologic grouping of reovirus by hemagglutination-inhibition*. American Journal of Hygiene, 1960. **71**: p. 242-249.
32. Lerner, A.M., et al., *Infections with reoviruses*. N Engl J Med, 1962. **267**: p. 947-52.
33. Tai, J.H., et al., *Prevalence of reovirus-specific antibodies in young children in Nashville, Tennessee*. J Infect Dis, 2005. **191**(8): p. 1221-4.
34. Weiner, H.L. and B.N. Fields, *Neutralization of reovirus: the gene responsible for the neutralization antigen*. J Exp Med, 1977. **146**(5): p. 1305-10.
35. Sabin, A.B., *Reoviruses: a new group of respiratory and enteric viruses formerly classified as ECHO type 10 is described*. Science, 1959. **130**: p. 1387-1389.
36. Breun, L.A., et al., *Mammalian reovirus L2 gene and lambda2 core spike protein sequences and whole-genome comparisons of reoviruses type 1 Lang, type 2 Jones, and type 3 Dearing*. Virology, 2001. **287**(2): p. 333-48.
37. Steyer, A., et al., *High similarity of novel orthoreovirus detected in a child hospitalized with acute gastroenteritis to mammalian orthoreoviruses found in bats in Europe*. J Clin Microbiol, 2013. **51**(11): p. 3818-25.
38. Reiss, K., et al., *The GM2 glycan serves as a functional co-receptor for serotype 1 reovirus*. PLoS Pathogens, 2012. **8**(12): p. e1003078.
39. Chappell, J.D., et al., *Identification of carbohydrate-binding domains in the attachment proteins of type 1 and type 3 reoviruses*. J Virol, 2000. **74**(18): p. 8472-9.
40. Reiter, D.M., et al., *Crystal structure of reovirus attachment protein sigma1 in complex with sialylated oligosaccharides*. PLoS Pathog, 2011. **7**(8): p. e1002166.
41. Barton, E.S., et al., *Utilization of sialic acid as a coreceptor enhances reovirus attachment by multistep adhesion strengthening*. J Biol Chem, 2001. **276**(3): p. 2200-2211.
42. Kobayashi, T., et al., *An improved reverse genetics system for mammalian orthoreoviruses*. Virology, 2010. **398**(2): p. 194-200.

43. Sutherland, D.M., et al., *Reovirus neurotropism and virulence are dictated by sequences in the head domain of the viral attachment protein*. J Virol, 2018. **92**(23): p. e00974-18.
44. Connolly, J.L., E.S. Barton, and T.S. Dermody, *Reovirus binding to cell surface sialic acid potentiates virus-induced apoptosis*. J Virol, 2001. **75**(9): p. 4029-39.
45. Barton, E.S., et al., *Utilization of sialic acid as a coreceptor is required for reovirus-induced biliary disease*. J Clin Invest, 2003. **111**(12): p. 1823-33.
46. Stencel-Baerenwald, J., et al., *Glycan engagement dictates hydrocephalus induction by serotype 1 reovirus*. MBio, 2015. **6**(2): p. e02356.
47. Furlong, D.B., M.L. Nibert, and B.N. Fields, *Sigma 1 protein of mammalian reoviruses extends from the surfaces of viral particles*. J Virol, 1988. **62**(1): p. 246-56.
48. Virgin, H.W., et al., *Antibody protects against lethal infection with the neurally spreading reovirus type 3 (Dearing)*. J Virol, 1988. **62**(12): p. 4594-604.
49. Mainou, B.A. and T.S. Dermody, *Transport to late endosomes is required for efficient reovirus infection*. J Virol, 2012. **86**(16): p. 8346-58.
50. Doench, J.G., et al., *Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9*. Nat Biotechnol, 2016. **34**(2): p. 184-191.
51. Dermody, T.S., J.S. Parker, and B. Sherry, *Orthoreoviruses*, in *Fields Virology*, D.M. Knipe and P.M. Howley, Editors. 2013, Lippincott Williams & Wilkins: Philadelphia. p. 1304-1346.
52. Schonberg, M., et al., *Asynchronous synthesis of the complementary strands of the reovirus genome*. Proc Natl Acad Sci U S A, 1971. **68**(2): p. 505-8.
53. Abeln, M., et al., *Sialylation is dispensable for early murine embryonic development in vitro*. Chembiochem, 2017. **18**(13): p. 1305-1316.
54. Schwarz, R.E., et al., *Wheatgerm agglutinin-mediated toxicity in pancreatic cancer cells*. Br J Cancer, 1999. **80**(11): p. 1754-62.
55. Marx, N., et al., *CRISPR-based targeted epigenetic editing enables gene expression modulation of the silenced Beta-Galactoside Alpha-2,6-Sialyltransferase 1 in CHO cells*. Biotechnol J, 2018. **13**(10): p. e1700217.
56. Han, J., et al., *Genome-wide CRISPR/Cas9 screen identifies host factors essential for Influenza virus replication*. Cell Rep, 2018. **23**(2): p. 596-607.
57. Lee, C.C., et al., *The role of N-alpha-acetyltransferase 10 protein in DNA methylation and genomic imprinting*. Mol Cell, 2017. **68**(1): p. 89-103 e7.
58. Chu, H.Y., et al., *Homology modeling and molecular dynamics study on N-acetylneuraminidase*. J Mol Model, 2009. **15**(3): p. 323-8.
59. Grove, J. and M. Marsh, *The cell biology of receptor-mediated virus entry*. J Cell Biol, 2011. **195**(7): p. 1071-82.
60. Goody, R.J., C.C. Hoyt, and K.L. Tyler, *Reovirus infection of the CNS enhances iNOS expression in areas of virus-induced injury*. Exp Neurol, 2005. **195**(2): p. 379-90.
61. Redka, D.S., et al., *Differential ability of proinflammatory and anti-inflammatory macrophages to perform macropinocytosis*. Mol Biol Cell, 2018. **29**(1): p. 53-65.
62. Frierson, J.M., et al., *Utilization of sialylated glycans as coreceptors enhances the neurovirulence of serotype 3 reovirus*. J Virol, 2012. **86**(24): p. 13164-13173.