Two Stories of Neurodevelopment: Exosome-Mediated Secretion of Sonic Hedgehog and Somatic Mutation in Disorders of DNA Damage Repair

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th>Citation</th>
<th>Coulter, Michael. 2016. Two Stories of Neurodevelopment: Exosome-Mediated Secretion of Sonic Hedgehog and Somatic Mutation in Disorders of DNA Damage Repair. Doctoral dissertation, Harvard University, Graduate School of Arts &amp; Sciences.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:33840698">http://nrs.harvard.edu/urn-3:HUL.InstRepos:33840698</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Two stories of neurodevelopment:

Exosome-mediated secretion of sonic hedgehog

and

Somatic mutation in disorders of DNA damage repair

A dissertation presented

by

Michael Edward Coulter

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Neurobiology

Harvard University

Cambridge, Massachusetts

June 2016
Two stories of neurodevelopment: Exosome-mediated secretion of sonic hedgehog and somatic mutation in disorders of DNA damage repair

Abstract

Human genetics of neonatal brain malformations has identified dozens of genes required for brain development that regulate diverse cellular processes. Recent evidence shows that somatic mutations, mutations that are only present in some cells of the body, can also cause brain malformations and disease. We have studied the function of a recently identified microcephaly and cerebellar hypoplasia gene, CHMP1A, and found that it is required for exosome-mediated secretion of sonic hedgehog (SHH). SHH is an essential growth factor in the developing brain, and our results reveal a novel mechanism for SHH secretion in the vertebrate brain. Somatic mutations are caused by incorrect repair of damaged DNA. Cockayne syndrome and xeroderma pigmentosum are disorders of DNA damage repair that cause microcephaly and early neurodegeneration. Using whole genome sequencing, we measured somatic mutation rate in post-mortem single cortical neurons from patients with these disorders and found a dramatic increase in the number of somatic mutations compared to normal individuals.
# Table of Contents

Abstract .............................................................................................................................................. iii

Table of Contents ............................................................................................................................... iv

List of Tables ....................................................................................................................................... v

List of Figures ..................................................................................................................................... vi

Acknowledgements ............................................................................................................................ viii

Chapter 1: Introduction – Brain malformations: from human genetics to molecular mechanisms ................................................................................................................................. 1
  1.1 Summary ..................................................................................................................................... 2
  1.2 Molecular mechanisms of brain development ........................................................................... 3
  1.3 Somatic mutation in brain development and degeneration ....................................................... 20
  1.4 Conclusion ................................................................................................................................. 33

Chapter 2: CHMP1A function reveals exosome-mediated secretion of sonic hedgehog in developing brain ...................................................................................................................... 34
  2.1 Summary ................................................................................................................................. 36
  2.2 Introduction ............................................................................................................................. 37
  2.3 Results .................................................................................................................................... 40
  2.4 Discussion ............................................................................................................................... 81
  2.5 Experimental Procedures ....................................................................................................... 91

Chapter 3: Somatic mutation in disorders of DNA damage repair ............................................. 104
  3.1 Summary ............................................................................................................................... 106
  3.2 Introduction ............................................................................................................................. 107
  3.3 Results ................................................................................................................................... 115
  3.4 Discussion ............................................................................................................................... 140
  3.5 Experimental Procedures ....................................................................................................... 150

Chapter 4: Discussion ....................................................................................................................... 154
  4.1 Summary ............................................................................................................................... 155
  4.2 CHMP1A is required for exosome-mediated SHH secretion .................................................. 156
  4.3 Neuronal somatic mutation is increased in disorders of DNA damage repair .................. 164
  4.4 Conclusion ............................................................................................................................... 168

References .......................................................................................................................................... 169
List of Tables

Table 1.1: Functions of sonic hedgehog during CNS and limb development ..................17
Table 1.2: Somatic mutations cause neurodevelopmental disorders ..........................21
Table 1.3: Somatic mutation rates from single cell WGS ........................................30
Table 2.1: Antibodies used in this study .................................................................103
Table 3.1: Post-mortem brains from patients with XP used in this study ....................116
Table 3.2: Post-mortem brains from CS patients used in this study ............................116
Table 3.3: Normal control brains used in this study ...............................................116
Table 3.4: Genome-wide coverage of single neurons analyzed in the study ..............119
Table 3.5: CS neuron truncating somatic mutations in disease genes ......................139
List of Figures

Figure 1.1: Cellular defects underlying brain malformations ............................................. 5
Figure 1.2: Model of exosome formation and release ......................................................... 11
Figure 1.3: Exosomes mediate Wg secretion at Drosophila NMJ ....................................... 13
Figure 1.4: Proposed mechanisms for specialized SHH secretion ....................................... 19
Figure 2.1: CHMP1A is essential for brain development in mice ........................................ 42
Figure 2.2: CHMP1A is essential for mouse postnatal survival and *Chmp1a* heterozygous mice are normal ........................................................................................................... 44
Figure 2.3: CHMP1A is essential for progenitor proliferation in human cerebral organoids .......................................................... 46
Figure 2.4: Disrupted gene expression in *CHMP1A* null cerebral organoids ...................... 48
Figure 2.5: CHMP1A is expressed in postmitotic neurons and ChP epithelial cells in developing brain .................................................................................................................. 51
Figure 2.6: Specificity of beta-gal and CHMP1A antibody staining ....................................... 53
Figure 2.7: CHMP1A is essential for cortex and dorsal telencephalon development .......... 57
Figure 2.8: Normal cell death and ventricular surface in *Chmp1a* null cortex ................. 58
Figure 2.9: MRI imaging of patients with *CHMP1A* null mutations ................................. 59
Figure 2.10: CHMP1A is essential for mouse cerebellum development and is required for SHH-mediated GCP proliferation ................................................................. 62
Figure 2.11: *Chmp1a* null cerebellum has normal number of Purkinje cells and downstream Shh activation is normal in the absence of CHMP1A ................................. 63
Figure 2.12: CHMP1A is essential for exosome-mediated SHH secretion ....................... 66
Figure 2.13: CHMP1A and SHH are present on the surface of extracellular vesicles ..... 68
Figure 2.14: CHMP1A is not required for cytokinesis or EGFR degradation.................69
Figure 2.15: Choroid plexus epithelial cell MVBs .................................................71
Figure 2.16: Choroid plexus epithelial surface is normal in Chmp1a null mice ..........72
Figure 2.17: Exosome are released from hindbrain choroid plexus .......................74
Figure 2.18: CHMP1A is essential for exosome release in hindbrain choroid plexus .....76
Figure 2.19: Exosomes in cerebellum and cortex ...................................................78
Figure 2.20: MVBs localize to dendrites in cerebellar Purkinje cells ......................80
Figure 3.1: Brain MRI of a Cockayne syndrome patient..................................110
Figure 3.2: Single neuron somatic SNV identification protocol ............................117
Figure 3.3: WGS coverage of brains used in this study ......................................120
Figure 3.4: Quality control measurements of neurons used in this study ...............123
Figure 3.5: Circos plot of illustrative Cockayne syndrome single neurons ..............125
Figure 3.6: Somatic mutation rates in CS and control neurons ............................127
Figure 3.7: Alternate allele frequency of somatic SNVs in CS and control neurons .....130
Figure 3.8: Strand bias of somatic SNVs from transcribed genomic regions ..........133
Figure 3.9: Somatic SNV substitution types in CS and control neurons .................134
Figure 3.10: Mutational signatures of SNVs from CS and control neurons ..........136
Acknowledgments

I would like to thank my adviser Chris Walsh for his support and guidance throughout my dissertation work. Chris is an incredible scientific mentor who leads by example. He cares about the career success and personal fulfillment of each one of the students and post-docs in his lab. He often shares insightful advice with us about navigating life outside the lab including grant writing, paper submission, and interaction with other scientists in the field. He also trusts each of member of the lab to follow his or her own passion and provides generous resources for each person’s creative scientific expression. During my PhD, Chris’ encouragement was crucial for completing my work studying CHMP1A although at times the results were discouraging or confusing. He also encouraged me to take on my second project, somatic mutation in DNA damage repair disorders, to ensure I had a successful graduate school experience. At the end of my PhD, thanks to my wonderful experience under Chris’ direction, I would eagerly do it again and am excited to continue my career in neuroscience.

I would also like to thank the members of the Walsh lab who have provided excellent advice and assistance during my time in the lab. The Walsh lab is a wonderful environment for scientific investigation because everyone is helpful, knowledgeable, and provides constructive criticism. I would like to thank my collaborators for contributing to the projects I will describe in this thesis. I would like to thank the DMS office, the PiN office, and the MD-PhD office for supporting me throughout my PhD.

Finally, I thank my family and friends for providing continuous support and encouragement throughout my life and especially during the last several years in graduate school.
To my parents, Anne and John,

and to my sister, Deborah
Chapter 1: Introduction

Brain malformations: from human genetics to molecular mechanisms
1.1 Summary

In this chapter, I will review recent progress towards understanding genetic mechanisms of brain development and highlight some open questions in the field. I have organized this introduction into two sections to reflect the projects that comprise my dissertation research: (1) CHMP1A’s role in a novel mechanism for sonic hedgehog secretion and (2) neuronal somatic mutation in disorders of DNA damage repair. Section 1.2 will review how genes identified in human genetic brain malformation studies both confirm current models of brain development and uncover new and unexpected mechanisms. This section will continue with specific introductions to exosome biology and sonic hedgehog function and secretion. Section 1.3 will highlight the role of somatic mutations in genetic brain malformations and then highlight DNA damage, the force behind somatic mutation. This section will continue by introducing inherited disorders of DNA damage repair and recent advances in single cell sequencing technology that enable us to answer new questions about the role of somatic mutation in brain development and degeneration.
1.2 Molecular mechanisms of brain development

An important finding from many years of identification of brain malformation and microcephaly proteins is that these proteins affect a vast array of different biological processes in the cell. Indeed, collecting families with brain malformations is an unbiased screen for genes required for brain development and the study of identified genes often uncovers new and unexpected mechanisms required for development of the brain. The first project of my thesis work focused on identifying the mechanism of a recently discovered microcephaly and cerebellar hypoplasia gene, charged multivesicular body protein 1A (CHMP1A). I have found that CHMP1A regulates a new and unexpected mechanism of brain development: exosome-mediated sonic hedgehog secretion.

Microcephaly proteins reveal molecular mechanisms of brain development

Identification of brain malformation genes through human genetics has been a fruitful endeavor since the 1990s when the first genes, \textit{LIS1}, \textit{DCX}, and \textit{FLNA} were identified (Reiner et al., 1993; Gleeson et al., 1998; Pilz et al., 1998; Portes et al., 1998). In the past two decades, dozens of brain malformation genes have been identified thanks to rapid advances in the application of human genetics (Faheem et al., 2015; Hu et al., 2014a).

These discoveries have impacted both clinical medicine and developmental neuroscience. Families in high-risk populations receive genetic screening for recurrent disease alleles in brain malformation genes and families with identified mutations can undergo prenatal testing to identify affected pregnancies. In addition, further
Investigation of brain malformation proteins has revealed several fundamental cellular and molecular mechanisms of cerebral cortex development.

Brain malformations include microcephaly, cerebellar hypoplasia, lissencephaly, double cortex syndrome, periventricular nodular heterotopia, focal cortical dysplasia, and hemimegalencephaly, among others. Identification of genetic causes of these malformations has revealed genes that regulate neural progenitor proliferation, postmitotic neuronal migration, and neuronal cell size (Figure 1.1). Starting at mid-gestation, the cerebral cortex develops as follows: (1) progenitor cell proliferation in the neuroepithelium surrounding the lateral ventricles, (2) differentiation of these progenitors into neurons, (3) migration of newborn neurons outward to the cortical plate, (4) birth and migration of glia, and finally (5) development of synaptic connections between neurons within the cortex and between cortical neurons and thalamus, cerebellum, and spinal cord neurons (Grieg et al., 2013). Cortical malformations fit well within this model of cortical development. Disorders of progenitor proliferation include microcephaly (small head and small cortex) and cerebellar hypoplasia (small cerebellum). Disorders of cortical neuron migration include lissencephaly (defined as a smooth cortex with few or no gyri and sulci), double cortex syndrome (defined by a band of neurons beneath the normal cortex, representing cells that fail to migrate from the ventricular zone all the way to the cortical plate) and periventricular nodular heterotopia (small clusters of neurons that never migrate from the ventricular surface).
Figure 1.1 Cellular defects underlying brain malformations

Different cortical malformations are caused by different cellular defects including progenitor proliferation, neuron migration, and neuronal connectivity (Hu et al., 2014).

This thesis will focus on a cause of microcephaly and cerebellar hypoplasia whose mechanism of action identifies new pathways not previously implicated by other brain malformation genes. Before discussing it, I will first highlight a few other microcephaly and cerebellar hypoplasia genes to illustrate the remarkable breadth of molecular mechanisms of brain development that have been identified through brain malformation genetics.
Molecular mechanisms of microcephaly proteins

Primary microcephaly, defined as head circumference of 4 SD or more below the age-adjusted mean, is one of the most prevalent neonatal brain malformations. The incidence of microcephaly ranges from 1.3/100,000 to 1/10,000 depending on the population being studied. The incidence is highest in consanguineous populations because inherited single-gene recessive mutations are a common cause of microcephaly. Patients with microcephaly have a small cerebral cortex and may also have other structural brain abnormalities such as agenesis or thinning of the corpus callosum, cerebellar hypoplasia, and brain stem hypoplasia. Microcephaly typically causes intellectual disability and physical impairment (delayed or inability to walk, for example) and severe microcephaly is often fatal in the first years of life (Faheem et al., 2015).

Proteins encoded by microcephaly genes regulate cerebral cortex development primarily through their function in cortical progenitor cells. Cortical progenitors undergo a finite number of cell divisions during embryonic development to generate all post-mitotic neurons (Greig et al., 2013). The proliferative capacity in the cortex is very tightly controlled: in mice, each cortical progenitor gives rise to 8-9 postmitotic neurons (Gao et al., 2014). If cortical progenitor proliferation is disrupted by any mechanism, fewer neurons are born and microcephaly results.

Microcephaly proteins regulate diverse processes in cortical development including progenitor proliferation and survival, newborn neuron migration into the cortical plate, and organization of the ventricular surface (Figure 1.1). Many proteins encoded by microcephaly genes localize to the centrosome, the cell’s microtubule organizing center that becomes the spindle pole during mitosis (Alkuraya et al., 2011;
Disruption of centrosome function prevents normal progression through cell division and often increases progenitor apoptosis. Together, these defects decrease the size of the cortical progenitor pool and, as a result, fewer neurons are born and cortical size is decreased. However, some microcephaly proteins do not localize to the centrosome; these include ZNF335, a nuclear protein that regulates gene expression (Yang et al., 2012), KIF5C and KIF2A, proteins that regulate the cytoskeleton (Poirier et al., 2013), QARS and VARS, tRNA synthetases (Karaca et al., 2015; Zhang et al., 2014), and PYCR2 and ASNS, enzymes in the proline and asparagine biosynthesis pathways, respectively (Nakayama et al., 2015; Ruzzo et al., 2013). Although non-centrosomal microcephaly proteins do not regulate mitosis directly, they are required for progenitor proliferation and survival and for post-mitotic neuron survival, and when absent, reduce the final number of post-mitotic neurons.

**Cerebellar hypoplasia proteins**

Another highly genetic brain malformation, cerebellar hypoplasia, can occur with microcephaly and patients with this disorder have similar symptoms, including intellectual disability, inability to walk, difficulty with speech, and seizures (Maricich et al., 2011). Proliferation of granule cells, the most abundant cell type in the cerebellum, is driven primarily by Purkinje cell secretion of sonic hedgehog (SHH), which acts as a mitogen to drive massive proliferation of granule cell precursors (GCPs). GCPs then differentiate into granule neurons, which migrate inward and form synaptic connections.
with Purkinje cells and deep cerebellar nuclei to form the cerebellar circuit (Roussel and Hatten, 2011). Mutations that disable cerebellar hypoplasia proteins disrupt GCP proliferation and survival through regulation of gene expression, RNA processing, protein synthesis, and autophagosome function. Mutations in CASK have been identified in patients with microcephaly and cerebellar hypoplasia. CASK translocates to the nucleus to interact with the transcription factor TBR1, which increases expression of TBR1 target genes such as RELN and CTIP1 that are expressed in cortical and cerebellar neurons. CASK is thus required for normal brain development because it promotes neuron survival through regulation of gene expression (Najm et al., 2008). Some patients with cerebellar hypoplasia and atrophy have mutations in EXOSC3, a component of the RNA exosome complex, an enzyme complex that processes and degrades mRNA with 3’-5’ exonuclease activity (Lubas et al., 2015). These patients have dysmorphic Purkinje cells and loss of granule cells, and partial loss of exosome complex function suggesting that disruption of RNA processing reduces neuron survival and progenitor proliferation (Wan et al., 2012). Mutations have been identified in SNX14 in patients with cerebellar hypoplasia (Akizu et al., 2015; Thomas et al., 2014). Mutations in SNX14 cause Purkinje cell loss and decrease the number of post-mitotic granule cells. The authors provide evidence that SNX14 is required for autophagosome function and disruption of this process leads to cell death (Akizu et al., 2015). Mutations in three tRNA transferases (TSEN54, TSEN2, and TSEN34) were found in families with pontocerebellar hypoplasia (Budde et al., 2008). These mutations may disrupt binding between tRNA splicing endonuclease components and cause a partial loss of protein synthesis. These findings suggest that the developing cerebellum and brainstem are particularly sensitive to
decreased protein synthesis. Although genetic causes of cerebellar hypoplasia do not always directly affect mitosis, many cause partial loss of function of essential cellular functions, which likely disrupt development by increasing cell death of Purkinje cells, GCPs, and post-mitotic neurons.

Cerebellar hypoplasia is also a feature of ciliopathies, syndromic disorders caused by defects of the primary cilium. In particular, cerebellar vermis hypoplasia is a common feature of Joubert, Bardet–Biedl, and Meckel–Gruber syndromes (Goetz and Anderson, 2010; Waters and Beales, 2011), disorders caused by mutations in ciliary proteins. Cerebellar hypoplasia in these disorders is likely a result of decreased GCP proliferation because SHH, the mitogen that drives GCP mitosis, activates its signaling cascade at the primary cilium in responsive cells. In support of this hypothesis, there is evidence of disrupted Shh signaling in cerebellum samples from Joubert and Meckel syndrome patients (Aguilar et al., 2012; Roosing et al., 2015).

Microcephaly and cerebellar hypoplasia proteins reveal several molecular mechanisms of brain development. Some of these, such as centrosome dysfunction impairing cell division or defective Shh signaling in ciliopathies, fit well with the current models of cortex and cerebellum development. However, some of these mechanisms are unexpected, such as defects in amino acid synthesis or RNA exosome complex. These unexpected findings highlight the idea that continued identification and characterization of new brain malformation proteins will reveal unanticipated biologic processes in brain development. For this reason, human genetics of brain malformation will continue to be of great interest to developmental neuroscience in the future. The first project of my dissertation research studied a microcephaly gene, *CHMP1A*, and revealed that it
regulates brain development though a novel and unexpected mechanism: exosome-mediated secretion of sonic hedgehog.

**Exosome biology: a different kind of cell-to-cell communication**

Cell-to-cell communication shapes development and regulates adult function in multicellular organisms. Classically, a soluble ligand is released by a secreting cell and then travels through the extracellular environment to bind its receptor on a receiving cell, which activates a transmembrane signaling cascade that can drive cell proliferation, cell fate change, cell migration, process extension or retraction, or formation of cell-cell contacts. Recently, it has been shown in several contexts that secreted molecules can be released inside or embedded in the membrane of small membrane-bound vesicles released by the secreting cells. These extracellular vesicles (EV) travel through the extracellular environment and deliver their cargo to receiving cells. In the 1980s, it was discovered that some EVs are released when endosomes fuse with the cell membrane and release preformed intracellular endosomal vesicles. The term exosome was coined in 1987 to describe this class of EVs. In the past 30 years, biologists have learned a great deal about the identity and origin of exosomes and the vast array of functions they serve during development and in adult life (Bobrie et al., 2012).

Exosomes are 50-100 nm diameter membrane bound vesicles released by endosomal compartment fusion with the cell membrane. They contain characteristic proteins that illustrate their origin in the cell. There is consensus that exosomes contain at least two markers proteins, CD63 (a tetraspanin) and TSG101. CD63 is a transmembrane protein that localizes to endosomal compartments in the cell called
multivesicular bodies (MVBs) (Heijnen et al., 1998). MVBs are large endosomes that are filled with membrane-bound vesicles called intraluminal vesicles (ILVs). When an MVB fuses with the cell membrane, ILVs are released into the environment and become exosomes (Figure 1.2, Cocucci and Meldolesi, 2015). During MVB biogenesis, ILVs are formed when a small section of the MVB limiting membrane is pinched off into the MVB lumen. Formation of ILVs requires endosomal sorting complex required for transport (ESCRT) proteins. Four ESCRT complexes coordinate to: recruit cargo, distort the limiting membrane, and then pinch off the newly formed ILV (McCullough et al., 2013). TSG101 is a member of ESCRT-I complex (Lötvall et al., 2014). Thus, the presence of CD63 and TSG101 on purified exosomes shows that these vesicles originate in MVBs; this property defines exosomes and differentiates them from other EVs.

![Exosome Formation Diagram](image)

**Figure 1.2: Model of exosome formation and release**

Exosomes are formed as ILVs in the MVB and are released when MVBs fuse with the cell membrane (Colombo et al., 2014).
Cells from many different tissues secrete exosomes, and their function has been studied in several contexts. Exosomes are secreted by lymphocytes and macrophages of the immune system, by cells in the heart and blood vessels in the context of coagulation, thrombosis, and angiogenesis (Sheldon et al., 2010), and by neurons and glia in the central nervous system. Exosomes have also been studied in disease and found to have potential roles in pathogenesis of diabetes and cardiovascular disease (Khan et al., 2015; Pironti et al., 2015), inflammation, cancer (Hoshino et al., 2015), and neurodegeneration (Basso and Bonetto, 2016). Here, I will focus on recent findings that highlight exosome functions in the nervous system.

The majority of work showing the functions of exosome secretion in the nervous system comes from Vivian Budnik’s lab. They have extensively studied exosome release in Drosophila neuromuscular junction (NMJ) and found that exosomes regulate two independent functions, WNT secretion and the retrograde signal synaptotagmin 4 (Syt4) (Figure 1.3). Wg (Drosophila homolog of WNT) signaling is required at the larval NMJ to coordinate pre- and post-synaptic expansion of the synapse during development. The Budnik lab has shown that Wg and its carrier protein Evi are released on exosomes at the NMJ. EM imaging shows that MVBs fuse with the motorneuron bouton membrane and that Evi positive exosomes are present in the sub-synaptic reticulum (SSR) (Koles et al., 2012). Wg receptor, DFz2 is located in the SSR far from the bouton and so Wg and Evi travel on exosomes to bind to DFz2. In addition, they show that in cultured cells, exosome secreted Evi and Wg can be transferred between cells, showing that exosome secretion mediates cell-to-cell communication (Korkut et al., 2009). Retrograde signaling modulates the synaptic strength of the NMJ. Upon electrical stimulation, the
muscle cell secretes Syt4, a transmembrane synaptic protein, which binds to the motor neuron bouton and modulates synaptic strength. They show that Syt4, which originates in the motor neuron, is secreted on motor neuron exosomes and then taken up by the muscle cell (Korkut et al., 2013). Together, their data show that exosome-mediated release of growth factors and synaptic proteins regulates neurodevelopment and synaptic plasticity in the *Drosophila* nervous system.

**Figure 1.3: Exosomes mediate Wg secretion at Drosophila NMJ**

Exosomes loaded with Evi and Wg are released from the motorneuron bouton into the SSR where Wg receptors (DFz2) are located (Budnik et al., 2016).

In contrast to *Drosophila*, less is known about exosome function in the mammalian CNS. In mammals, there is evidence that both neurons and glia secrete exosomes and that they may regulate synaptic strength, cell repair and cell survival (Budnik et al., 2016). When depolarized, cultured neurons release exosomes containing cargo such as AMPA receptor subunits L1CAM, or miRNAs. These exosomes can be taken up by other neurons and may function to modulate synaptic strength or cell adhesion.
Oligodendrocytes also release exosomes carrying proteolipid protein, myelin, and oxidative stress protection proteins in response to neuron-glia interactions. Exosome release requires calcium entry through oligodendroglial NMDA and AMPA receptors and is stimulated by neurotransmitter release from depolarized neurons. These exosomes can be taken up by neurons and protect neurons from cell death during environmental stress such as oxygen and glucose deprivation. When added to cultured neurons, oligodendrocyte exosomes can also increase the neuron’s firing rate (Budnik et al., 2016). In addition, microglia use exosomes to secrete WNT3A and a lactate transporter (Gross et al., 2012; Zhang and Wrana, 2014). Although many functions have been suggested for nervous system exosomes from vertebrate neuron and glia cultures, to date, none of these functions have been shown in vivo. My in vivo and in vitro results in Chapter 2 show a new role for exosomes in the mammalian CNS: regulation of sonic hedgehog secretion.

**Sonic hedgehog: a global driver of CNS development**

Secreted proteins shape development by enabling cell-to-cell communication. Sonic hedgehog (SHH) is a secreted growth factor that has many roles during development including functions as a morphogen, a mitogen, an axon guidance molecule, and a regulator of synapse formation. SHH signals through two transmembrane proteins, Patched (PTCH) and Smoothened (SMO) and three transcription factors, GLI1, GLI2 and GLI3. GLI1 is an activator, GLI3 is a repressor, and GLI2 can be an activator or repressor. SHH initiates its signaling cascade by binding to its receptor PTCH on the membrane of receptive cells. SHH-PTCH binding relieves PTCH inhibition of the transmembrane effector SMO, which then activates the GLI transcription factor cascade.
When SHH is absent, GLI1 expression is low and PTCH inhibits SMO, which allows proteasome cleavage of GLI2 and GLI3 into C-terminal deleted repressor proteins. When SHH is present, it binds to PTCH at the primary cilium and relieves inhibition of SMO. This inhibits cleavage of GLI2 and GLI3 and full-length GLI2 protein functions as an activator that drives expression of GLI1. GLI1 then acts as a strong activator driving further expression of GLI1, PTCH, and context dependent target genes (Ruiz i Altaba et al., 2007). In addition to PTCH, SHH has another receptor, BOC, which also activates the signaling cascade. BOC can act synergistically with PTCH as well as independently (Mille et al., 2014).

SHH has several functions in the developing CNS (Table 1.1). It first acts in the ventral (Hensen’s) node to create lateral asymmetry in the developing embryo. SHH is secreted from cells on the left side of node and activates downstream targets in cells on the right side; this process helps define left and right in the embryo (Tanaka et al., 2005). Next in development, SHH expression in the notochord and the floor plate of the neural tube acts as a morphogen to direct cell fate in the ventral neural tube and create cell type specific dorsal-ventral organization of the developing spinal cord and inner ear (Roelink et al., 1995; Brown and Epstein, 2011; Martí et al., 1995; Riccomagno et al., 2002). In parallel, SHH expression at the midline is required for midline crossing of commissural axons from newly born spinal cord neurons (Bourikas et al., 2005; Charron et al., 2003; Okada et al., 2006; Wilson and Stoeckli, 2013). Also at this time, SHH is expressed by posterior cells in the limb bud for specification of posterior digits (Chiang et al., 2001; Lewis et al., 2001; Li et al., 2006). At mid-gestation, SHH continues to shape the CNS. Expressed from the floor of the third ventricle, SHH acts as a morphogen and mitogen to
drive formation and growth of the MGE and LGE of the ventral telencephalon, which eventually gives rise to the basal ganglia (Huang et al., 2007). In the dorsal telencephalon, very recent evidence also shows that Shh signaling drives proliferation of intermediate progenitors and basal radial glia during cortical neurogenesis (Wang et al., 2016). In the fourth ventricle, SHH expressed by choroid plexus epithelial cells mediates proliferation of pericytes in the hindbrain choroid plexus (Huang et al., 2009; Lun et al., 2015; Nielsen and Dymecki, 2010). Starting in late embryogenesis, SHH acts as a mitogen for cortical SVZ progenitors (Palma, 2005; Traiffort et al., 1999) and for granule cell precursors, progenitor cells in the developing cerebellum (Dahmane and Ruiz i Altaba, 1999; Kenney and Rowitch, 2000; Wallace, 1999; Wechsler-Reya and Scott, 1999). In these roles, Shh mRNA is expressed in the SVZ and cerebellar Purkinje cells, respectively. SHH reprises its role as an axon guidance molecule in early postnatal life in the olfactory bulbs where it is expressed by mitral tufted cells (Gong et al., 2009) and then takes on a new function as a regulator of projection neuron synapse formation in the cortex when it is expressed by layer V corticofugal projection neurons in the second week of life (Harwell et al., 2012). Shh signaling is less well understood in adult CNS but can be pathologically activated in medulloblastoma, a granule cell tumor of the cerebellum (Eisner et al., 2015; Goodrich et al., 1997).
<table>
<thead>
<tr>
<th>Target tissue</th>
<th>SHH source</th>
<th>Age</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lateral asymmetry</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ventral node - L</td>
<td>Ventral node - R</td>
<td>E8</td>
<td>Tanaka, 2005</td>
</tr>
<tr>
<td><strong>Cell fate specification</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ventral neural tube</td>
<td>Notocord, floor plate</td>
<td>E8.5 - 10.5</td>
<td>Roelink, 1995; Marti, 1995</td>
</tr>
<tr>
<td>Cochlea</td>
<td>Notocord, floor plate</td>
<td>E9.5 - 10.5</td>
<td>Riccomagno, 2002; Brown and Epstein, 2011</td>
</tr>
<tr>
<td>Limb bud</td>
<td>Posterior cells</td>
<td>E10</td>
<td>Chiang 2001; Lewis, 2001; Li, 2006</td>
</tr>
<tr>
<td>Ventral telencephalon</td>
<td>3rd ventricle floor plate</td>
<td>E11.5 - 13.5</td>
<td>Haung 2007</td>
</tr>
<tr>
<td><strong>Progenitor proliferation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choroid plexus pericytes</td>
<td>Choroid plexus epithelial cells</td>
<td>E12 - 18.5</td>
<td>Nielsen, 2010; Huang 2009; Lun, 2015</td>
</tr>
<tr>
<td>Cerebellar granule cell precursors</td>
<td>Purkinje cells</td>
<td>E18.5 - P8</td>
<td>Dahmane, 1999; Kenney, 2000; Wallace, 1999; Wechleser-Reya, 1999</td>
</tr>
<tr>
<td>Prenatal cortical progenitors</td>
<td>Progenitors or neurons</td>
<td>E15.5 - 18.5</td>
<td>Palma, 2004; Wang, 2016</td>
</tr>
<tr>
<td>Postnatal SVZ progenitors</td>
<td>SVZ</td>
<td>P3-adult</td>
<td>Traiffort, 1999; Palma 2005</td>
</tr>
<tr>
<td><strong>Tumor cell proliferation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medulloblastoma cells</td>
<td>Unknown</td>
<td>2 - 9 months</td>
<td>Goodrich, 1997; Hahn 1998; Eisner, 2015</td>
</tr>
<tr>
<td><strong>Axon guidance</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spinal cord midline</td>
<td>Commisural axons</td>
<td>E11.5</td>
<td>Charron, 2003; Okada, 2006; Bourikas, 2005; Wilson, 2013</td>
</tr>
<tr>
<td>Olfactory glomeruli</td>
<td>Mitral and tufted cells</td>
<td>P2</td>
<td>Gong, 2009</td>
</tr>
<tr>
<td><strong>Circuit formation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axons of collosal and local projection neurons</td>
<td>Corticofugal projection neuron dendrites</td>
<td>P4 - 14</td>
<td>Harwell, 2012</td>
</tr>
</tbody>
</table>
**Sonic hedgehog requires specialized secretion**

Sonic hedgehog signaling begins with SHH secretion from producing cells, but SHH is a cholesterol-modified protein that is hydrophobic and so cannot simply be released into the aqueous extracellular space. Consequently, SHH requires a specialized secretion mechanism (Figure 1.4). Several specialized mechanisms for Shh secretion have been identified in invertebrates, but little is known about SHH secretion in vertebrate development. Studies of *Drosophila* imaginal disk, *Drosophila* embryos and *C. elegans* (Liégeois et al., 2006) have shown that Hh can be secreted as oligomeric complexes, in clusters, in lipoprotein particles, and in exosomes (Briscoe and Thérond, 2013; Thérond, 2012). Of note, recent work showed that Hh is secreted on exovesicles in the *Drosophila* imaginal disk during wing development (Matusek et al., 2014).

Exovesicles are large (100-400 nm) membrane bound vesicles that bud from the plasma membrane, in contrast to exosomes, which are released upon MVB fusion. In vertebrate development, SHH secretion has only been examined in the embryonic mouse ventral node (Tanaka et al., 2005). In early mouse embryonic development, left-right asymmetry is established by the flow of extraembryonic fluid over the surface of the ventral node. This flow is required for heart looping and left-right localization of visceral organs. In E8 embryos, it was found that the right side of the node releases large membrane-bound particles (300-500 nm in diameter) that contain SHH and retinoic acid and activate downstream signaling in the left side of the node (Tanaka et al., 2005). These particles provide a mechanism for SHH secretion in the ventral node during early embryonic development. Recent *in vitro* experiments show that SHH can be secreted on exosomes in vertebrates (Vyas et al., 2014). Specifically, SHH secreted from transfected HEK293T
cells was detected in two populations of purified exosomes. In addition, SHH was isolated in the same exosome populations from the supernatant of cultured chick notochord (HH stage 21-24), raising the possibility that exosome mediated SHH may occur in vivo (Vyas et al., 2014).

Whereas specialized SHH secretion has been well studied in invertebrates and some work has been done in early embryonic vertebrate development, an important unanswered question in the field involves the mechanism of SHH secretion during brain development. As detailed above, SHH signaling has several essential functions during CNS development and so identifying the mechanisms for Shh secretion in the brain is an important priority. Chapter 2 of this thesis presents a novel mechanism for SHH secretion in the developing CNS by study of CHMP1A, a microcephaly and cerebellar hypoplasia gene, which encodes a component of the ESCRT-III complex.

![Diagram of SHH secretion mechanism](image)

**Figure 1.4: Proposed mechanisms for specialized SHH secretion**

Several proposed mechanisms for SHH secretion including oligomeric complexes, lipoprotein complexes, and extracellular vesicles (Briscoe and Therond, 2013).
1.3 Somatic mutation in brain development and degeneration

The second project of my dissertation analyzed somatic mutations in single neurons from patients with DNA damage repair disorders that cause microcephaly and early neurodegeneration. This analysis revealed a dramatic increase in somatic mutation rate at the single cell level in disease neurons compared to cells from normal brain. To provide the biological context for this project, I will highlight recent work showing that somatic mutations can disrupt brain development, introduce the role of DNA damage repair in brain development and degeneration, and highlight recent progress in somatic mutation analysis in single cells.

**Brain malformations can also be caused by somatic mutations**

Each year, additional genes and alleles are identified in new families with brain malformations that deepen our understanding of the genetic mechanisms of brain development. However, in the last few years, it has been increasingly recognized that somatic mutations, sometimes in the same genes that cause germline disorders, are also an important cause of brain malformations, especially milder or asymmetrical ones (Table 1.2). Unlike traditional inherited mutations, present in one or both parents before fertilization, somatic mutations arise after fertilization and so are only present in some cells of the body. Depending on when a somatic mutation occurs, it can be present in tissues throughout the body (early mutation) or it can be restricted to a specific organ or a specific region of tissue within an organ (late mutations) (Poduri et al., 2013). Although somatic mutations have been studied for many years in the context of cancer
development and growth, the recent discovery of somatic genetic causes of brain malformations shows that somatic mutations can also disrupt normal brain development. This discovery has opened a new field in brain malformation genetics and has encouraged us to pay much more attention to the role of somatic mutation in the developing brain.

### Table 1.2: Somatic mutations cause neurodevelopmental disorders

<table>
<thead>
<tr>
<th>Syndrome</th>
<th>Gene</th>
<th>Alternate allele frequency</th>
<th>Tissue</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double cortex</td>
<td>DCX, LIS1</td>
<td>5-26%</td>
<td>Peripheral blood</td>
<td>Jamuar et al., 2014</td>
</tr>
<tr>
<td>Pachygyria</td>
<td>TUBB2B</td>
<td>23%</td>
<td>Peripheral blood</td>
<td>Jamuar et al., 2014</td>
</tr>
<tr>
<td>Periventricular nodular heterotopia</td>
<td>FLNA</td>
<td>35%</td>
<td>Peripheral blood</td>
<td>Jamuar et al., 2014</td>
</tr>
<tr>
<td>Megalencephaly</td>
<td>PIK3CA</td>
<td>1-43%</td>
<td>Peripheral blood, lymphocytes, buccal, saliva</td>
<td>Riviere et al., 2012; Jensen et al., 2015</td>
</tr>
<tr>
<td>Hemi-megalencephaly</td>
<td>1q trisomy, AKT3, PIK3CA, MTOR</td>
<td>8-40%</td>
<td>Cortex</td>
<td>Poduri et al., 2012; Lee et al., 2012; Jansen et al., 2015</td>
</tr>
<tr>
<td>Focal cortical dysplasia</td>
<td>PIK3CA, MTOR</td>
<td>5-44%</td>
<td>Cortex, blood, buccal</td>
<td>D'Gama et al., 2015a; Jansen et al., 2015</td>
</tr>
<tr>
<td>Autism spectrum disorder</td>
<td>SCN1A, SETD2</td>
<td>16-32%</td>
<td>Cortex</td>
<td>D'Gama et al., 2015b</td>
</tr>
</tbody>
</table>

Somatic mutations that cause brain malformations can be organized into two broad categories: obligatory somatic mutations, mutations that can only be identified
when somatic because presence of the mutation in the germline would be lethal, and somatic mosaic mutations, mutations that occur in inherited brain malformations genes but that are not present in all cells of the body. Next generation sequencing and single cell technology catalyzed recent advances in identification of somatic mutations in brain malformations because high coverage sequencing allows mutations to be identified that are present in a small fraction of cells and would be missed by traditional Sanger sequencing and single cell sequencing allows direct measurement of the percentage of cells that carry a mutation (Poduri et al., 2013).

The first evidence that obligatory somatic mutations cause brain malformations was found in cortex overgrowth disorders. Mutations in mTOR pathway genes (AKT, PIK3CA, MTOR, 1q duplication) that are likely incompatible with life as germline mutations, because they would cause multiple organ defects, were discovered in resected brain tissue from patients with megalencephaly and hemimegalencephaly (malformations in which both hemispheres or one hemisphere of the cortex, respectively, grow abnormally large). These mutations were often found only in brain tissue but sometimes also at a low level in leukocytes (D'Gama et al., 2015a; Jansen et al., 2015; Lee et al., 2012; Poduri et al., 2012; Rivière et al., 2012). Follow up studies showed that somatic mutations in this pathway also cause focal cortical dysplasia, a malformation with one or a few small regions of abnormal neurons in the cortex (D'Gama et al., 2015a; Jansen et al., 2015). These studies suggest a model in which somatic mutations can be restricted to the brain, the cortex, or a specific region of the cortex based on when they occur during development, and that this timing determines the type of malformation, earlier mutations: megalencephaly, later: hemimegalencephaly, and latest: focal cortical dysplasias.
The second category of somatic mutations that cause brain malformations are somatic mosaic mutations that cause milder versions of conditions most commonly caused by germline mutations. In the early years after identification of *DCX* and *LIS1*, patients with somatic mosaic mutations in both of these genes were identified and found to have milder forms of disease, subcortical band heterotopia in somatic *LIS1* mutations (Sicca et al., 2003) and double cortex syndrome or asymptomatic disease in *DCX* somatic mutations (Gleeson et al., 2000). Recent work used high coverage NGS to show that somatic mutations in several different malformation genes are present in patients without a diagnosed germline mutation. Patients with double cortex syndrome, lissencephaly, pachygyria, and periventricular nodular heterotopia without a genetic diagnosis were found to have somatic mutations in *DCX*, *LIS1*, *TUBB2B*, and *FLNA* (Jamuar et al., 2014). Further highlighting the widespread nature of somatic mutations in neurodevelopmental disorders, a recent report found somatic mutations in *SCNIA* and *SETD2* in patients with autism (D’Gama et al., 2015b).

The recent accumulation of evidence for obligatory somatic and somatic mosaic mutations as causes of structural brain malformations and autism shows that somatic mutations are an important genetic cause of neurodevelopmental disorders. Somatic mutations arise as a result of unrepaired DNA damage, and disorders of deficient DNA damage repair cause microcephaly and premature neurodegeneration. I have found that neuronal DNA damage and somatic mutation rate are increased in Cockayne syndrome, a disorder of DNA damage repair (Chapter 3).
DNA damage in brain development and degeneration

Throughout development and during adult life, the nervous system is constantly surveying the genome for DNA damage and repairing the damage it finds. In one day, the genome of a single mammalian cell likely experiences 200 cytosine deaminations, 3,000 guanidine methylations, 10,000 spontaneous depurinations, 10,000-100,000 oxidative DNA lesions, 10,000 single-strand DNA breaks and 10 – 50 double-strand breaks (Madabhushi et al., 2014). And this does not count the many dipyrimidine dimers in cells exposed to UV damage from sunlight. As a result of this constant degradation of the genome, several DNA damage repair pathways have evolved. These include nucleotide excision repair (NER) for helix-distorting DNA damage from UV, polycyclic aromatic hydrocarbons, and reactive oxygen species (Nickoloff and Hoekstra, 1998, p. 335), base excision repair (BER) for ionizing radiation, oxidizing and methylating agents, and deamination and depurination (Nickoloff and Hoekstra, 1998, p. 336), double stranded break repair (DSBR) for double stranded DNA breaks, and single stranded break repair (SSBR) for single stranded breaks (Madabhushi et al., 2014). These DNA repair pathways maintain the fidelity of the genome for accurate DNA replication during cell division and for accurate mRNA transcription throughout the cell’s life.

A major theory of aging postulates that DNA repair efficiency decreases over time leading to an accumulation of unrepaired DNA damage and somatic mutations that lead to senescence, cell death, and neoplastic transformation (Maynard et al., 2015). Genetic defects in DNA damage repair pathways cause catastrophic disorders, many of which include neurodegeneration. For example, defective NER in Cockayne syndrome and xeroderma pigmentosum is associated with decreased growth and early degeneration
of the cortex, cerebellum, and basal ganglia; defective DSBR in ataxia telangiectasia and related disorders is associated with cerebellar atrophy and ataxia, defective DSBR in XLF syndrome causes microcephaly, and defective SSBR in spinocerebellar ataxia causes cerebellar atrophy (Madabhushi et al., 2014). In all these disorders, despite the clear relationship between defective DNA repair and degeneration, the mechanism and time-course of the postulated accumulated DNA damage, and how this damage contributes to the clinical phenotype, are still unclear. In addition, there is evidence for increased DNA damage in several neurodegenerative diseases including Alzheimer’s disease, Parkinson’s disease, and amyotrophic lateral sclerosis (ALS) (Coppedè and Migliore, 2015). Therefore, understanding the mechanism of DNA damage repair can not only lead to improved understanding of mechanisms of neuronal degeneration, it also promises to help in developing potential treatments for many disorders including, potentially, Alzheimer’s disease, dementia, and ALS.

Cockayne syndrome (CS) and xeroderma pigmentosum (XP), disorders of DNA damage repair characterized by early neurodegeneration, are caused by loss of function mutations in genes that encode components of the nucleotide excision repair (NER) pathway. NER is split into two components: transcription-coupled repair (TCR) for lesions in transcribed DNA and global genome repair (GGR) for lesions throughout the genome. These two types of NER have separate enzymes for lesion recognition and then recruit a shared complex for lesion repair. NER is activated by proteins that survey the genome for changes in conformation of the DNA helix caused by dipyrimidine dimers from UV exposure or DNA adducts from reactive oxygen species. In transcribed DNA, this function is performed by CSA and CSB (which detect stalled RNA polymerase) and
in non-transcribed DNA this function is performed by DDB1, DDB2 (encoded by \textit{XPE}), and XPC. When a DNA lesion is identified, the TFIIH transcription factor complex is recruited to the lesion. This complex includes a helicase, a 3’ nuclease, and a 5’ nuclease. XPD unwinds DNA at the lesion and then XPG and XPF excise the damaged nucleotide and several surrounding nucleotides. Then a DNA polymerase (delta, epsilon, or kappa) synthesizes new DNA to fill the patch and this DNA is ligated by XRCC1, ligase I, or ligase III (Cleaver et al., 2009).

CS and XP are caused by mutations in genes that encode components of the NER pathway. Each disease has different clinical manifestations that are specific to the location of the genetic mutation in the pathway. Mutations in \textit{XPA-XPQ} cause XP. XP proteins are responsible for GGR and TCR and patients with XP have three types of symptoms: skin, constitutional, and neurological. Skin findings include a 2,000-fold increase in skin cancer, with median age of onset at 8 years; skin sensitivity to sunlight; and skin pigmentation changes. Constitutional findings include slow growth and delayed sexual maturation. Neurological findings (present in 20% of patients) include neurodegeneration causing microcephaly and intellectual disability, decreased reflexes, and sensorineural deafness (Nickoloff and Hoekstra, 1998, p. 338). Mutations in \textit{CSA} and \textit{CSB} and some alleles of \textit{XPB}, \textit{XPD}, and \textit{XPQ} cause CS. CS is characterized by specific loss of TCR and patients have constitutional and neurologic symptoms but do not have increased incidence of skin cancer. Constitutional symptoms include profound postnatal growth failure, skin sensitivity to light, dental caries and median lifespan of 12 years. Neurologic symptoms include widespread neurodegeneration, demyelination, deafness, cataracts, and pigmentary retinopathy (Nickoloff and Hoekstra, 1998, p. 343).
The absence of skin cancer in CS suggests that defects in GGR and TCR in XP cause growth failure, neurodegeneration, and skin cancer, while defects in TCR only cause growth failure and neurodegeneration. Indeed, patients with defects in GGR only (from mutations in XPE, the lesion survey protein for non-transcribed DNA) do not have neurodegeneration or growth failure, suggesting these symptoms are primarily a result of TCR disruption (Cleaver et al., 2009). The mechanism underlying these symptom differences is unknown. One hypothesis is that skin cancer does not develop before CS patients die in their teenage years, and another is that keratinocytes from CS patients are more sensitive to UV damage and undergo cell death after exposure instead of transforming to become cancerous (Cleaver et al., 2009).

DNA damage and resulting somatic mutations are hypothesized to be a driver of aging in the brain and when DNA damage repair is genetically disrupted, as in CS and XP, neurodegeneration occurs at a very young age. Understanding how DNA damage causes brain aging and degeneration is very important and investigating the mechanisms of degeneration in CS and XP can provide broad insights into the effects of DNA damage on the brain.

**Somatic mutation analysis in single cells**

DNA damage that is left unrepaired or is repaired incorrectly causes somatic mutations. Recent work has discovered that human brain malformations can be caused by somatic mutations in addition to inherited germline mutations (Poduri et al., 2013). Studying the generation and consequences of somatic mutation is a powerful tool because it reveals genetic mechanisms that act at the level of individual cells. The recent
availability of new methods for DNA amplification and next generation DNA sequencing
technologies has dramatically lowered the cost of sequencing and made it feasible to
amplify and sequence the whole genome of many single cells, enabling comprehensive
genome-wide characterization of somatic variation at a single cell level (Hou et al., 2012;
Xu et al., 2012). Somatic mutation rate and identification of somatic variants in normal
individuals provides a developmental history of that tissue, including how closely related
individual cells are to one another, what mutational pressure the tissue has experienced,
and how the tissue ages over time. In disease, somatic mutation identification has been
used to identify driver mutations in cancer and genetic causes of brain malformations.
However, the availability of single cell methods enables new insight into the effects of
disease on the genome, even if those effects are diverse and target different cells at
different loci, by being able to compare genetic lesions in one cell to lesions in another.
This is particularly important in the nervous system, since neurons are postmitotic, and
genetic lesions in neurons tend to be unique to one cell, rather than shared by large
numbers of adjacent cells with common clonal origin as occurs in cancer.

Normal nervous system function requires a huge number of different neuronal and
glial subtypes enabling complex cellular interactions and highly specialized functions.
One hypothesis is that somatic mutation is used to generate neuronal diversity in the brain
in general and the cerebral cortex in particular (Muotri et al., 2005). This hypothesis has
been under investigation for the last 20 years. Experiments with vastly different
estimates of somatic mutation rate in neurons have been published that support or
question specific aspects of this hypothesis (Muotri et al., 2009; Evrony et al., 2012;
Upton et al., 2015; Hazen et al., 2016) and so determining whether this hypothesis is true
for specific times or tissues in the brain or for specific types of somatic variation remains of great scientific interest today. In particular, work from our lab and others show that mobile element insertions are relatively uncommon, copy number variants are more common, and SNVs are the most common source of somatic genetic variation (Evrony et al., 2012; Cai et al., 2013; Lodato et al., 2015). Our finding that SNVs are common in single neurons suggests that this type of somatic variation could contribute to neuronal diversity or normal neuron function.

Previously in our lab, Mike Lodato and Mollie Woodworth characterized somatic SNVs in post-mortem human neurons. They performed fluorescence-assisted nuclear sorting (FANS) to isolate single neuronal nuclei from fresh frozen human cortex, used multiple displacement amplification (MDA) to amplify the genome of each isolated nucleus, and finally performed whole genome sequencing (WGS) (Evrony et al., 2012). They identified somatic SNVs using a conservative variant calling pipeline that was sufficiently specific that 90% of called variants shared between two or more neurons could be validated from primary brain DNA using other methods. They performed WGS on 10-16 prefrontal cortex neurons each from three normal individuals (aged 17, 19, and 42) and found an average of 1458-1580 somatic SNVs per neuron (Lodato et al., 2015). Nearly all of these variants are private variants present only in a single neuron, highlighting the enormous mutational burden neurons experience after they become post-mitotic. SNVs in coding genes were significantly enriched in relatively highly expressed genes and variants occurred more frequently at CpG islands. Together these three pieces of data suggest that a significant fraction of somatic SNVs in neurons result from transcriptional DNA damage. In addition, somatic SNVs were significantly enriched in
exons of coding genes and in genes expressed in neurons, highlighting the functional role of these SNVs and suggesting that these variants could contribute to neuronal diversity in the human cortex.

With the recent decrease in cost of WGS, many similar experiments are underway and some have already been reported. Three papers in addition to the work from our lab estimate somatic mutation rate based on single cell WGS (Table 1.3).

**Table 1.3: Somatic mutation rates from single cell WGS.**

<table>
<thead>
<tr>
<th>Somatic SNVs per cell</th>
<th>Cell type</th>
<th>Genome amplification</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>~1500</td>
<td>Human, cortical neuron</td>
<td>MDA</td>
<td>Lodato et al., 2015</td>
</tr>
<tr>
<td>609</td>
<td>Mouse, intestinal stem cells</td>
<td>Organoid culture</td>
<td>Behjati et al., 2014</td>
</tr>
<tr>
<td>391</td>
<td>Human, endothelial progenitor cell</td>
<td>iPSC reprogramming and cell culture</td>
<td>Rouhani et al., 2016</td>
</tr>
<tr>
<td>86</td>
<td>Mouse, mitral and tufted neuron</td>
<td>SCNT and cell culture</td>
<td>Hazen et al., 2016</td>
</tr>
</tbody>
</table>

Kristin Baldwin’s lab measured somatic mutation rate in mouse mitral and tufted neurons. Her group dissected mouse mitral and tufted cells, large projection neurons of the olfactory bulb, and then used somatic cell nuclear transfer to clone mice from these cells. They then used WGS to sequence tissue from the donor animal and from ESCs of the cloned embryo. Variants found in ESCs that were absent from the donor tissue were interpreted to be somatic variants from the original single neuron. They created ESC lines from 6 mitral and tufted cells and found an average of 86 somatic SNVs per neuron. All of these SNVs were private to each cell; there were no shared variants (Hazen et al.,
The large difference in neuronal somatic SNV rate between our lab’s results and the Baldwin lab’s results shows that more work is necessary to improve the accuracy of neuronal somatic mutation rate measurements.

There are several differences between these two experiments that may contribute to the large discrepancy in estimated somatic mutation rate. Specifically, one experiment was in mouse and the other in human; one experiment measured olfactory bulb mitral and tufted neurons and the other measured cortical neurons; and one experiment used ES culture to amplify the genome from a single neuron and the other used MDA whole genome amplification. In the Walsh lab, we are confident in our ability to minimize MDA amplification errors when calling somatic SNVs, though these are acknowledged as a potential source of false positive SNV calls, and we are continually improving our methods for SNV calling and SNV validation to further increase the accuracy of our somatic mutation rate measurements.

Allen Bradley’s group and Michael Stratton’s group have reported single cell somatic mutation rates for non-neuronal cells. The Bradley lab measured the number of somatic SNVs in a single human endothelial progenitor cell (EPC). They established a cell line starting from a single EPC from a 57-year-old healthy man and then reprogrammed this line into three independent iPSC lines. They then performed WGS of the three iPSC lines and the original EPC line and cultured fibroblasts from the individual. Because the fibroblasts were generated from many cells in a skin biopsy, they were used as the germline reference genome. Somatic SNVs were identified in the EPC line and three iPSC lines. The intersection of these four groups of somatic SNVs represents the group of somatic SNVs from the single starting EPC, rather than SNVs
that arose during cell culture and reprogramming. They found that this single EPC had 391 somatic SNVs (Rouhani et al., 2016). Using 2-year-old mice, the Stratton lab collected clonal adult stem cells from the stomach, small intestine, colon, and prostate, and grew organoids from single cells or a single group of clonal cells (Behjati et al., 2014). They then performed WGS on 25 organoids and identified somatic SNVs from the original stem cells by comparing variants called in an organoid to germline variants called in DNA from the donor’s tail. They found that the number of somatic SNVs from each original stem cell ranged from 179 – 1,190 with an average of 609 SNVs / cell.

WGS of a single cell to estimate somatic mutation rate is now feasible, but more research is needed to determine the accuracy of this method. Our lab found 1458 – 1580 somatic SNVs per cortical neuron in human, Kristin Baldwin’s lab found 86 somatic SNVs per mitral tufted neuron in mouse, Allen Bradley’s lab found 391 somatic SNVs per EPC in human, and Michael Stratton’s lab found 609 somatic SNVs per intestinal stem cell in mouse. These differences may simply reflect the different cell types and species used in each experiment. However, each method also used a different approach for genome amplification of a single cell: we used MDA, Baldwin used SNCT and ESC culture, Bradley used iPS reprogramming and culture, while Stratton used organoid culture. It is important as the field moves forward to confirm that each of these methods can accurately measure somatic mutation rate. One of the motivations for my work sequencing single neurons from CS and XP is to use neurons from diseases with increased somatic mutation rate to show that MDA amplification can accurately measure the number of single cell somatic SNVs.
1.4 Conclusion

Human genetics of brain malformation began 20 years ago with cloning of two X-linked genes \textit{DCX} and \textit{FLNA} using yeast artificial chromosome mapping. DNA sequencing technology has improved so much in the past two decades that is now possible to sequence the entire genome of several single neurons to identify somatic mutations that alter brain development. Studying the function of brain malformation genes has revealed many surprising mechanisms of neurodevelopment. In Chapter 2, I will present evidence for a surprising mechanism in brain development, exosome-mediated secretion of SHH, discovered by studying \textit{CHMP1A}, a gene mutated in microcephaly with cerebellar hypoplasia. In Chapter 3, I will show that the somatic mutation rate is greatly increased in Cockayne syndrome cortical neurons, highlighting the importance of DNA damage repair in both brain development and degeneration.
Chapter 2: CHMP1A function reveals exosome-mediated secretion of sonic hedgehog in developing brain
Publication

This chapter is currently being submitted to Cell as an article entitled “CHMP1A is essential for exosome-mediated secretion of sonic hedgehog in developing brain.”

Attribution

Michael Coulter wrote the dissertation and planned, conducted, and analyzed all experiments except as noted below. Raphael Gaudin performed EV purification and assisted with Propidium iodine staining and EGFR degradation experiments. Frank Jacobs and Gerrald Lodewijk differentiated and grew cortical organoids, performed RNA sequencing, and helped with RNA sequencing analysis. Elaine Lim grew cerebral organoids used for immunofluorescence. Wei Lee provided mouse cortex electron microscopy data for analysis. Hart Lidov and Monica Calicchio provided human choroid plexus tissue. Jennifer Partlow and Edward Yang collected and read brain MRIs. Thorsten Schlaeger established patient fibroblast line and derived CHMP1A null (Q30X) iPSC line. Margaret Thompson and Vijay Ganesh generated Chmp1a GT mouse line. Stan Hollenberg provided the rabbit Chmp1a antibody, Adrian Salic provided the SHH plasmid, and Chad Cowan and Kirin Musunuru provided CRISPR plasmids. Maria Ericsson, Louise Trakimas, and Elizabeth Benecchi assisted with TEM sample preparation and imaging. Bill Fowle assisted with SEM sample preparation and imaging. Hannah Somhegyi drew the illustration in Figure 2.19D. Christopher Walsh, Tomas Kirchhausen, and David Haussler, Maria Lehtinen, and Ganesh Mochida oversaw the project.
2.1 Summary

Exosomes, small membrane-bound vesicles released when multivesicular bodies (MVB) fuse with the cell membrane, are essential for cell-to-cell communication because they allow hydrophobic signaling molecules to travel through the extracellular environment. In the nervous system, exosome secretion has been proposed to regulate synapse growth and retrograde signaling at the neuromuscular junction, but other roles have not been identified. Our data show that CHMP1A, a member of ESCRT-III complex recently implicated in microcephaly and cerebellar hypoplasia, is required for intraluminal vesicle (ILV) budding in the MVB, and absence of CHMP1A reduces ILVs inside MVBs and impairs exosome release. Sonic hedgehog (SHH) secretion is reduced in the absence of CHMP1A, and Chmp1a null mice exhibit features of disrupted Shh signaling including growth failure, microcephaly, and cerebellar and basal ganglia hypoplasia. We show that SHH secretion is exosome-mediated and requires CHMP1A. Our data reveal a novel mechanism of localized, as well as diffuse, growth factor secretion and cell-to-cell communication in developing mammalian brain.
2.2 Introduction

Exosomes are increasingly recognized as essential mediators of specialized cellular secretion, but the mechanisms of exosome function are not well understood. Exosomes represent a specific class of extracellular vesicles (EVs), 30-100 nm in diameter, defined by their origin as MVB intraluminal vesicles (ILVs), which are subsequently released upon MVB fusion with the cell membrane (Cocucci and Meldolesi, 2015; Pan et al., 1985). Exosomes are essential for cell-to-cell communication because they allow hydrophobic signaling molecules (Korkut et al., 2009), RNA (Tietje et al., 2014), and other specialized cargo (Budnik et al., 2016) to travel through an aqueous extracellular environment. At *Drosophila* neuromuscular junction (NMJ), exosome-mediated wingless (Wg/Wnt) secretion is required for synapse growth, and exosome-mediated Synaptotagmin 4 (Syt4) secretion is required for retrograde signaling (Koles et al., 2012; Korkut et al., 2009; Korkut et al., 2013). Exosome-mediated secretion is essential in these functions because Wg is hydrophobic and requires a transmembrane carrier, Evi, for secretion, and Syt4 is an intramembrane protein. Cultured mammalian neurons (Lachenal et al., 2011), oligodendrocytes (Fruhbeis et al., 2013), and microglia (Antonucci et al., 2012) secrete exosomes that have been proposed to similarly modulate synaptic strength and maintain cell viability, but this has not been proven because of a lack of *in vivo* vertebrate models for selective removal of exosome function. Exosomes are also proposed to mediate pathological transfer of RNAs, prion-like proteins, and Tau proteins (Asai et al., 2015), and so might play key roles in neurodegeneration as well.
Loss of function (LOF) mutations in \textit{CHMP1A} cause microcephaly with pontocerebellar hypoplasia and short stature (Mochida et al., 2012), but previous work suggested several possible mechanisms of CHMP1A in the brain. First, \textit{CHMP1A} encodes a protein that is a member of the endosomal sorting complexes required for transport (ESCRT) III complex (Howard et al., 2001), and ESCRT complexes have several cellular functions that might cause microcephaly, including exosome secretion (Colombo et al., 2013), regulation of receptor degradation at the MVB (Spitzer et al., 2009), and cellular abscission during cytokinesis (Carlton et al., 2012). Second, some CHMP1A protein localizes to the nuclear matrix (Stauffer et al., 2001) and CHMP1A decreases expression of the pro-senescence gene \textit{INK4A} through regulation of the nuclear repressor BMI1 (Mochida et al., 2012), suggesting a potential nuclear function of CHMP1A. Third, the severe cerebellar hypoplasia seen with \textit{CHMP1A} mutations resembles cerebellar hypoplasias seen in disorders of cilia function (Goetz and Anderson, 2010) and disrupted \textit{Shh} signaling (Corrales et al., 2006).

\textit{SHH} is an essential growth factor that functions during development as a morphogen (Cohen et al., 2015; Roelink et al., 1995), a mitogen (Nielsen and Dymecki, 2010; Dahmane and Ruiz i Altaba, 1999), an axon guidance molecule (Wilson and Stoeckli, 2013; Charron et al., 2003), and a regulator of synapse formation (Harwell et al., 2012). In developing cerebellum, \textit{SHH} stimulates proliferation of granule cell precursors (GCPs) that generate granule neurons, the most abundant type of neuron in the brain (Zhou et al., 2007). Disruption of \textit{Shh} causes profound cerebellar hypoplasia and partial loss of \textit{Shh} causes milder hypoplasia (Corrales et al., 2006).
Whereas the source of secreted SHH that regulates GCP proliferation appears to be Purkinje neurons (Wechsler-Reya and Scott, 1999), large principal cells of the cerebellum, the exact mechanism of SHH secretion by Purkinje cells is unknown. SHH is a hydrophobic, cholesterol-modified protein that requires specialized mechanisms for secretion into the aqueous extracellular space (Briscoe and Therond, 2013). Studies in invertebrates have proposed mechanisms for hedgehog secretion including oligomeric complexes, lipoprotein particles, exosomes, and recently exovesicles (Therond, 2012; Matusek et al., 2014). In vertebrate development, SHH secretion at the mouse ventral node appears to occur via large membrane-bound particles (300-500 nm in diameter) that contain SHH and retinoic acid (Tanaka et al., 2005). Although several mechanisms for SHH secretion have been proposed, few have been validated beyond the specific context reported and no mechanism has been identified for SHH secretion during brain development. Here we show that SHH is secreted on exosomes in the developing mammalian brain and that exosome biogenesis requires CHMP1A, providing the first genetic model to dissect exosome function in mammalian brain. We show that Chmp1a null mice, similar to CHMP1A mutant humans, have widespread defects in forebrain and hindbrain development, implicating exosome function in many aspects of central nervous system development and potentially adult function as well.
2.3 Results

A Chmp1a gene trap mouse creates a null mutation and highlights CHMP1A expression

To model Chmp1a LOF, we created a Chmp1a gene trap (GT) mouse line as shown in Figure 2.1 that completely removes CHMP1A protein. A gene trap cassette (Stryke et al., 2003) inserted in intron 1 of Chmp1a contains a strong splice acceptor from En2 fused to the coding sequence for beta-galactosidase (Figure 2.1A). ES cells containing this GT allele were injected into mouse blastocysts to generate Chmp1a GT chimeric mice, which were outcrossed to create germline Chmp1a GT allele carriers. DNA sequencing confirms that in homozygous GT embryos, Chmp1a intron 1 is fused to En intron 1 (Figure 2.2A). Heterozygous GT mice show reduced CHMP1A protein expression compared to wild type (WT) and homozygous GT mice express no detectable CHMP1A protein (Figure 2.1B), confirming that Chmp1a GT mice are null for CHMP1A.

CHMP1A is required for embryonic development

CHMP1A is required for normal embryonic development and postnatal survival in mice, with Chmp1a null embryos being significantly smaller than littermate controls during embryogenesis and more so at birth, when Chmp1a null mice are 40% smaller than controls (two-tailed t-test, p = 2x10^{-5}; Figures 2.1C, 2.1E). Chmp1a null mice die at or soon after birth (Figure 2.2B), reflecting increasingly severe defects in brain development during embryogenesis. By P0, Chmp1a null brains are 14% smaller than
controls (two-tailed t-test, p = 0.01) with a generalized reduction in brain size (Figures 2.1D, 2.1F). Since there are no detectable differences in survival, body weight, or brain weight between wild type and Chmp1a heterozygous mice (Figures 2.2C, 2.2D), wild type and heterozygous mice have been combined in all analyses and referred to as controls. Chmp1a null mice have smaller olfactory bulbs, smaller and thinner cerebral cortex, smaller striatum, and the developing cerebellum is smaller and less well foliated than normal (Figure 2.1G). Together, reduced body size, microcephaly, reduced basal ganglia, and cerebellar hypoplasia in Chmp1a null mice closely models the phenotype of CHMP1A null patients (Mochida et al., 2012).
Figure 2.1: CHMP1A is essential for brain development in mice

(A) Schematic of Chmp1a gene trap allele design and insertion. (B) Immunoblot of MEFs from Chmp1a GT mice show reduced CHMP1A protein expression in het GT mice and undetectable CHMP1A protein in homozygous GT mice. (C) P0 Chmp1a null pups are smaller than littermate controls. (D) P0 Chmp1a null pups have smaller brains than controls. Note smaller olfactory bulbs and shorter A-P length of the cortex (dashed white lines). (E) Chmp1a null embryos grow slowly. P0 body mass is reduced 35%. (F) Chmp1a null brain grows slowly, with P0 brain mass reduced 14%. (G) P0 Chmp1a null pups show thin cerebral cortex, small basal ganglia and olfactory bulbs, and small and less foliated cerebellum. Two-tailed t-test, * p < 0.05, *** p < 0.001.
Figure 2.1 (Continued)
Figure 2.2: CHMP1A is essential for mouse postnatal survival and Chmp1a heterozygous mice are normal

(A) Sanger sequencing of genomic DNA in Chmp1a null mice shows fusion of Chmp1a intron 1 to En2 intron 1 in GT cassette. (B) Mendelian ratios of Chmp1a GT litters during embryonic development. Expected Mendelian ratio is shown in the left bar. At late embryogenesis Chmp1a gt/gt animals begin to die and few are recovered at P0. (C) Chmp1a +/gt embryos do not have a defect in embryonic development compared to Chmp1a +/+ embryos. Below: mass and p-values of +/+ vs +/gt, two-tailed t-test. (D) Chmp1a +/gt embryos do not have a defect in embryonic brain development compared to Chmp1a +/+ embryos. Below: mass and p-values of +/+ vs +/gt, two-tailed t-test.

CHMP1A is essential for human cerebral organoid formation

CHMP1A is also required for human cerebral organoid formation and progenitor proliferation (Figure 2.3). We differentiated cerebral organoids (Figure 2.3A) from WT iPSCs, and two CHMP1A null iPSC lines, patient mutation Q30X (Mochida et al., 2012) and CRISPR-induced mutation V33fs (Figure 2.4A). We made cerebral organoids using directed differentiation (Experimental Procedures and Lancaster et al., 2013). Immunostaining 4 week organoids for PAX6 (to mark cortical progenitors) and TUJ1 (to
mark post-mitotic neurons) showed less well-developed cortical-like regions and fewer progenitor cells in \textit{CHMP1A} null organoids compared to controls, providing additional evidence of defective neural progenitors (\textbf{Figure 2.3A}). The average length of the proliferative zone (defined as regions of continuous PAX6-positive cells lining the organoid’s internal surface) was reduced by 35\% in \textit{CHMP1A} null organoids (apical surface length, 459 vs 295 um, WT: n = 10, V33fs: n = 10, two-tailed t-test, p = 0.0015, \textbf{Figure 2.4B}).

RNA sequencing revealed pronounced changes in gene expression that became apparent as organoids matured (\textbf{Figure 2.3B, 2.4C}). GO pathway analysis of mature organoids showed categories “cell cycle” and “regulation of neurogenesis” were significantly decreased in \textit{CHMP1A} null organoids while “synaptic transmission” was significantly increased (\textbf{Figure 2.3C, 2.4E}), confirming defects in neurogenesis and showing premature neuronal maturation. In particular, canonical markers of proliferating progenitor cells were decreased in \textit{CHMP1A} null organoids, including \textit{PAX6}, \textit{TBR2}, \textit{FOXG1}, \textit{SOX1},\textit{2,3,5}, \textit{HMGA2}, and \textit{NR2E1}, while markers of synapse formation, \textit{SYN1}, \textit{SLC12A5}, \textit{SLC6A5}, and \textit{NPTX2}, were increased (\textbf{Figure 2.3D, 2.4D}). Targets of \textit{WNT} signaling, \textit{LEF1}, \textit{FZD7}, \textit{NEUROG2}, \textit{SFRP2}, \textit{MYCN}, and \textit{CCND1}, were downregulated in the absence of \textit{CHMP1A} suggesting specific pathway defects responsible for impaired progenitor proliferation (\textbf{Figure 2.3E}). Together, RNA sequencing and immunostaining of \textit{CHMP1A} null cerebral organoids suggests CHMP1A is essential for neural progenitor proliferation and function by regulating proliferative signaling cascades such as \textit{WNT}. 

45
Figure 2.3: CHMP1A is essential for progenitor proliferation in human cerebral organoids

(A) Appearance of human cerebral organoids with characteristic cortical morphology (boxed). Postmitotic neurons (immunostained for Tuj1, red) surround cortical progenitors (immunostained for Pax6, green) that surround central cavities. Some CHMP1A null cortical regions have neurons among progenitors (arrowhead). (B) RNA sequencing of Control and CHMP1A null cerebral organoids shows altered gene expression in mature organoids (week 5) lacking CHMP1A. (C) GO pathway enrichment analysis of mature organoids. (D) By week 5, several proliferative marker genes have decreased expression in CHMP1A null organoids while synapse makers have increased expression. (E) WNT signaling target genes are downregulated in week 5 CHMP1A null organoids.
Figure 2.3 (Continued)
Figure 2.4: Disrupted gene expression in *CHMP1A* null cerebral organoids

(A) Human cerebral organoids were differentiated from two *CHMP1A* null iPS lines, V33fs and Q30X. Sanger sequencing shows the mutation in each line. Immunoblot shows that V33fs and Q30X iPSCs produce no detectable CHMP1A protein. (B) Cortical proliferative zone is 35% shorter in *CHMP1A* null organoids compared to controls. (C) All genes with greater than 2 fold changed expression in control vs *CHMP1A* null organoids. Week 2 are differentiating organoids, week 5 are mature. (D) Expression of neural stem cell genes and postmitotic neuron genes. In mature organoids, most stem cell markers are downregulated in *CHMP1A* null organoids and some neuron markers are upregulated. In differentiating organoids, ventral stem cell markers are upregulated and caudomedial markers are downregulated. (E) GO pathway enrichment analysis in week 2 differentiating organoids. “Synaptic transmission” is upregulated in the absence of CHMP1A as it is in week 5. Two-tailed t-test, **p < 0.01. 
Figure 2.4 (Continued)
CHMP1A is expressed in post-mitotic neurons and choroid plexus during brain development

The Chmp1a GT allele inserts lacZ into the Chmp1a locus and provides crucial information about the normal expression pattern of the Chmp1a gene. Beta-gal staining in heterozygous GT mice reveals expression of Chmp1a in the developing cerebellum and hindbrain choroid plexus (Figures 2.5A, 2.6A). In the cerebellum at P4, the peak of GCP proliferation, Chmp1a driven lacZ is specifically expressed in the Purkinje cell layer (Figure 2.5A). Building on this finding, we then used an anti-CHMP1A antibody to localize CHMP1A protein to specific cell types in the developing cerebellum, choroid plexus, cerebral cortex, and olfactory bulb. Two results show the specificity of this antibody. First, immunoblot of Chmp1a null MEF cell lysate with this antibody shows no reactivity (Figure 2.1B). Second, tissue sections from Chmp1a null mice incubated with this antibody have nearly complete loss of signal compared to wild type littermates (Figure 2.6B). Immunostaining with this antibody confirms that CHMP1A is highly expressed in Purkinje cells in the developing cerebellum with little expression in GCPs (Figure 2.5B). In the choroid plexus, CHMP1A is expressed in epithelial cells, co-stained with AQP1 (Figure 2.5C). In developing cerebral cortex, CHMP1A protein is enriched in post-mitotic neurons of the cortical plate compared to ventricular zone progenitors (Figure 2.5D). Finally, CHMP1A is expressed in the mitral layer of the olfactory bulb at P0 (Figure 2.5E). Immunostaining of human choroid plexus reveals CHMP1A subcellular localization. CHMP1A protein is localized in the cytoplasm and colocalizes with CD63 positive MVBs near the ventricular surface (Figure 2.5F), suggesting a role in exosome secretion, which occurs via MVB fusion at the surface.
Figure 2.5: CHMP1A is expressed in postmitotic neurons and ChP epithelial cells in developing brain

(A) In P4 cerebellum, *Chmp1a*-driven LacZ is detected in Purkinje cells of the cerebellum and absent from EGL. LacZ is also strongly expressed in hindbrain choroid plexus (ChP). (B) CHMP1A immunoreactivity in cerebellar Purkinje cells. Merge of Calbindin and CHMP1A highlights CHMP1A in Purkinje cell cytoplasm and dendrites (white arrow) with absence in the nucleus (asterisk). (C) CHMP1A immunoreactivity is punctate in epithelial cells of hindbrain choroid plexus at P0. AQP1 labels ventricular surface. (D) CHMP1A protein is expressed in post-mitotic neurons (CTIP2) but not cortical progenitors (SOX2) at E14.5. (E) CHMP1A protein in mitral cell layer but not in progenitor cells in P0 olfactory bulb. (F) CHMP1A immunoreactivity in human hindbrain ChP (11 d.o.) is cytoplasmic but not nuclear. MVB marker CD63 localizes to ventricular surface. There is partial colocalization of CHMP1A and CD63 puncta at the ventricular surface (inset and white arrows).
Figure 2.5 (Continued)

A. Cerebellum (P4)

B. Cerebellum (P4)

C. Mouse hindbrain ChP (P0)

D. Cerebral cortex (E14)

E. Olfactory bulb (P0)

F. Human ChP (11 d.o.)
Figure 2.6: Specificity of beta-gal and CHMP1A antibody staining

(A) Beta-gal staining in wild type mouse shows no signal in cerebellum or choroid plexus at P4. (B) Immunostain for CHMP1A shows no signal in Chmp1a null cerebellum at P0.
**CHMP1A is essential for cortical progenitor proliferation**

*Chmp1a* null embryos show defects in forebrain development and progenitor proliferation (Figure 2.7). *Chmp1a* null embryos have decreased anterior to posterior cerebral cortex length, which suggests a defect in proliferation of cortical progenitors (Figure 2.1D). We found evidence for decreased cortical neurogenesis by observing that *Chmp1a* null embryos (E17.5-18.5) have a thinner cortex than controls (13% thinner, two-tailed t-test, p = 0.03) and defects in cortical layers (Figure 2.7A). Superficial cortical layer (II-IV, CUX1 positive neurons) thickness is reduced by 25% in the absence of *Chmp1a* (two-tailed t-test, p = 0.002) (Figure 2.7A). In contrast, deep cortical layer (V-VI, CTIP2 positive neurons) thickness is less affected as thickness is only reduced by 9% (two-tailed t-test, p = 0.08) (Figure 2.7A). Preferential reduction of upper cortical layer thickness has been observed in other mouse models of microcephaly and can accompany a defect in cortical neurogenesis (Lizarraga et al., 2010) since cortical layers form in an inside-out manner with deep layer neurons born first and upper layer neurons born last (Greig et al., 2013). Quantification of mitotic ventricular zone progenitors using phosphorylated histone H3 (pH3) at E14.5 was used to measure cortical progenitor proliferation. As expected, we found fewer mitotic ventricular zone progenitors in *Chmp1a* null embryos compared to controls (20% reduction, two-tailed t-test, p = 0.003) (Figure 2.7B) suggesting CHMP1A is essential for neural progenitor proliferation. On the other hand, other mechanisms can underlie microcephaly, specifically increased apoptosis and apical surface malformation (Lizarraga et al., 2010; Hu et al., 2014; Kim et al., 2010). We found no detectable increase in the number of apoptotic cells (immunostain for cleaved caspase 3) during neurogenesis in the cortex in *Chmp1a* null
embryos (Figure 2.8A) and found that the apical surface is intact (immunostain for aPKC) and has normal localization of key signaling proteins such as beta-catenin (Figure 2.8B). Together these results argue that the microcephaly and cortical thinning observed in the absence of CHMP1A is a result of decreased cortical progenitor proliferation.

*Chmp1a null telencephalon phenocopies Shh hypomorphic mice*

The combination of cerebellar hypoplasia and growth failure in *Chmp1a* null human patients and mice and CHMP1A expression in Purkinje cells and choroid plexus, two important sources of the growth factor Sonic hedgehog (SHH) during brain development (Corrales et al., 2004; Dahmane and Ruiz i Altaba, 1999; Lewis et al., 2004; Nielsen and Dymecki, 2010), suggests that CHMP1A may regulate SHH secretion. The ventral telencephalon, which requires Shh for normal development (Palma and Ruiz i Altaba, 2004), is reduced in the absence of CHMP1A. Mice with hypomorphic transgenic alleles of *Shh* can form a dorsal telencephalon but show defects in basal ganglia development (Huang et al., 2007). We found two defects in basal ganglia formation in *Chmp1a* null embryos, suggesting that loss of CHMP1A causes partial loss of *Shh* signaling (Figure 2.7). First, by late embryogenesis, the striatum is substantially reduced in size in the absence of CHMP1A (25% reduction, two-tailed t-test, p = 0.0004) (Figure 2.7C). Second, this phenotype follows from an early defect in progenitor proliferation in the MGE and LGE, two structures that provide neurons for the basal ganglia. Immunostaining for pH3 during MGE and LGE neurogenesis at E12.5 and showed a dramatic 43% reduction in pH3 positive progenitors in *Chmp1a* null embryos (two-tailed t-test, p = 0.016) (Figure 2.7D). These two observations are consistent with
loss of CHMP1A causing defective Shh signaling. Matching this phenotype from
Chmp1a null mice, we reviewed brain MRIs of patients with null mutations in CHMP1A
and found that some patients have basal ganglia hypoplasia (Figure 2.9).
Figure 2.7: CHMP1A is essential for cortex and dorsal telencephalon development

(A) The cortical plate is 13% thinner in *Chmp1a* null embryos at E18.5. (B) 19% reduction in pH3 positive mitotic cortical progenitors at E14.5 *Chmp1a* null embryos. (C) Striatum area is reduced by 25% at E18.5 in *Chmp1a* null embryos compared to controls. (D) 43% fewer pH3 positive mitotic progenitors in the MGE and LGE in *Chmp1a* null embryos. Two-tailed, unpaired t-test, * p < 0.05, ** p < 0.01.
Figure 2.8: Normal cell death and ventricular surface in Chmp1a null cortex

(A) Apoptotic cells in developing cortex (E12.5) were labeled with immunostaining for cleaved caspase 3. There is no detectable difference between Chmp1a null and littermate controls (6.25 vs 7.5 per hemisphere, Control: n=6, Chmp1a null: n=3, two-tailed t-test, p = 0.52). (B) Immunostaining for beta-catenin and atypical protein kinase C (E14.5) shows no detectable defect in cortical ventricular surface formation.
Figure 2.9: MRI imaging of patients with CHMP1A null mutations

(A) MRI images of one control and two patients with CHMP1A null mutations. Control is a neurologically normal 2 year old male, CH-2701 is 6 month old female with Q30X CHMP1A mutation, and CH-3105 is 3 month old male with c.28-13G>A V33fs CHMP1A mutation (Mochida et al., 2012). Sagittal images show decreased head circumference (microcephaly) and severe cerebellar hypoplasia (arrow) in absence of CHMP1A. Coronal images show hypoplasia of caudate (arrows) in absence of CHMP1A.
CHMP1A is essential for Shh-mediated GCP proliferation

Cerebellar hypoplasia in Chmp1a null mice and co-expression of Chmp1a and Shh in Purkinje cells suggests that CHMP1A regulates SHH-mediated progenitor proliferation (Corrales et al., 2004, Dahmane and Ruiz i Altaba, 1999, Lewis et al., 2004). GCP proliferation is significantly impaired in Chmp1a null mice (Figure 2.10). At P0, the latest date we can study Chmp1a null mice, we observed a substantial decrease in mitotic GCPs (Figure 2.10A). Chmp1a null mice have 40% fewer mitotic GCPs than littermate controls (two-tailed t-test, p = 0.0001) (Figure 2.10C). This deficit is twice as large as the reduction in mitotic progenitors in the developing cortex and matches human CHMP1A null patients, whose cerebellar hypoplasia is strikingly severe in relation to their more modest microcephaly (Mochida et al., 2012). The overall size of the cerebellum is already reduced at P0 in Chmp1a null mice (Figures 2.10A, B). To isolate an effect of the absence of CHMP1A on GCP proliferation, we normalized the number of mitotic GCPs to cerebellum diameter. Even with this conservative correction, Chmp1a null mice have a 20% reduction in mitotic GCP density compared to controls (two-tailed t-test, p = 0.0009) (Figure 2.10C). Since SHH is the primary mitogen that drives GCP proliferation (Corrales et al., 2004; Dahmane and Ruiz i Altaba, 1999; Lewis et al., 2004) this reduction suggests that CHMP1A is necessary at some level for SHH-mediated proliferation. At E18.5, Shh signaling is more active in the anterior cerebellum and less active in the dorsal cerebellum (Corrales et al., 2004). We measured GCP proliferation in areas with high and low Shh signaling. In controls, we found a nearly two-fold increase in mitotic GCPs (pH3+) at P0 in the anterior EGL compared to dorsal EGL. In Chmp1a null mice, this difference was substantially reduced (1.75 vs 1.25, two-tailed t-test, p =
0.005) (Figures 2.10A, D) consistent with loss of Shh signaling. *In situ* hybridization for *Ptch*, a downstream target of Shh signaling, revealed lower *Ptch* expression in GCPs and Purkinje cells in *Chmp1a* null mice at P0 (Figure 2.10E), providing further evidence for decreased Shh signaling in the absence of CHMP1A.

We hypothesize that decreased Shh signaling is a result of a defect in SHH secretion. This is supported by a lack of detectable difference in Purkinje cell number in *Chmp1a* null mice (area: 0.36 vs 0.33, Control: n=4, *Chmp1a* null: n=2, two-tailed t-test, p = 0.16, Figure 2.11A), and by the finding that smoothened agonist stimulation of *Chmp1a* null and control MEFs showed no detectable difference in *Gli1* activation, a primary downstream target of Shh signaling, (76.9x vs 51.5x, Control: n=4, *Chmp1a* null: n=4, two-tailed t-test, p = 0.39, Figure 2.11B). Finally, we directly measured SHH protein concentration in embryonic CSF with ELISA, a highly sensitive method. Choroid plexus epithelial cells secrete Shh into the CSF during mid- to late-embryogenesis (Nielsen and Dymecki, 2010). CSF from the fourth ventricle of E14.5 embryos showed a 40% decrease in secreted SHH concentration in *Chmp1a* null embryos compared to controls (47 vs 29 pg/ml, Control: n=10, *Chmp1a* null: n=5, two-tailed t-test, p = 0.009) (Figures 2.12A, B), showing that CHMP1A is required for SHH secretion in the developing brain.
Figure 2.10: CHMP1A is essential for mouse cerebellum development and is required for SHH-mediated GCP proliferation

(A) P0 midline section shows cerebellar hypoplasia in Chmp1a null pups compared to control littermates. pH3+ mitotic GCPs were counted in the entire EGL (white lines) and specifically in the dorsal and anterior lobes. (B) Midline perimeter is decreased 18% in Chmp1a null pups. (C) Chmp1a null pups have 38% fewer mitotic GCPs than control littermates. (D) SHH drives GCP proliferation more in anterior lobe than in dorsal lobe. Control littermates have 75% more mitotic GCPs in the anterior lobe compared to the dorsal lobe; in contrast, Chmp1a null pups have only 25% more mitotic GCPs in anterior lobe. (E) ISH for Shh target gene Ptch shows reduced expression in Chmp1a null P0 cerebellum. Two-tailed, unpaired t-test, * p < 0.05, ** p < 0.01.
Figure 2.11: *Chmp1a* null cerebellum has normal number of Purkinje cells and downstream *Shh* activation is normal in the absence of Chmp1a

(A) Purkinje cell layer is intact in *Chmp1a* null P0 pups and unchanged compared to littermate controls. Immunostain for Calbindin labels Purkinje cells in control and *Chmp1a* null P0 cerebellum. Fraction of cerebellum area covered by the Purkinje cell has no detectable difference in *Chmp1a* null pups compared to controls (Control = 0.36 (n=4) vs *Chmp1a* null (n=2) 0.33, two-tailed t-test, p = 0.16) (B) *Shh* signaling is intact in *Chmp1a* null MEFs. MEFs were stimulated with SAG and *Gli1* expression increase was used to measure *Shh* pathway activation. There is no detectable difference between *Chmp1a* null and littermate control MEFs (Control: 76.9x (n=4) vs *Chmp1a* null: 51.5x (n=4), two-tailed t-test, p = 0.39).
**CHMP1A is required for exosome and SHH secretion *in vitro***

To identify the mechanism of regulation of SHH secretion by CHMP1A, we assayed exosome release in *CHMP1A* null HEK293 cells, generated by CRISPR-Cas9 mutagenesis (Figure 2.13A), and found a dramatic defect in exosome secretion (Figure 2.12). We transfected cells with a human *SHH* expression plasmid and collected extracellular vesicles (EV) from the supernatant by sequential ultracentrifugation (Figure 2.12C). We found no detectable difference in total protein in the EV pellet from *CHMP1A* null cells (2.51 vs 2.93 mg/ml, WT: n=5, *CHMP1A* null, n=5, two-tailed t-test, p = 0.46; Figures 2.12D, E); however, we found a 49% reduction in TSG101 in EVs from *CHMP1A* null cells suggesting a defect in exosome secretion (ratio: *CHMP1A* null / WT, n=4, two-tailed t-test, p = 0.0012; Figures 2.12D, E). TSG101 is a classical marker of exosomes (Bobrie et al., 2012), though some recent evidence suggests it can be found on microvesicles as well, another type of EV (Cocucci and Meldolesi, 2015). In parallel, we found a 55% reduction in SHH in the EV pellet suggesting SHH is secreted on exosomes (ratio: *CHMP1A* null / WT, n=5, two-tailed t-test, p = 0.0014; Figures 2.12D, E). We then examined the EV pellet with immunogold transmission EM (TEM) and visualized vesicles with CD63 protein on the surface that were about 70 nm in diameter (Figure 2.12F), two specific findings showing purification of exosomes (Cocucci and Meldolesi, 2015). Immunogold TEM also showed SHH immunoreactivity present on the surface of ~70 nm vesicles (Figures 2.12F, 2.13D), further confirming exosome-mediated SHH secretion, consistent with a recent report of *SHH*-expressing HEK293 cells releasing SHH on exosomes (Vyas et al., 2014). CHMP1A protein is also present on the surface of some of these exosomes (Figures 2.13B, C).
Exosome are EVs secreted when an MVB fuses with the plasma membrane and releases its ILVs into the extracellular space (Budnik et al., 2016). We measured exosome formation in the absence of CHMP1A by counting the number of ILVs in each MVB. We found a 32% reduction in ILVs per MVB in CHMP1A null 293HEK cells compared to WT (19 vs 13; WT: n=41, CHMP1A null: n=46, Mann-Whitney test, p = 0.0012; Figures 2.12G, H) providing evidence for impaired exosome biogenesis. These data provide evidence that SHH is secreted on exosomes through a process that requires CHMP1A.

Other ESCRT functions are unaffected by absence of CHMP1A

ESCRT-III has several functions in the cell but other functions appeared intact in the absence of CHMP1A. To evaluate ESCRT-III regulation of membrane abscission during cytokinesis (Carlton et al., 2012), we measured the fraction of mitotic cells in cytokinesis by immunostaining early mitotic cells with pH3 and cells in cytokinesis with aurora A at the midbody, and detected no abnormalities in cytokinesis progression in Chmp1a null mouse cortical progenitors or CHMP1A null human iPSCs (Figure 2.14A). In addition, Propidium iodide staining for DNA content was unchanged in Chmp1a null and control MEFs (Figure 2.14B), providing further evidence for normal cytokinesis in the absence of CHMP1A. Next, we measured ESCRT-mediated EGFR receptor degradation (Slagsvold et al, 2006) by transfecting HeLa cells with CHMP1A or control siRNA and then incubating the cells with EGF, which induces rapid removal of EGFR from the cell membrane and degradation over several hours (Ma et al., 2007). EGFR degradation was unaffected by CHMP1A depletion (0.40 vs 0.35 EGFR remaining, Control: n=2, CHMP1A siRNA: n=2, two-tailed t-test, p = 0.82; Figures 2.14C, 2.14D).
Figure 2.12: CHMP1A is essential for exosome-mediated SHH secretion

(A) CSF was collected from the 4th ventricle of control and Chmp1a null embryos at E14.5 and SHH concentration measured by ELISA. (B) CSF concentration of SHH is reduced 38% in Chmp1a null embryos. (C) EVs isolated from HEK293 cell culture supernatant via ultracentrifugation. (D) Ponceau stain of EV pellet shows no change in total protein between wild type and CHMP1A null cells, while immunoblot of exosome-associated protein TSG101 and immunoblot of SHH show reduced secretion in the EV pellet of CHMP1A null cells. (E) Quantification of immunoblots shows no detectable difference in total EV protein between CHMP1A null cells and control, TSG101 is reduced by 49%, and SHH is reduced by 55%. (F) Immunogold-EM of EV pellet shows a mixture of CD63-positive exosomes (black arrowhead) and CD63-negative EVs (white arrowhead). Immunogold-EM shows SHH on the surface of EVs. (G and H) MVBs in HEK293 cells (arrowheads). CHMP1A null HEK293 cells have 32% fewer ILVs per MVB than wild type. Two-tailed t-test (B and E): ** p < 0.01; Mann-Whitney test (H): ** p < 0.01.
Figure 2.12 (Continued)

A

E14.5
CSF
CSF collection
Sonic hedgehog
ELISA

B

CSF Shh (pg/ml)

Control
Chmp1a null

C

293HEK

300-g
2000-g
10Kg
100Kg
100Kg

PBS

D

CHMP1A
WT
Null

EV pellet protein
(Ponceau)

TSG101

SHH

E

EV pellet protein
(normalized)

Wild type
CHMP1A null

WT null
WT null

Blot: TSG101
Blot: SHH

F

CD63

SHH

G

Wild-type

CHMP1A null

H

Number of ILVs per MVB

Wild type

CHMP1A null
Figure 2.13: CHMP1A and SHH are present on the surface of extracellular vesicles

(A) CHMP1A null HEK293 cell line Sanger sequencing shows an early compound 1 bp and 2 bp frameshift deletion. This mutation creates a stop codon at amino acid 33 (V33X) and a peptide with a different sequence that is terminated past the original stop codon. No protein is detected by immunoblot. (B) CHMP1A Immunoblot of EVs purified from HEK293 supernatant shows CHMP1A protein is present in EVs. Staining in CHMP1A null EVs detects no protein. (C) Immunogold staining for CHMP1A on purified EVs shows CHMP1A protein present on the surface of EVs. Staining of CHMP1A null EVs shows no staining. (D) Immunogold for SHH on EVs purified from cells transfected with hsSHH shows SHH present on the surface of EVs (similar to Figure 2.12) and staining of mock-transfected cells shows no signal.
Figure 2.14: CHMP1A is not required for cytokinesis or EGFR degradation

(A) Cytokinesis disruption was not detected in the absence of CHMP1A. Ventricular surface of mouse cortex and cultured iPSCs were immunostained for pH3 and aurora A (images show mouse cortex ventricular surface). pH3 labels early mitotic cells and aurora A labels the midbody of late mitotic cells. The fraction of mitotic cells in cytokinesis was calculated as: (Aurora A) / (Aurora A + pH3). There was no detectable difference in the fraction of cells in cytokinesis in ventricular cortical progenitors of Chmp1a null embryos compared to littermate controls (0.43 vs 0.48, Control: n=8, Chmp1a null: n=7, two-tailed t-test, p = 0.093). There was no detectable difference in fraction of cells in cytokinesis in CHMP1A null and WT cultured iPSCs (0.64 vs 0.67, Control: n=15, Chmp1a null: n=15, two-tailed t-test, p = 0.59). (B) Propidium iodide staining of Chmp1a null and control MEF lines. There was not detectable difference in distribution of DNA content between Chmp1a null and littermate control MEFs. (C) There was no detectable difference in EGFR degradation in CHMP1A depleted HeLa cells compared to control. (D) CHMP1A siRNA depletes all CHMP1A detected by immunoblot.
Defective ILV and MVB structure in Chmp1a null embryos

To estimate exosome release in vivo, we counted the number of ILVs within choroid plexus epithelial cell MVBs with transmission electron microscopy (Figure 2.15A-C, 2.16C). In Chmp1a null embryos, the average number of ILVs per MVB is reduced from 18 to 12 (Control: n=88, Chmp1a null: n=39, Mann-Whitney test, p = 0.0003; Figures 2.15C, 2.15D, 2.16C). In addition, some Chmp1a null MVBs contain very large ILVs (Figure 2.15C), another consequence of impaired ILV budding (Lee et al., 2007). We did not observe any defects of the epithelial surface or microvilli in Chmp1a null embryos (Figures 2.16A, 2.16B), showing that the ChP is otherwise normal. These results provide in vivo evidence for decreased SHH secretion from ChP in the absence of CHMP1A as a result of decreased ILV generation within MVBs and thus, exosome release.
Figure 2.15: Choroid plexus epithelial cell MVBs

(A) Hindbrain choroid plexus from Chmp1a null and littermate controls at E17.5 was imaged with transmission electron microscopy. (B) MVB in an epithelial cell approaching the ventricular surface where it will fuse and release exosomes into the CSF. (C) TEM reveals defects in MVB formation in Chmp1a null embryos. Chmp1a null MVBs have fewer ILVs than controls and contain occasional abnormally large ILVs. (D) The number of ILVs in each MVB is reduced by 33% in Chmp1a null choroid plexus. Two-tailed t-test, *** p < 0.001.
Figure 2.16: Choroid plexus epithelial surface is normal in Chmp1a null mice

(A) There is no detectable defect in ventricular surface of the hindbrain choroid plexus in Chmp1a null embryos at E17.5. (B) There is no detectable defect in choroid plexus epithelial cell microvilli in Chmp1a null embryos at E17.5. (C) More examples of choroid plexus epithelial cell MVBs in control and Chmp1a null littermates. Chmp1a null MVBs contain fewer ILVs and often contain large ILVs. Both these findings point to a defect in ILV budding, a consequence of ESCRT-III dysfunction.
Scanning electron microscopy reveals MVB fusion and exosome release in vivo

To better characterize the mechanism of MVB fusion and exosome release in choroid plexus epithelial cells, we turned to scanning electron microscopy (SEM) to image the choroid plexus ventricular surface. SEM revealed the beautiful three-dimensional organization of the choroid plexus (Figure 2.17A). A single layer of epithelial cells forms into grape-like clusters that maximize ventricular surface area. Higher magnification reveals that each epithelial cell’s ventricular surface is covered in microvilli, to dramatically increase ventricular surface area, as well as clusters of motile cilia to drive CSF flow (Figures 2.17B-D). Exosomes are secreted when a MVB fuses with the plasma membrane and releases its ILVs into the extracellular space. SEM showed rare exosome release events on the ventricular surface in which clusters of small vesicles were visible protruding out into the ventricular space (Figures 2.17E, F). In the absence of CHMP1A, fewer vesicles were seen at these release sites (Figure 2.18A). Capturing the exact moment of exosome release was quite rare however, and so we looked for structures that might represent MVB fusion prior to exosome release (Figures 2.17G, 2.18A). We observed abundant spherical membranous protrusions, whose shape and size resembled MVB fusion events and hence, may represent MVBs prior to fusion. These structures were 63% smaller in CHMP1A null embryos consistent with our observation of fewer exosomes at active release sites (1.22 vs 0.44 um diameter, Control: n=13, Chmp1a null: n=3, two-tailed t-test, p = 0.030; Figure 2.18A). Aside from exosome release, the choroid plexus ventricular surface is normal in Chmp1a null embryos (Figure 2.18B).
Figure 2.17: Exosome are released from hindbrain choroid plexus

(A) SEM reveals three-dimensional organization of mouse embryonic hindbrain choroid plexus at E17.5. A single layer of epithelial cells forms into grape-like clusters that maximize ventricular surface area. (B-D) The ventricular surface of single epithelial cells is covered with microvilli and motile cilia. (E, F) Exosome release from MVB fusion with ventricular surface of a choroid plexus epithelial cell. (G) Spherical protrusion on choroid plexus ventricular surface, which is likely precursor to exosome release.
Figure 2.17 (Continued)
Figure 2.18. CHMP1A is essential for exosome release in hindbrain choroid plexus

(A) Spherical protrusions on ventricular surface are fusing MVBs that open to release exosomes. Fewer exosomes are released by fusing MVBs in Chmp1a null ChP. Spherical protrusion diameter is decreased by 64% in Chmp1a null ChP providing evidence for decreased exosome release. (B) The choroid plexus ventricular surface is grossly normal in Chmp1a null embryos. Motile cilia and microvilli are unchanged compared to control.

Abnormal MVBs and exosomes in neuronal dendrites in Chmp1a mutants

In addition to the choroid plexus, SHH and CHMP1A are co-expressed in cerebellar Purkinje cells and layer V pyramidal neurons during brain development. SHH acts as a mitogen to drive GCP proliferation during cerebellum development (Corrales et al., 2004, Dahmane and Ruiz i Altaba, 1999, Lewis et al., 2004) and is required for synapse formation between layer V corticofugal projection neurons and colossal projection neurons (Harwell et al., 2012). TEM shows that MVBs are abundant in dendrites of Purkinje cell and cortical projection neurons. In mouse P4 cerebellum,
Purkinje cells MVBs are common in dendrites pointing towards the EGL where SHH responsive GCPs are located (Figure 2.19A, 2.20A). In mouse cerebral cortex, we examined serially reconstructed TEM images of cortical pyramidal neurons (Lee et al., 2016) from adult cortex, and found that MVBs were surprisingly abundant in pyramidal cell dendrites (Figure 2.19B). The dendritic arbor of a single pyramidal cell contained at least 80 MVBs, often present near synapses, suggesting that exosome-mediated secretion may be very active in dendrites.

Like the choroid plexus, Purkinje cell MVBs in Chmp1a null mice show a decrease in the number of ILVs per MVB, providing further evidence that defective SHH secretion in Chmp1a mutants reflects abnormal exosome function. TEM of Chmp1a null P0 pups showed that MVBs in Purkinje cells have 25% fewer ILVs per MVB compared to control littermates (Figures 2.19C, 2.20B, 10.4 vs 8.1, Control: n=88, Chmp1a null: n=74, Mann-Whitney test, p = 0.0005), further supporting the model that SHH secretion from Purkinje cells is exosome mediated and disrupted in the absence of CHMP1A.

Together our results suggest a model for SHH secretion during CNS development. During embryogenesis, SHH is secreted from choroid plexus epithelial cells when MVBs fuse with the ventricular surface and release exosomes into the CSF (Figure 2.19D), providing a pool of CSF-SHH that can regulate proliferation of the basal forebrain and early cerebellum. During postnatal cerebellum development, exosomes are released by Purkinje cell dendrites directly into the EGL, with SHH on these exosomes driving GCP proliferation in the EGL. We postulate that later, during cortical synapse development, MVBs in pyramidal cell dendrites release exosomes to secrete SHH or other growth factors to drive synapse formation.
Figure 2.19: Exosomes in cerebellum and cortex

(A) MVBs are in distal (2) and proximal (3) Purkinje cell dendrites in developing cerebellum (P4). (B) MVBs in dendrites of layer II/III cortical pyramidal neurons. High magnification image shows a MVB (white arrowhead) near an excitatory synapse in a pyramidal cell dendritic spine and near an inhibitory synapse in a dendritic shaft. MVBs locations are represented by red dots in a rendered 3D TEM reconstruction of a serially sectioned pyramidal neuron. (C) Chmp1a null Purkinje cell MVBs contain 22% fewer ILVs than littermate controls. (D) Model of CHMP1A and exosome-mediated SHH secretion. Mann-Whitney test, *** p < 0.001.
Figure 2.19 (Continued)
Figure 2.20: MVBs localize to dendrites in cerebellar Purkinje cells

(A) MVBs are located in Purkinje cell dendrites in developing cerebellum. Transmission EM of the Purkinje layer and EGL at P0 shows MVBs in Purkinje cell dendrites. (B) Additional examples of Purkinje cell MVBs from \textit{Chmp1a} null P0 pups and littermate controls. \textit{Chmp1a} null MVBs contain fewer ILVs than controls.
2.4 Discussion

**CHMP1A function in brain development**

Loss of function mutations in human *CHMP1A* cause a syndrome of microcephaly, severe pontocerebellar hypoplasia and short stature (Mochida et al., 2012). Using *CHMP1A* null human cerebral organoids and *Chmp1a* null mice, we have demonstrated that CHMP1A is essential for brain development because it is required for progenitor proliferation in the cortex, ganglionic eminences, and cerebellum. CHMP1A is required for progenitor proliferation because it regulates exosome-mediated SHH secretion. We have shown that SHH is secreted on exosomes, and that in the absence of CHMP1A fewer exosomes are released and MVBs contain fewer intraluminal vesicles. Our findings provide evidence for a new mechanism of SHH secretion in vertebrate brain development that has broad potential relevance to understanding secretion of other growth factors and bioactive molecules in the CNS.

**CHMP1A specifically impairs exosome secretion**

We provide evidence that CHMP1A is essential for brain development through a specific function regulating exosome release. CHMP1A is a member of the ESCRT-III complex but our evidence suggests that the defect in ESCRT function in the absence of CHMP1A is specifically exosome release. In particular, we have not detected a defect in activated receptor degradation or in cytokinesis, two other ESCRT complex functions that could regulate progenitor proliferation. Instead, we observe a dramatic decrease in exosome-mediated SHH secretion. *Hrs* and *Tsg101* mouse knockouts, with nearly
complete loss of ESCRT complex function have very early embryonic lethality (Komada and Soriano, 1999; Ruland et al., 2001; Wagner et al., 2003). This contrasts with Chmp1a null mice, which survive through late embryogenesis. It is possible that because loss of CHMP1A causes only partial loss of ESCRT function (25-35% reduction of ILVs compared to over 80% reduction in Chmp1a/1b double knockout Arabidopsis (Spitzer et al., 2009) exosome mediated secretion of growth factors is dosage-sensitive and is disrupted before other canonical ESCRT functions. Chmp1a loss of function likely only causes partial loss of ESCRT function because of redundancy with other ESCRT-III components, specifically Chmp1b. CHMP1B protein has 56% amino acid similarity to CHMP1A, and so its expression likely spares complete loss of ESCRT function and early embryonic lethality of Chmp1a null mice.

The existence of some specificity in CHMP1A function is suggested by the early embryonic lethality in mice lacking all ESCRT function (Komada and Soriano, 1999; Ruland et al., 2001; Wagner et al., 2001), and the finding that mice with LOF of Chmp2b, another ESCRT-III member, are viable with postnatal defects in dendrite morphology (Chassefeyre et al., 2015). CHMP2B protein localizes in dendrites of cultured neurons and Chmp2b mutant mice have decreased spine density and spine maturation in the stratum radiatum of the hippocampus (Chassefeyre et al., 2015). However, the authors argue that this is not mediated by CHMP2B’s function in MVB biogenesis or exosome release. Instead they propose CHMP2B complexes with other ESCRT-III proteins to form stable helices at the postsynaptic density of dendritic spines, which promote spine stability (Chassefeyre et al., 2015). To our knowledge, our finding that CHMP1A
mediated exosome secretion is necessary for brain development is specific to Chmp1a loss of function.

**Exosome release and exosome-mediated SHH secretion are specifically disrupted by loss of CHMP1A**

CHMP1A appears to be required for exosome secretion far out of proportion to its requirement for release of other extracellular vesicles (EV). Exosomes are a specific class of EVs defined by marker proteins CD63 and TSG101. In contrast to other EVs, exosomes release requires ESCRT-III function (Budnik et al., 2016). In the absence of CHMP1A, we found no detectable difference in total EV secretion (*Figure 2.12*). Instead, we found a 49% reduction in TSG101 positive exosomes, showing that CHMP1A is preferentially, if not specifically, required for exosome secretion. We found a parallel 55% reduction in secreted SHH in purified EVs and that SHH protein is present on EV surface. Together these results show that SHH secretion is mediated at least in part by exosomes and requires CHMP1A.

Our *in vivo* results also provide strong evidence that SHH secretion is mediated by exosomes. During mid embryogenesis, hindbrain choroid plexus is one source of CSF SHH (Nielsen and Dymecki, 2010). In Chmp1a null embryos, the concentration of secreted SHH in the CSF is reduced by 40%; in parallel, hindbrain choroid plexus epithelial cells have a 35% reduction in exosome release as quantified by the number of ILVs per MVB. Together this data strongly suggests that during mouse brain development Shh secretion is exosome mediated.
Mammalian exosome-mediated SHH secretion is distinct from previously reported Hh secretion via exovesicles

Exosome-mediated SHH secretion in cerebellum and choroid plexus seems to differ from Hedgehog (Hh) secretion in Drosophila imaginal disk, where a recent report showed that Hh is secreted on larger exovesicles from posterior cells (Matusek et al., 2014). When ESCRT function is impaired, Hh accumulates at the surface of posterior cells and cannot be released, consistent with what is known about exovesicles (also called ectosomes or microvesicles), a class of large EVs that bud from the cell membrane through ESCRT-mediated vesicle formation (Budnik et al., 2016). Exovesicle-mediated impairment of Hh secretion is not associated with defects in MVB biogenesis and in turn, disruption of MVB biogenesis does not impair Hh secretion in flies. Our findings suggest a slightly different mechanism for SHH secretion in vertebrate brain. We show that Chmp1a null mice have fewer MVB ILVs, indicating a defect in MVB biogenesis, and that SHH is secreted on exosomes, ILVs that are released when an MVB fuses with the membrane, as opposed to exovesicles, which bud from the cell surface. Consistent with a slightly different mechanism for SHH secretion in vertebrate brain, exovesicles described in fly as Hh carriers are 100-400 nm in diameter, while the exosomes we observe as SHH carriers are much smaller, 50-100 nm, a size difference that reflects different vesicle origins. Larger vesicles (300-500 nm) reported as SHH carriers in the mouse ventral node, also likely correspond to exovesicles (Tanaka et al., 2005), suggesting at least two different specialized secretion mechanisms that both contribute to SHH/Hh secretion. Perhaps exovesicle Hh secretion is specific to invertebrates, limb/wing development, and the vertebrate ventral node, while exosome-mediated SHH
secretion occurs in vertebrate brain. Such multiple mechanisms might explain the absence of SHH-related limb and L-R body axis defects in *Chmp1a* mutant mice.

**Exosome-mediated SHH secretion is consistent with requirement for DISP1 and SCUBE2**

In addition to specific physical mechanisms for SHH secretion, previous work has identified two proteins, DISP1 (dispatched) and SCUBE2, that are required for SHH secretion. Although we have not looked specifically at these proteins, our results are consistent with what is known about DISP1 and SCUBE2’s role in SHH secretion. DISP1 and SCUBE2 are transmembrane proteins that bind to SHH and are required for SHH release (Creanga et al., 2012; Ma et al., 2002). SHH is hydrophobic, and so tethering to the membrane via DISP1 and SCUBE2 enables SHH targeting to lipid membranes. We propose that DISP1 and SCUBE2 incorporate into ILV membranes and keep SHH attached to the outside of ILVs, which would be consistent with previous results of WNT secretion. WNT, a hydrophobic protein like SHH, requires EVI, a transmembrane protein, for release. EVI is incorporated into ILV membranes and found on released exosomes along with WNT (Korkut et al., 2009). This suggests that EVI is required for WNT secretion because it targets and/or tethers WNT to exosomes. As DISP1 and SCUBE2 are transmembrane proteins like EVI, and are required for SHH secretion, it is reasonable to propose that they tether SHH to the ILV/exosome membrane. This model shows that exosome mediated SHH secretion is compatible with previous work on the biochemical components of SHH release.
CHMP1A regulates INK4A indirectly

Previous work from our lab found that CHMP1A is required for repression of the pro-senescence protein INK4A, mediated by a genetic interaction with the nuclear repressor BMI1 (Mochida et al., 2012). With improved methods for localizing CHMP1A in the cell, our current work suggests that CHMP1A protein is not detectable in the nucleus, and that it regulates INK4A indirectly. SHH signaling is known to activate BMI1 repression of INK4A to maintain proliferation in progenitor cells (Subkhankulova et al., 2010), and so a defect in SHH secretion in the absence of CHMP1A could decrease SHH signaling in responsive cells and, in turn, prevent INK4A repression. Thus our current findings complement our previous work, but suggest that CHMP1A regulation of INK4A repression is actually indirect.

Chmp1a null phenotype is consistent with hypomorphic Shh signaling

The several developmental defects—including a thin cerebral cortex, small basal ganglia, cerebellar hypoplasia, small somatic size and perinatal lethality—are all consistent with partial loss of Shh, due to CHMP1A’s function as a member of ESCRT-III that is required for exosome-mediated SHH secretion. Small basal ganglia, cerebellar hypoplasia, small somatic size, and perinatal lethality are all consistent with previous mouse models of decreased Shh signaling through hypomorphic alleles, Shh-N and Shh-flox, in particular (Huang et al., 2007). Although thinning of the cerebral cortex has not been specifically described, basal forebrain defects are present in mouse models with defective Shh signaling (Huang et al., 2007), and SHH has recently been implicated in regulating proliferation of progenitors in the subventricular zone of developing cerebral
cortex (Wang et al, 2016). On the other hand, we do not see digit or spinal cord patterning defects in Chmp1a null mice that are characteristic of null Shh mutations (Zhu and Mackem, 2011; Fuccillo et al., 2006). This is consistent with the fact that we do not observe a complete blockage of SHH secretion, or a complete absence of SHH protein in CSF, suggesting that Shh signaling is only partially reduced in the absence of CHMP1A. We have not formally defined whether the Chmp1a null phenotype reflects loss of only exosome-mediated SHH release, with other mechanisms regulating SHH release in other contexts, or whether Chmp1a mutation may partially impair SHH secretion broadly.

**Future experiments to test CHMP1A regulation of Shh signaling**

We have presented evidence that SHH secretion is reduced in the absence of CHMP1A during mouse brain development and demonstrated defects in cortex, basal ganglia, and cerebellum development in Chmp1a null mice consistent with decreased Shh signaling. We also showed that expression of Ptch, a target of Shh signaling, is decreased in Chmp1a null cerebellum. There are two primary future experiments to provide further evidence for CHMP1A regulation of Shh signaling.

First, I will determine if defects in Chmp1a null embryonic brain development can be rescued by increasing Shh signaling with the Ptch1 mutant mouse (Goodrich et al., 1997). When one allele of the downstream Shh repressor Ptch1 is mutated, Shh signaling increases and this has been shown to partially rescue cerebellar hypoplasia in mice with defective Shh signaling (Dutka et al., 2015). I will establish a Chmp1a/ Ptch1 double heterozygous mouse line and cross these mice to obtain Chmp1a null mice heterozygous for the Ptch1 mutation. I will measure survival, brain anatomy, and progenitor
proliferation in these mice to determine if increased Shh signaling rescues the defects caused by absence of CHMP1A.

Second, I will confirm that CHMP1A acts in Purkinje cells to regulate cerebellum development. My results show that CHMP1A is required for SHH secretion and CHMP1A is expressed in Purkinje cells; these findings suggest that defective cerebellum growth and GCP proliferation in Chmp1a null mice is a result of decreased SHH secretion from Purkinje cells. I will test this conclusion by creating conditional null mice with Chmp1a specifically deleted from Purkinje cells. I will order ES cells for a floxed conditional Chmp1a deletion allele and then cross carriers of this allele to L7-Cre mice. L7 driven Cre expression will remove CHMP1A specifically from Purkinje cells in the cerebellum (Lewis et al., 2004). I will measure cerebellum size and GCP proliferation in these mice to confirm removal of CHMP1A from Purkinje cells decreases cerebellum size and reduces GCP proliferation. In parallel, I will use Math1-Cre (Machold and Fishell, 2005) mice to remove CHMP1A specifically from GCPs, the SHH-receiving cells, to confirm that Chmp1a null cerebellar hypoplasia is not a result of impaired Shh signaling activation in receiving cells.

**Defective exosome release also disrupts WNT secretion**

We present evidence that CHMP1A is required for secretion of SHH during brain development. In addition to SHH, CHMP1A likely regulates exosome-mediated secretion of other growth factors. In particular, there is evidence that WNT3A and WNT family proteins, secreted growth factors essential for cortical progenitor proliferation, are secreted on exosomes in *Drosophila* and in human cells (Gross et al., 2012; Korkut et al.,
Like SHH, WNT proteins are cholesterol modified and lipophilic, and require a specialized secretion mechanism. In Chmp1a null embryos, reduced cerebral cortical progenitor proliferation and cortical thinning are likely caused by a combination of decreased SHH and WNT secretion, as both signaling pathways drive proliferation in the cortex. Consistent with a potential defect in WNT secretion, we found decreased expression of several downstream WNT signaling targets in CHMP1A null cerebral organoids.

**Neuronal SHH secretion via exosomes**

Our results show exosome-mediated SHH secretion occurs in cerebellar Purkinje cells in addition to choroid plexus epithelial cells. SHH secreted by Purkinje cells is the primary mitogen for GCP proliferation in the developing cerebellum (Corrales et al., 2004; Dahmane and Ruiz i Altaba, 1999; Lewis et al., 2004), and GCP proliferation is substantially reduced in Chmp1a null mice. MVBs are widespread in Purkinje cell dendrites, and in the absence of CHMP1A, Purkinje cell MVBs have 25% fewer ILVs. Together, these results provide evidence that Purkinje cells dendrites secrete SHH on exosomes to drive GCP proliferation and that this process requires CHMP1A. Exosome mediated SHH secretion from post-mitotic neurons is likely widespread in the developing brain. In addition to Purkinje cells, SHH protein is expressed in dendrites of mitral and tufted cells in the developing olfactory bulb (Gong et al., 2009) and dendrites of layer V pyramidal cells in the developing cortex (Harwell et al., 2012). During cortical synapse formation (second postnatal week in mice), SHH protein is expressed in the dendrites of layer V corticofugal projection neurons and BOC, a SHH co-receptor, is expressed in
callosal and local projection neurons which form synapses with corticofugal projection neurons (Harwell et al., 2012). Analysis of conditional KO mice of cortical Shh showed that SHH expression in corticofugal projection neurons is required for formation of these synapses. We have shown that CHMP1A is expressed in post-mitotic cortical neurons and that MVBs are abundant in dendrites of cortical pyramidal cells. These results suggest a potential mechanism of spatially constrained retrograde SHH secretion that might have widespread importance for neuronal connectivity.
2.5 Experimental Procedures

**Human Tissue**

Human tissue collection and use was approved by Boston Children’s Hospital IRB.

**Animal Use**

Animals were cared for humanely and all experiments were approved by Boston Children’s Hospital IACUC.

**Patient CH-2701**

A dermal skin biopsy was obtained from Patient CH-2701. This biopsy was used to generate a fibroblast cell line.

**iPS line generation**

Skin fibroblasts from Patient CH-2701 were transformed in iPS cells with an integration free episomal introduction of the Yamanaka factors described previously (Schlaeger et al., 2014). iPS cells were identified by morphology. IMR-90 iPS cells were obtained from WiCell. All iPSC lines were maintained in mTESR media and grown on matrigel.

**pH3 analysis**

Matching coronal telencephalon sections at E12.5 or E14.5, or midline sagittal cerebellum sections at P0 were immunostained for pH3. pH3+ cells lining the cortical ventricular surface were counted at E14.5, pH3+ cells in the MGE and LGE were counted at E12.5, and pH3+ cells in the cerebellar EGL were counted at P0.

**Immunostaining and immunoblot**
Immunostain of human choroid plexus, organoids, and mouse brain and immunoblot of purified exosomes was done according to standard protocols. See Supplemental information.

LacZ staining

Chemical staining for beta-galactosidase activity was performed with the beta-gal staining kit from Invitrogen (K146501). Briefly, tissue was fixed overnight in 2% gluteraldehyde, microtome sectioned (70 uM), and then stained for beta-gal activity according to the kit instructions.

In situ hybridization

In situ hybridization was performed as previously described (Arlotta et al., 2005). Ptch in situ probe was a gift from C Cepko and A Joyner. RNA was synthesized with the Megascript kit from Invitrogen. DIG dNTPS and anti-DIG Fab fragments were ordered from Roche.

CSF collection and Shh ELISA

CSF was collected from the 4th ventricle of E14-E15 Chmp1a null mouse embryos and littermate controls using a pulled micropipette. CSF was centrifuged at 10000 G for 5 minutes and then used in the Shh ELISA. Shh ELISA kit was purchased from R and D systems (MSHH00) and used according to the manufacture’s instructions.

Statistics

Mean values are presented for pooled data and errors bars are SEM. Statistical significance was determined using a two-tailed, unpaired t-test or a Mann-Whitney test (* p < 0.05, ** p < 0.01, *** p < 0.001). All counts of ILVs per MVB failed the D’Agostino and Pearson normality test (p < 0.01, expect HEK293 WT: p = 0.17). As a result, the
nonparametric Mann-Whitney test was used to test significance of ILV per MVB differences. All other experiments used a two-sample t-test, except Figure 5F TSG101 and SHH, which used a one-sample t-test.

**Chmp1a GT mouse generation and mouse breeding**

Mouse ES cells with a gene trap cassette inserted into Chmp1a were obtained from BayGenomics (B6;129P2-Chmp1a<sup>Gr(XC472)Byg/Mmucd</sup>) and injected into blastocysts of WT mice. Resulting chimeras were out crossed with WT C57/Bl6 mice to generate heterozygous GT mice. Heterozygous GT mice were backcrossed to C57/Bl6 for 7-8 generations. Mouse DNA was genotyped with the following primers: WT primer F: GAGACAGCGGGGTCCGTAAC, WT primer R: AACACACACTCGAACCGAAAG, GT primer F: GAGACAGCGGGGTCCGTAAC, GT primer R: GGTCCTAGTGGGAGGTCTCG.

**CHMP1A null cell line generation**

HEK293 CHMP1A null line and iPSC CHMP1A null lines were generated using CRISPR-Cas9 mutagenesis (Veres et al., 2014) Cas9 was expressed from a plasmid encoding Cas9 and GFP (Gift of Chad Cowan and Kirin Musunuru). gRNA targeting CHMP1A was expressed from a co-transfected plasmid, hsCHMP1A protospacer sequence: GAAGGACTCCAAGGCGGAGC. GFP positive cells were grown as single colonies isolated and then sequenced to identify homozygous frame shift mutations in CHMP1A. Controls were cells from a similar CRISPR experiment targeting an unrelated gene EXOC7. Sanger sequencing showed that these cells were WT for both the target gene and CHMP1A, and so were used as controls. Primers used to genotype were F: GAAGACAGACACTGGGAGAAAACC R: CAGAAGACAAACCAGGAGGTCA.
**Human choroid plexus immunostaining**

Choroid plexus tissue was fixed in 4% PFA, frozen, and sectioned at 15 um on a cryostat. Cultured iPSCs were fixed with 4% PFA for 10 minutes at RT. Tissue or cells were permeabilized with 0.04% Tween in PBS and blocked in 0.04% tween, 2.5% donkey serum, and 2.5% goat serum in PBS. Sections or cells were incubated with primary antibody diluted in blocking buffer overnight at 4C and then stained with Alexa secondary antibodies and Hoechst. Imaging was done on Zeiss 510 confocal microscope.

**Mouse brain and human organoid immunostaining**

Tissue was fixed overnight at 4 C in 4% PFA and sectioned at 70um using a Vibratome. Antigen retrieval was performed with Retrievagen A. Tissue was permeabilized and blocked in 3% BSA, 0.3% Triton X-100, 0.3% sodium azide in PBS. Primary antibodies were diluted in blocking buffer and incubated overnight at 4 C. Sections were then stained with Alexa secondary antibodies and Hoechst. Imaging was done on Zeiss 510 confocal microscope.

**Immunoblot**

Total cell protein was isolated following cell lysis with RIPA buffer. EVs were purified as described above. SDS sample buffer was added to the sample and then protein was subjected to gel electrophoresis. Protein was transferred to Immobulin (Millipore) membrane and incubated with primary and secondary antibodies in Licor blocking buffer. A Licor Odyssey fluorescent imaging system was used to image membranes and quantify protein. Primary antibodies used: CHMP1A, SHH, and TSG101. Secondary antibodies were purchased from Licor for use with the Odyssey system. For TSG101 and SHH blots and quantification, equal amount of EV protein was
blotted for WT and Chmp1a null cell lines. For SHH quantification, band intensity was normalized between the two cell lines by correcting for any differences in transfection efficacy detected by SHH immunoblot of whole cell lysate for each experiment.

**Tissue culture**

HEK293 and HeLa cells were maintained in DMEM high glucose supplemented with 10% FBS and 1% penicillin streptomycin. iPSCs were grown on Matrigel (BD) in mTESR media (Stem Cell Tech). For culture of fibroblast from skin biopsy media used was high-glucose DMEM supplemented with 15% FBS, L-glutamine, MEM-NEAA, sodium pyruvate, pen/strep, 50 µg/mL gentamicin and 2.5 µg/mL amphotericin B.

**Cerebral organoid culture**

iPSC lines were differentiated into cerebral organoids. Organoids for immunostaining were differentiated according to Lancaster et al., 2013.

Organoids for RNA sequencing were differentiated as follows. Full grown iPSC colonies were lifted of with a cell scraper and transferred to ultra low attachment culture dishes (Corning) containing EB Neural Differentiation media without additives (DMEM-F12 (Invitrogen) supplemented with 20% Knockout Serum Replacement (Invitrogen), 100 units/ml Penicillin/100 µg/ml streptomycin (Gibco), 2 mM Glutamax (Gibco), 1 x MEM non essential amino acids (Gibco), 1 mM sodium pyruvate (Invitrogen) and 0.1 uM beta mercaptoethanol (Gibco). Embryoid bodies were formed overnight and after 24 hours (day 0) media was replaced with EB Neural Differentiation media + additives (500 ng/ml DKK1 (peprotech), 500 ng/ml human recombinant NOGGIN (R & D Systems), 10 uM SB431542 (Sigma) and 1 uM Cyclopamine V. californicum (VWR). EB Neural Differentiation media + additives was replaced every other day. On Day 8, single
embryoid bodies (cortical neurospheres) were transferred to ultra low attachment U-shaped bottom 96 well plates (Corning) and cultured in 100 ul EB Neural Differentiation Media + additives. On day 18, media was replaced with Neurobasal/N2 media (Neurobasal (Invitrogen), 1x N2 supplement (Gibco), 2 mM Glutamax (Gibco), 100 units/ml Penicillin/100 ug/ml streptomycin (Gibco)) supplemented with 1 uM Cyclopamine. On day 26 media was replaced with Neurobasal/N2 media without cyclopamine. Cerebral organoids were further maintained in Neurobasal/N2 media until harvest.

Organoids were harvested at 2 or 5 weeks post differentiation (RNA sequencing) or 4 weeks post differentiation (immunostaining).

**RNA sequencing library preparation**

RNA was treated with RQ1 DNaseI (Promega) for 1 hour at 37 C and total RNA was cleaned up using the RNAeasy Mini kit (Qiagen). For each sample, the non-ribosomal fraction of 5 ug of total RNA was isolated using a Ribo-Zero rRNA removal Kit (Epicentre) following the manufacturer’s protocol (Lit.#309-6/2011). For the non-ribosomal fraction of RNA, 200 ng RNA was fragmented by incubating 30 min. at 98°C in RNA storage buffer (Ambion) and then double stranded (ds) cDNA was synthesized as described previously (Parkhomchuk et al., 2009) using dUTP in the second strand synthesis and USER digest before amplification to retain strand specificity. Clean-up steps were performed using RNA Clean & Concentrator or DNA Clean & Concentrator kits (Zymo research). Double stranded cDNA was used for library preparation following the Low Throughput guidelines of the TruSeq DNA Sample Preparation kit (Illumina), with the following additions. Size selections were performed before and after cDNA
amplification on an E-gel Safe Imager (Invitrogen) using 2% E-gel SizeSelect gels (Invitrogen). The cDNA fraction of 300-400 bp in size (including adapters) was isolated and purified. For adapter ligations, 1 ul instead of 2.5 ul of DNA Adapter Index was used. Indexed libraries were pooled and sequenced on the Illumina HiSEQ platform.

**Mapping of RNA sequencing data**

All samples were mapped using Tophat2 (Kim et al., 2011) using Bowtie2 (Langmead et al., 2012) as the underlying alignment tool. The input Illumina fastq files consisted of paired end reads with each end containing 100bp. The target genome assembly for the human samples was GRCh37/UCSC hg19, and Tophat was additionally supplied with a gene model (using its “-GTF” parameter) with data from the hg19 UCSC KnownGenes track (Hsu et al., 2006). For the macaque samples, the target genome assembly was MGSC Merged 10/UCSC rheMac2, and Tophat was supplied with gene model data from the rheMac2 RefSeq Annotations track (Pruitt et al., 2005). For multiply-mapped fragments, only the highest scoring mapping determined by Bowtie2 was kept. Only mappings with both read ends aligned were kept. Potential PCR duplicates (mappings of more than one fragment with identical positions for both read ends) were removed with the samtools “rmdup” (Li et al., 2009) function, keeping only one of any potential duplicates. The final set of mapped paired end reads for a sample was converted to position- by-position coverage of the relevant genome assembly using the bedtools “genomeCoverageBed” (Quinlan et al., 2010) function. To determine the count of fragments mapping to a gene, the position-by-position coverage was summed over the exonic positions of the gene. This gene total coverage was divided by the factor 200, to account for the 200 bp of coverage induced by each mapped paired end fragment
(100bp from each end), and rounded to an integer. This was calculated for each gene in the UCSC Known Gene set. For input to DESeq (Anders and Huber, 2010) all genes with non-zero counts in any sample were considered. Replicates of each week’s samples were combined per the DESeq methodology. Pairwise comparisons were made to determine significant differences in control and CHMP1A null organoids.

**GO pathway enrichment analysis**

For genes with at least a 2-fold change between wild type and CHMP1a null, enrichment of gene sets for GO- Biological Processes (BP-FAT) and KEGG-pathways was determined by DAVID (Database for Annotation, Visualization and Integrated Discovery) Functional Annotation Clustering (v6.7). For each analysis raw RNAseq read counts were normalized using DEseq2. For week 2 and week 5 respectively, all genes with mean read count >64 and read count >32 were included as background. Clusters of GO-terms and KEGG pathways were manually given a name that best represented all individual GO-term categories within each cluster. The DAVID Functional Annotation Tool provides an ‘enrichment score’ but does not provide a measure of statistical significance for Functional Annotation Clusters. For each Functional Annotation Cluster, the benjamini-corrected p-value for the highest ranking individual GO-term in each cluster is displayed next to the bar graphs.

**Generation of Chmp1a null MEFs**

*Chmp1a* GT heterozygous mice were crossed and E14-15 embryos were collected. Embryos were decapitated and visceral organs removed. The remaining tissue was dissociated with Trypsin and then plated and maintained in DMEM high glucose
supplemented with 10% FBS, 1% P/S, and 2 mM L-glutamine. MEFs were assayed before passage 4.

**Shh signaling in MEFs**

MEFs from *Chmp1a* null embryos and littermate controls were cultured as described above. The cells were then cultured in 0.5% FBS with 1 uM SAG for 48 hours. Then total RNA was isolated with RNeasy kit (Ambion) and cDNA was created with Superscript III kit (Invitrogen). Taqman qRT-PCR assay was run on StepOnePlus (Applied Biosystems) with probes for Axin2 and Beta-actin to measure gene expression.

**Exosome purification**

Extracellular vesicles were purified from cultured HEK293 cells according as previously described (Théry et al., 2006). Cells were transfected with human *SHH* expressing plasmid (gift of A Salic) using Lipofectamine 2000. The following day media was replaced with DMEM supplemented with exosome depleted FBS (Systems Biosciences: EXO-FBS-50A-1) and P/S. After 48 hours in exosome depleted media, the supernatant was collected. Cell culture supernatant was centrifuged at 300 g for 5 minutes to eliminate floating cells, 2,000 g for 20 minutes to remove dead cells, and 10,000 g for 40 minutes to remove cell debris. The cleaned supernatant was centrifuged at 100,000 g for 90 minutes to pellet EVs. EVs were resuspended in PBS and then centrifuged again at 100,000 g for 90 minutes to eliminate contaminating proteins. Purified EVs were resuspended in PBS.

**Cytokinesis analysis**

Mouse cortex at E14.5 and cultured iPSCs were fixed in 4% PFA. Cortical ventricular surface or cultured iPSCs were immunostained with antibodies against aurora
A kinase, pH3 and Hoechst. Aurora A labels the midbody in anaphase through telophase while pH3 labels histones from prophase through metaphase. A high power field was imaged and the number of pH3 positive early mitotic cells and the number of aurora A positive late mitotic cells was counted. The ratio of midbodies to pH3 positive cells was calculated as measure of cytokinesis duration.

**EGFR degradation assay**

Wild type HeLa cells were transfected with siRNA targeting *CHMP1A* (Thermo, #4392420) or a negative control (Thermo, #4390843) from Ambion Silencer Select according to the manufacturer’s protocol. Cells were grown for 48 hours, and then placed in serum free media for 2 hours. EGF was added to the media at 250 ng/mL and cells were fixed at 0 and 2 hours. Cells were stained with anti EGFR antibody (13A9, Genentech) for 1 hour and then Alex fluorescent secondary for 30 minutes. Fluorescence was then quantified at the single cell level for hundreds to thousands of cells using a BD FACS Aria flow cytometer. For *VPS4* experiment, HeLa cells were transfected with wild type or dominant negative (E228Q) human *VPS4* (gift from W. Sundquist) for 48 hours before the EGF stimulation.

**Electron microscopy**

Embryonic mouse choroid plexus and P0 mouse cerebellum was fixed in 2.5% Glutaraldehyde and 2% Paraformaldehyde in 0.1 M sodium cacodylate buffer (pH 7.4) at 4C overnight. 70 um thick tissue sections were cut on a vibratome. Sections were washed in 0.1 M cacodylate buffer and postfixed with 1% Osmiumtetroxide (OsO4) and 1.5% Potassiumferrocyanide (KFeCN6) for 1 hour, washed in water 3x and incubated in 1% aqueous uranyl acetate for 1 hour followed by 2 washes in water and subsequent
dehydration in grades of alcohol (10 min each; 50%, 70%, 90%, 2x 10 min 100%). The samples were then infiltrated for 15 min in a 1:1 mixture of propyleneoxide and TAAB Epon (Marivac Canada Inc. St. Laurent, Canada). The samples were embedded in drops of TAAB Epon between two sheets of aclar plastic (Electron Microscopy Sciences) and polymerized at 60°C for 48 hours. Ultrathin sections (about 80 nm) were cut on a Reichert Ultracut-S microtome, placed onto copper grids, stained with uranyl acetate and lead citrate and examined in a JEOL 1200EX transmission electron microscope. Images were recorded with an AMT 2k CCD camera.

SEM samples were postfixed in 1.0% osmium tetroxide in 0.1M cacodylate buffer (pH 7.4) for 1 hour at room temperature. Following postfixation, the samples were rinsed with buffer then dehydrated through a graded series of ethanol. The specimens were then critical point dried with CO2 using a Samdri PVT-3 critical point dryer (Tousimis Corp. Rockville, MD). The specimens were attached to specimen mounts using conductive adhesive tabs, coated with 5nm platinum using a Cressington 208HR sputter coater (Cressington Scientific Instruments, Ltd. Walford, UK).

TEM images were collected on a JEOL 1200EX at the HMS Electron Microscopy facility. SEM images were collected on a Hitachi S-4800 at Northeastern University.

**Immunogold EM**

Purified exosomes were fixed in 4% Paraformaldehyde in 0.1 M sodium cacodylate buffer (pH 7.4) for 10 minutes at RT. Primary antibodies (anti-CD63 and anti-SHH) were diluted in 1% BSA in PBS and centrifuged 1 minute at 14 000 rpm. Exosomes were applied to copper grid then grid was floated on a drop of 1% BSA for 10 minutes to block. Next, grids were moved to 5µl drops of diluted primary antibody and
incubated for 30 minutes at RT. Grids were washed with PBS for 15 minutes then transferred to 5µL drops of Protein-A gold (diluted in 1% BSA) for 20 minutes. Finally, grids were washed in PBS followed by distilled water. To image, immunogold stained exosomes were stained with uranyl acetate and examined in a JEOL 1200EX transmission electron microscope with an AMT 2k CCD camera.
Table 2.1: Antibodies used in this study

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Ab species</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1CAM</td>
<td>Rat</td>
<td>Millipore MAB5272</td>
</tr>
<tr>
<td>TUJ1</td>
<td>Mouse</td>
<td>Biolegend 801201</td>
</tr>
<tr>
<td>CHMP1A (blot)</td>
<td>Rabbit</td>
<td>Gift from Stan Hollenberg</td>
</tr>
<tr>
<td>CHMP1A (immuno)</td>
<td>Rabbit</td>
<td>ProteinTech 15761-1-AP</td>
</tr>
<tr>
<td>Beta actin</td>
<td>Mouse</td>
<td>Abcam AC-15</td>
</tr>
<tr>
<td>Calbindin</td>
<td>Mouse</td>
<td>Abcam AF2E5</td>
</tr>
<tr>
<td>AQP1</td>
<td>Mouse</td>
<td>Santa Cruz SC25287</td>
</tr>
<tr>
<td>CTIP2</td>
<td>Rat</td>
<td>Abcam AB18465</td>
</tr>
<tr>
<td>SOX2</td>
<td>Goat</td>
<td>Santa Cruz SC17330</td>
</tr>
<tr>
<td>CD63</td>
<td>Mouse</td>
<td>Pelicluster M1544</td>
</tr>
<tr>
<td>CUX1</td>
<td>Rabbit</td>
<td>Santa Cruz SC13024 M-222</td>
</tr>
<tr>
<td>phosho Histone H3</td>
<td>Rat</td>
<td>Sigma H9908</td>
</tr>
<tr>
<td>TSG101</td>
<td>Mouse</td>
<td>Thermo MA1-23296</td>
</tr>
<tr>
<td>SHH (blot)</td>
<td>Rabbit</td>
<td>Abcam AB73958</td>
</tr>
<tr>
<td>SHH (immunoEM)</td>
<td>Mouse</td>
<td>Developmental Studies Hybridoma Bank 5E1</td>
</tr>
<tr>
<td>Cleaved caspase 3</td>
<td>Rabbit</td>
<td>Abcam AB13847</td>
</tr>
<tr>
<td>Beta catenin</td>
<td>Mouse</td>
<td>BD 610153</td>
</tr>
<tr>
<td>Atypical PKC</td>
<td>Rabbit</td>
<td>Santa Cruz SC216</td>
</tr>
<tr>
<td>Aurora A</td>
<td>Mouse</td>
<td>BD 611082</td>
</tr>
<tr>
<td>EGFR</td>
<td>Mouse</td>
<td>Gift from Genentech 13A9</td>
</tr>
<tr>
<td>PAX6</td>
<td>Rabbit</td>
<td>Covance PRB-278P-100</td>
</tr>
</tbody>
</table>
Chapter 3: Somatic mutation in disorders of DNA damage repair
Publication


Attribution

Michael Coulter wrote the dissertation and planned, conducted, and analyzed all experiments except as noted below. Neuronal nuclei from brains 1465, 4638, and 4643 were isolated and whole genome sequenced by Gilad Evrony, Xuyu Cai, Michael Lodato, and Mollie Woodworth. Semin Lee called somatic SNVs in these neurons. Daniel Kwon and Alison Barton called somatic SNVs in Cockayne syndrome and xeroderma pigmentosum brains, calculated MAPD scores, and generated mutational signatures. Amir Karger wrote the script for calculating allele and locus dropout. Christopher Walsh and Peter Park oversaw the project.
3.1 Summary

Human disorders of proteins in the nucleotide excision repair (NER) pathway, including Cockayne syndrome (CS) and xeroderma pigmentosum (XP), cause increased rates of DNA mutation, and are characterized by premature aging and early neurodegeneration. We sequenced single cortical neurons from neurotypical individuals and from patients with CS and XP and somatic single nucleotide variant (SNV) calling revealed that CS neurons have about 4 times more somatic SNVs per cell than neurons from normal individuals. The alternate allele frequency (AAF) of somatic SNVs from CS neurons is most commonly $\approx 0.25$, rather than the expected 0.50 for a fixed, heterozygous mutation, suggesting that single-stranded, non-repaired DNA lesions may make up a majority of detected SNVs in these neurons, though SNVs with higher and lower AAF are all increased. Consistent with previous work on CS, protection of the transcribed strand by transcription-coupled repair is lost in CS so that, in contrast to controls, the transcribed and non-transcribed strands have indistinguishable somatic SNV rates. Somatic SNVs from CS neurons have a distinct mutational signature from normal neurons, consistent with a defect in nucleotide excision repair. Finally, CS neurons contain several truncating mutations in essential genes, suggesting that somatic mutations may impair neuronal function. Together, our findings suggest a model for neurodegeneration: increased DNA damage and somatic mutations impair genes important for neuron function and survival, causing cell death, atrophy, and degeneration.
3.2 Introduction

Somatic mutations, mutations that are only present in some cells of an individual, are ubiquitous in the human brain. Recent work has demonstrated that adult cortical neurons from neurologically normal individuals each carry approximately 1,500 somatic single nucleotide variants (SNVs) (Lodato et al., 2015) not found in other cells or tissues. Several recent reports have also established that somatic mutation of disease genes in the brain can cause neurodevelopmental disorders ranging from cortical malformations to autism when they occur in progenitor cells and are shared by many brain cells (D’Gama et al., 2015a; Jansen et al., 2015; Lee et al., 2012; Poduri et al., 2012; Rivière et al., 2012 Jamuar et al., 2014; D’Gama et al., 2015b). Further study of somatic mutation in the human brain will help reveal the role of somatic mutation in normal brain function and disease and the mechanisms driving these mutations.

Recent collaborative work from our lab and Peter Park’s lab measured the rate of somatic SNVs in single neurons from post-mortem human cortex of neurotypical individuals (Lodato et al., 2015). By isolating single neuronal nuclei and performing multiple displacement amplification (MDA) followed by whole genome sequencing (WGS), Lodato et al. showed that an adult cortical neuron contains an average of 1458 – 1580 somatic SNVs called in common by three different somatic mutation calling algorithms. They found that over 99% of these variants are unique to each neuron and that these variants likely arise from DNA damage during transcription. On the other hand, two other groups have also recently published estimates of somatic SNV rate in single mouse cells grown up into large clones and found substantially different results.
Baldwin and colleagues used somatic cell nuclear transfer of the nuclei of mouse mitral and tufted cells to create a new cloned mouse, with the genome of that mouse representing an amplified version of the genome of the original single neuron; they found an average of 86 somatic SNVs per neuron (Hazen et al., 2016). Bradley and colleagues reprogrammed human endothelial precursor cells (EPC) into induced pluripotent stem cells and used WGS following stem cell culture to identify 391 somatic SNVs per EPC (Rouhani et al., 2016). These two groups used cell culture to amplify single cell DNA prior to WGS, whereas our lab used MDA. There are likely several reasons for the higher rates of called SNVs in single human neurons compared to cells whose genome has been amplified by cell division, including both technical and species differences. However, single cell genome amplification is an artifact-prone procedure and could also potentially contribute to these differences. If single cell MDA does indeed accurately capture rates of somatic mutation however, it makes a strong prediction that disorders of DNA damage repair should show higher rates of somatic mutation than normal cells. In order to further test this hypothesis, we sequenced single neurons from two human disorders of DNA damage repair, diseases with evidence of increased somatic mutation. Not only does this experiment allow us to test the biological plausibility of somatic mutation rates in MDA amplified DNA, it also provides a potentially revolutionary ability to track the number and distribution of somatic mutation in various diseases to elucidate mechanisms of DNA damage and somatic mutation. Successfully demonstrating the ability of MDA/WGS of single cells to capture rates and types of mutations potentially provides a powerful new tool to study diverse neurodegenerative disorders to understand the mechanisms of DNA damage.
Cockayne syndrome (CS) and xeroderma pigmentosum (XP) are inherited disorders caused by loss of function mutations in genes that encode for components of the nucleotide excision repair (NER) pathway. CS and XP patients have several neurologic deficits including microcephaly, neurodegeneration, and hearing and vision problems. Brain MRI shows cortical and cerebellar atrophy over the course of 8 years in a young CS patient (Figure 3.1). Constitutional symptoms differ between the two disorders; CS patients have growth failure and death in the first or second decade of life, while XP patients have greatly increased risk of skin cancer and live to the third to fifth decade (Nickoloff and Hoekstra, 1998). These disorders also show evidence of progeria, i.e., premature aging, suggesting that better understanding the mechanisms and consequences of DNA mutation in these conditions might enlighten our understanding of the aging process in general.
Figure 3.1: Brain MRI of a Cockayne syndrome patient

(A) and (B): coronal and horizontal images at 4 years of age. (C) and (D): coronal and horizontal images at 13 years of age. Cortex and cerebellum showed marked atrophy over the course of 8 years (Laugel et al., 2010).

The molecular mechanisms that underlie CS and XP have been studied since the disease genes were discovered 20 years ago. There is consensus that accumulated mutations in tumor suppressor and oncogenes underlie the huge increased risk of skin cancer in XP (Cleaver et al., 2009), but there is no consensus on how defective NER causes neurodegeneration. There is evidence for several different cellular dysfunctions in
CS, but which, if any of these, is most relevant for explaining neurodegeneration is unknown.

The primary molecular defect in CS and XP is loss of NER function. NER removes bulky DNA lesions that distort the DNA helix including pyrimidine dimers caused by UV irradiation (Hoejimakers, 2001), cyclo-adenosine and cyclo-guanosine residues created by oxidative damage (Brooks 2008), and bulky DNA adducts caused by DNA modification by polycyclic aromatic hydrocarbons (Hoejimakers, 2001). In the absence of NER, the cell cannot remove these lesions and DNA damage accumulates. Early experiments showed that CS Chinese hamster ovary (CHO) cells have substantial impairment in removing DNA damage (pyrimidine dimers) from the transcribed strand following UV treatment (Vreeswijk et al., 1998). In 24 hours, control cells removed 90% of pyrimidine dimers, while CS cells removed only 35%. Accumulation of DNA damage can cause two molecular defects that may underlie neurodegeneration in CS and XP: increased somatic mutation rate and decreased gene transcription.

Rates of somatic mutation are increased in CS and XP cell lines. Csb loss of function increases in vitro mutation rate following UV irradiation. Following UV exposure, somatic mutation rate at the Hprt locus was 3 times higher in Csb null CHO cells compared to controls (Vreeswijk et al., 1998). Somatic mutation rate was also increased in fibroblasts from mouse models of CS and XP exposed to DMBA, a polycyclic aromatic hydrocarbon. Somatic mutation rate was 3 – 7 times higher at the Aprt locus in CS and XP cells compared to controls (Wijnhoven et al., 2001). These results show that defective NER increases somatic mutation rate in CS and XP, but they do not profile the landscape and mechanism of these somatic mutations. Increased
somatic mutation rate in the brain could cause neuron dysfunction and cell death leading to degeneration as mutations accumulate in genes required for neuron function and survival, but the postmitotic state of CNS neurons, and their diverse clonal origins, has heretofore prevented direct study of rates and patterns of mutations in CS and XP neurons.

Endogenous increased somatic mutation rate has been detected in mouse models of CS and XP. Hprt sequencing in T cells from Csb null and Xpa null mice did not detect an increase in endogenous mutation rate. However, when apoptosis was inhibited, by creating a Xpa and p53 double knockout mouse, the somatic mutation rate at the Hprt locus was 5 times higher than in controls (Wijnhoven et al., 2000). This finding provides evidence that disrupting the common NER pathway does increase somatic mutation frequency but also increases cell death, which can mask this effect. In fact, T cells may not be the best cell type for detecting somatic mutations in CS and XP. T cells are constantly proliferating, so are more likely to activate apoptosis, and have a very short half-life of about 90 days (Hellerstein et al., 1999), whereas neurons are post-mitotic from mid-gestation development (when cortical neurogenesis ends) to the end of an individual’s life. Thus, one possibility is that neurons may accumulate somatic mutations that would trigger apoptosis in proliferating cells and so provide a better cell type for measuring somatic mutation rate, though this hypothesis has not been tested.

In addition to creating somatic mutations, excess DNA damage in CS impairs gene transcription. In transcription coupled NER (TCR), bulky DNA lesions in transcribed regions stall RNA polymerase and stop transcription, which recruits CSA and CSB to sites of DNA damage for repair (Cleaver et al., 2009). Consistent with this
model, CS cell lines lacking TCR have reduced transcription recovery following UV exposure (Friedberg, 1996; Mayne and Lehmann, 1982; Vreeswijk et al., 1998). In addition to increased somatic mutation rate, globally reduced gene expression as a result of excess RNA polymerase stalling in CS could cause neurodegeneration because of decreased expression of neuron function and survival genes. Beyond defective TCR, additional molecular defects have been found in CS that could contribute to disease pathogenesis including mitochondrial dysfunction from mutations in mitochondrial DNA (Kamenish et al., 2010), and reduced telomere length and stability (Batenburg et al., 2012). Our incomplete knowledge of the magnitude and time-course of these various mutational processes prevents us from gaining a systematic understanding of the disease processes that drive the various phenotypic features of the disease, and a clearer understanding of these pathogenic mechanisms would ultimately guide treating or preventing disease manifestations.

As CS and XP are primarily disorders of NER, we think single neuron WGS is an excellent method to investigate neurodegeneration in these disorders. Single cell sequencing potentially allows us to accurately measure somatic mutation rate genome-wide, characterize the type of SNVs present in neurons with CS and XP, and identify genes that are disrupted by these somatic mutations. Two advantages of this approach are that we will directly identify endogenous somatic mutations and that we will study these mutations in neurons, the most appropriate cell type for understanding neurodegeneration.

Since previous work has suggested that somatic mutation rate is increased in CS and XP (Vreeswijk et al., 1998), measuring somatic SNV rate in single CS and XP
neurons by MDA genome amplification and whole genome sequencing has the additional utility that it helps develop the technology and calling algorithms for utilizing this sequencing technology in a wider range of disease processes. Studying a condition where there is strong prior evidence for increased somatic mutation produces a strong hypothesis that we expect to see increased rates of somatic mutation at the single cell level; hence, studying CS and XP will test the accuracy of this method for single cell somatic mutation identification. Testing our protocol to detect an expected biologic difference in mutation rate is important because our estimate for somatic SNV rate in normal neurons is much higher than studies published using cell culture for genome amplification rather than MDA.
3.3 Results

**Post mortem cortex from patients with Cockayne syndrome and XP**

We queried the NIH NeuroBioBank, a repository of donated human post-mortem brains, for Cockayne syndrome and XP. We found a total of 3 brains for XP and selected 6 brains with Cockayne syndrome that showed the shortest post-mortem interval, suggesting that they were likely to be the best preserved. For all experiments we used fresh frozen tissue from the prefrontal cortex (or general cortex for brains in which further location information was not available). As controls, we used fresh frozen tissue from three brains donated to the NeuroBioBank from neurotypical individuals. **Tables 3.1 – 3.3** show donor information for each sample.
Table 3.1: Post-mortem brains from patients with XP used in this study

<table>
<thead>
<tr>
<th>ID</th>
<th>Age</th>
<th>Gender</th>
<th>Clinical features</th>
</tr>
</thead>
<tbody>
<tr>
<td>UMB5316</td>
<td>45</td>
<td>F</td>
<td>Not available</td>
</tr>
<tr>
<td>UMB5379</td>
<td>24</td>
<td>F</td>
<td>Developmental delay, ataxia, brain atrophy, facial squamous cell carcinoma</td>
</tr>
<tr>
<td>UMB5416</td>
<td>46</td>
<td>F</td>
<td>Learning and socialization difficulty, ataxia, deafness, skin tumors</td>
</tr>
</tbody>
</table>

Table 3.2: Post-mortem brains from patients with CS used in this study

<table>
<thead>
<tr>
<th>ID</th>
<th>Age</th>
<th>Gender</th>
<th>Clinical features</th>
</tr>
</thead>
<tbody>
<tr>
<td>UMB580</td>
<td>8</td>
<td>F</td>
<td>Microcephaly</td>
</tr>
<tr>
<td>UMB682</td>
<td>33</td>
<td>M</td>
<td>Not available</td>
</tr>
<tr>
<td>UMB1124</td>
<td>5</td>
<td>F</td>
<td>Seizures, deafness, cataracts</td>
</tr>
<tr>
<td>UMB1286</td>
<td>6</td>
<td>M</td>
<td>Microcephaly, small somatic size, seizures, strokes, pontine nuclei mineralization</td>
</tr>
<tr>
<td>UMB1762</td>
<td>4</td>
<td>F</td>
<td>Microcephaly, cerebellar atrophy, strokes, decreased myelination</td>
</tr>
<tr>
<td>UMB5105</td>
<td>9</td>
<td>M</td>
<td>Developmental delay, seizures, deafness, cataracts, brain atrophy, mineralization, decreased myelination</td>
</tr>
</tbody>
</table>

Table 3.3 Normal control brains used in this study

<table>
<thead>
<tr>
<th>ID</th>
<th>Age</th>
<th>Gender</th>
<th>Clinical features</th>
</tr>
</thead>
<tbody>
<tr>
<td>UMB1465</td>
<td>17</td>
<td>M</td>
<td>Neurologically normal, died in motor vehicle accident (MVA)</td>
</tr>
<tr>
<td>UMB4638</td>
<td>15</td>
<td>F</td>
<td>Neurologically normal, died in MVA</td>
</tr>
<tr>
<td>UMB4643</td>
<td>42</td>
<td>F</td>
<td>Neurologically normal, died of possible stroke</td>
</tr>
</tbody>
</table>
Fluorescence activated isolation of neuronal nuclei

We isolated nuclei from each sample using tissue disassociation and ultracentrifugation. We stained nuclei for immunoreactivity to NEUN antiserum, a nuclear marker of post-mitotic neurons, and then used fluorescence assisted nuclear sorting (FANS) to deposit one neuronal nucleus in each well of a 96 well plate. We then performed whole genome amplification using multiple displacement amplification (MDA) on each nucleus to generate several micrograms of whole genome-amplified DNA (Figure 3.2).

Figure 3.2 Single neuron somatic SNV identification protocol

Bulk nuclei are isolated from frozen cortex and neuronal nuclei are purified with FANS, DNA is amplified with MDA and then sequenced (adapted from Lodato et al., 2015).
WGS coverage statistics

We isolated and MDA-amplified DNA from 150 nuclei from each CS and XP cortex sample and then selected 4 – 6 nuclei with the most complete coverage for WGS, as judged by PCR amplification of 16 – 24 microsatellite loci distributed across the genome. 10 – 16 cortical neuron nuclei selected and sequenced from three normal brains as part of a previously published study were used as control cells (Lodato et al., 2015). After selecting cells with the most complete and uniform amplification from each brain we performed WGS (40-45x target coverage). In total, we sequenced 42 disease neurons and 36 control neurons (Table 3.4). Normal neurons have excellent WGS coverage with individual brain averages ranging from 84 – 89 percent of the genome covered with more than 10 reads (>10x, Table 3.4, Figure 3.3). The coverage for the CS and XP brains was considerably more variable, ranging from 42 – 76 percent covered at 10x for CS brains and 34 – 56 percent covered at 10x for XP brains (Table 3.4, Figure 3.3). The reasons for consistently less complete amplification in these two diseases are not known. It may reflect many technical variables, since with the disease brains we have a narrower range of postmortem intervals and quality characteristics to choose from. However, it may also reflect effects of DNA damage on the amplification process itself.
Table 3.4 Genome-wide coverage of single neurons analyzed in the study

<table>
<thead>
<tr>
<th>Normal</th>
<th>Brain ID</th>
<th>1465</th>
<th>4638</th>
<th>4643</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of cells</td>
<td></td>
<td>16</td>
<td>10</td>
<td>10</td>
<td>36</td>
</tr>
<tr>
<td>Covered at 10x (%)</td>
<td></td>
<td>83</td>
<td>84</td>
<td>89</td>
<td>85    (0.67)</td>
</tr>
<tr>
<td>Covered at 20x (%)</td>
<td></td>
<td>69</td>
<td>72</td>
<td>78</td>
<td>73    (0.92)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cockayne syndrome</th>
<th>Brain ID</th>
<th>580</th>
<th>682</th>
<th>1124</th>
<th>1286</th>
<th>1762</th>
<th>5105</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of cells</td>
<td></td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>27</td>
</tr>
<tr>
<td>Covered at 10x (%)</td>
<td></td>
<td>63</td>
<td>76</td>
<td>42</td>
<td>57</td>
<td>74</td>
<td>69</td>
<td>66   (2.14)</td>
</tr>
<tr>
<td>Covered at 20x (%)</td>
<td></td>
<td>51</td>
<td>63</td>
<td>30</td>
<td>45</td>
<td>59</td>
<td>56</td>
<td>52   (2.02)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Xeroderma pigmentosum</th>
<th>Brain ID</th>
<th>5316</th>
<th>5379</th>
<th>5416</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of cells</td>
<td></td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>Covered at 10x (%)</td>
<td></td>
<td>34</td>
<td>55</td>
<td>56</td>
<td>51    (2.65)</td>
</tr>
<tr>
<td>Covered at 20x (%)</td>
<td></td>
<td>26</td>
<td>43</td>
<td>44</td>
<td>40    (2.26)</td>
</tr>
</tbody>
</table>
Figure 3.3 WGS coverage of brains used in this study

(A) and (B): percent of genome covered with indicated number of reads for each brain.
Quality control assessment of WGS nuclei

Although WGS of neurons from neurotypical brains showed excellent genome-wide coverage, coverage varied among the CS and XP brains, and so we chose brains with coverage statistics close to the control brains for comparison and further analysis. To select the highest-quality CS and XP neurons, we first measured locus dropout (LD) by identifying loci that were heterozygous and well covered in bulk brain but that had no reads in WGS of a single neuron. Control brains had 2 – 5% locus dropouts per cell while CS and XP brains ranged from 4 – 33% (Figure 3.4). CS brains 682, 1762, and 5105 had much lower LD than the other disease brains and the average LD in these brains was not dramatically higher than control brains (Control average = 3.5%, 682, 1762, and 5105 average = 5.1%). Once again, the reasons for higher LD might reflect both technical issues as well as biological issues, so that as we become more comfortable with analysis of the higher-quality cells and disease brains, we plan to return to the analysis of these lower quality cells in the future.

We examined several other quality control metrics across all brains and found that brains 682, 1762, and 5105 consistently showed more complete genome coverage than other disease brains (Figure 3.4). First, genome wide coverage at 10x ranged from 84 – 89% for control neurons, while neurons from 682, 1762, and 5105 had average genome wide coverage at 10x ranging from 69 – 76%, while neurons from the other six disease brains ranged from 34 – 63%. MAPD score (median absolute pairwise deviation) is a measurement of evenness of coverage in which values close to 0 indicate more even coverage (Cai et al., 2014). Average MAPD of control neurons ranged from 0.5 – 0.6, average MAPD of 682, 1762, and 5105 ranged from 0.98 – 1.13, and average MAPD of
the remaining six disease brains ranged from 1.22 – 1.94. These three measures of completeness and evenness of coverage all show that neurons from CS brains 682, 1762, and 5105 had the most complete and even coverage of the nine CS and XP brains and were closest to the high coverage of control neurons.

One additional measurement for even coverage is allele dropout (AD). AD is present when one allele of a heterozygous variant from bulk DNA is absent from the reads of a single neuron, while the other allele is captured. This measurement shows a large difference between neurons from the three control brains (5 – 13%) and neurons from all CS and XP brains (27 – 35%). The causes for this discrepancy are currently under examination, and once again, include potential technical as well as biological explanations. MDA is driven by phi29 polymerase DNA synthesis. This polymerase is highly processive and generates fragments of DNA that are several kilobases in length. One possible explanation is that increased DNA damage in CS and XP creates more DNA lesions, which cause early termination of phi29. When this occurs at the location of a germline SNP, the allele on the lesioned strand will be dropped and this would increase the chance of observing AD. All CS and XP brains have similar AD rate, so we did not use this feature to select high quality brains for comparison to normal.

Based on all quality control measurements, comparison between CS and normal in the following experiments uses CS brains 682, 1762, and 5105, and normal brains 1465, 4638, and 4643.
Figure 3.4: Quality control measurements of neurons used in this study

Gray indicates neurotypical control brains, orange indicates CS, and purple indicates XP.
**Somatic SNVs identified by single neuron WGS**

To determine the role of CS mutations on somatic mutation rate in the brain, we measured the rate of somatic SNVs in control and CS neurons. We identified somatic SNVs based on a protocol previously developed in collaboration between the Walsh lab and Peter Park’s lab at HMS (Lodato et al., 2015). Several algorithms are available to identify SNVs in WGS data. We use three independent algorithms: GATK, MuTect and VarScan, to detect variants that are present in WGS from a single neuron and absent from WGS of bulk tissue from the same individual. We intersect these three lists of somatic variants, and within “triple-called” variants, keep only variants located at bases that have a called genotype (not dropped out) in all neurons of a particular brain. Previous analysis of somatic SNVs identified in this way from control brains 1465, 4638, and 4643 showed that about 90% of “triple called” SNVs validate by Sanger resequencing of the single neuron’s amplified DNA (Lodato et al., 2015). A circos plot shows that triple called SNVs from 4 example CS neurons are evenly distributed throughout the genome with no large missing regions (**Figure 3.5**). All 36 normal neurons had very similar number of somatic SNVs. All but one CS neuron also had somatic SNV rates similar to each other (**Figure 3.6**). There was one outlier neuron from brain 5105 that had about 6 times as many variants as all other CS neurons (Grubbs’ outlier test, p < 0.01). Because this cell is such an extreme outlier we excluded it from all analyses; since this very high rate is not obviously due to a technical explanation, it might reflect something about the mechanism of disease we do not yet understand. CS patients exhibit cerebral atrophy, so a plausible explanation for the massive number of somatic SNVs in this cell is that it was apoptotic or necrotic at the time of collection and its DNA started to fragment.
Figure 3.5: Circos plot of illustrative Cockayne syndrome single neurons

Circos plot of triple-called SNVs from example neurons from brains 1762 and 5105 show that somatic variants in CS and XP were evenly spread across the genome at a large scale and that genome amplification did not miss any very large regions (e.g. chromosome arm).
Dramatic increase in somatic SNVs in CS neurons

We compared the number of triple-called SNVs per genome in CS and normal neurons and discovered that CS neurons have about 4 times as many somatic SNVs per genome as normal neurons (Figure 3.6). The three control brains have a range of 1684 – 1794 triple-called SNVs per genome and cells from the three CS brains show a range of 6115 – 7869 (Control = 36 cells, CS = 15 cells, two-tailed t-test, p = 1.82 x 10^{-30}) triple-called SNVs using the same algorithm. Including the estimated 23% false discovery rate previously reported for the control neurons (Lodato et al., 2015), control brains range from 1458 – 1580 SNVs per neurons and CS brains range from 4709 – 6059. This finding supports our hypothesis that impaired DNA damage repair in CS increases somatic mutation rate in the cerebral cortex and suggests that increased somatic mutation may underlie cortical degeneration in CS patients. The large increase in triple-called somatic SNVs in CS neurons also provides evidence that our protocol for somatic mutation analysis identifies bona fide somatic variants rather than simply identifying technical artifacts created by MDA genome amplification.
Figure 3.6 Somatic mutation rates in CS and control neurons

(A) Number of triple-called somatic SNVs for individual neurons from each CS brain.

(B) Average somatic SNVs per neurons for control and CS brains. Pooled data illustrates 4-fold increase in SNVs per neuron in CS.
**Alternate allele fraction and DNA lesions**

To better understand the increased rate of somatic SNVs in CS neurons, we measured the alternate allele frequency (AAF) of somatic SNVs in CS and normal neurons. The AAF represents the proportion of DNA sequence reads in which the somatic mutation is captured. For a somatic, heterozygous mutation, the expected AAF is 50% assuming perfect sequencing technology, whereas artifacts arising during amplification and/or sequencing might be expected to have very low AAFs. In our previous work on neurons from neurotypical individuals, we had found that somatic SNVs shared between two neurons showed average AAF of about 50%, though with a wide range from about 25-80% (Lodato et al., unpublished). On the other hand, many somatic triple-called SNVs had AAF around 25%, for reasons that are not yet clear. We also modeled “false-positive” SNV, which are more likely to be amplification artifacts, and found that these most commonly showed AAF <25% (Lodato et al., 2015).

CS neurons showed increased SNV rate throughout the spectrum of AAF compared to normal control neurons (Figure 3.7). Somatic SNVs with AAF ≥50% were increased 2.6 fold (Control average = 140, CS average = 368), suggesting a substantial increase in fixed, double-stranded mutations caused by faulty repair of a DNA lesion to the incorrect base. On the other hand, when comparing the overall AAF in normal and CS neurons, we found important differences in average AAF and AAF distribution. Average somatic SNV AAF = 41% in control neurons and 36% in CS neurons (two-tailed t-test, p = 1.74 x 10⁻¹⁴). In addition, there is a striking change in shape of the AAF curve between control and CS (Figure 3.7). In particular, CS neurons have a 6 fold increase in SNVs at AAF = 0.25 compared to controls (Control average = 296, CS
average = 1,780). We quantified this difference by calculating the ratio AAF = 0.25 to AAF = 0.5. The ratio of 0.25 to 0.5 was 2.1:1 in control neurons and more than doubled to 4.8:1 in CS neurons (two-tailed t-test, p = 0.01).

Although we are not yet certain of the reasons for the large peak of 25% SNV in normal neurons, and the further increase in the size of this peak in CS, the most parsimonious explanation is that these 25% AAF SNVs represent predominantly single-stranded DNA lesions, and that these “unrepaired” single stranded lesions are particularly increased in the presence of faulty NER. Post-mitotic neurons are diploid and hence have four DNA strands (two sets of complementary strands) at each locus. Heterozygous or “fixed” somatic SNVs will be present on two strands of DNA and have AAF centered at 0.5, and this is well reflected in the average 50% AAF of shared somatic SNVs, suggesting that the sequencing technology accurately captures AAF. In contrast, DNA lesions that are unrepaired, or destined to be repaired, are present on one strand of DNA and hence when amplified during MDA would be expected to show an AAF centered at 0.25. During in vitro MDA, it is likely that many DNA lesions will be converted into SNVs because the cellular machinery for DNA damage repair is absent. Thus, we interpret the increased ratio of AAF = 0.25 to AAF = 0.5 in CS neurons as an indication of increased DNA lesions. This finding is consistent with several previous studies that have found increased DNA damage in cells and tissue with CS (Vreeswijk et al., 1998), which provides evidence that the differences we see in AAF and by extension somatic SNV rate are a result of biological differences between CS and normal neurons rather than a result of MDA artifacts.
Figure 3.7: Alternate allele frequency of somatic SNVs in CS and control neurons

(A) Alternate allele frequency is displayed for somatic SNVs from each control (gray) and each CS brain (orange). The y-axis of each plot is number of SNVs. (B) Overlay of average AAF of all CS and normal neurons shows more SNVs at 0.5 and 0.25 in CS, with a prominent shift towards 0.25.
Figure 3.7 (Continued)

A

Control - 1465

Cockayne - 682

Control - 4643

Cockayne - 1762

Control - 4638

Cockayne - 5105

B

Control vs CS

CS

Control
Evidence for loss of TCR in CS neurons

To determine whether differences between somatic SNVs in CS and control neurons are consistent with disease pathology, we estimated the rate of SNVs on each strand in regions of transcribed DNA. CS proteins are responsible for transcription-coupled repair (TCR), i.e., repair of DNA lesions on single stranded DNA that is being actively transcribed (Hu et al., 2015). As a result of TCR, there are fewer mutations on the transcribed strand in normal cells compared to the non-transcribed strand (Hu et al., 2015). Without TCR in CS, this preferential repair is abolished and the expectation is that the repair rate would become more equal between the transcribed and non-transcribed strands.

Whereas previous analysis of neurons from control brains showed that somatic SNVs are significantly depleted on the transcribed strand, consistent with intact TCR (Lodato et al., 2015; Figure 3.8), in CS neurons we found no detectable difference between SNV rate on the transcribed and non-transcribed strands in transcribed regions (Figure 3.8). Two brains (1762 and 5105) showed a trend towards increased SNV rate on the non-template strand, a finding that has been previously reported in CS cells (Hu et al., 2015). The loss of strand bias, suggesting impairment in TCR in neurons from patients with CS, is consistent with what is previously known about the disease and provides further evidence that the differences we observe between somatic SNV rates in CS neurons and control neurons are likely to represent biological phenomena. Together, our findings of overall increased somatic SNV rate, an especially large increase in SNVs with AAF near 0.25, and the loss of SNV strand bias in WGS of CS single neurons all support the hypothesis that single cell MDA-WGS can identify signatures of mutation.
that differ in a biologically plausible fashion, and support the validity of our method to identify somatic SNVs in post-mortem human neurons.

Figure 3.8: Strand bias of somatic SNVs from transcribed genomic regions

(A) A>G and G>A transitions are depleted from the transcribed (-) strand in control neurons reflecting intact TCR. (B) A>G and G>A transitions rates are indistinguishable between transcribed (-) and non-transcribed (+) strand showing loss of TCR.

SNV substitution type and mutational signature

To learn more about the pathogenesis of CS, we examined the type of base substitutions in CS SNVs and the mutational signature of CS SNVs, and we found that the frequency of substitution types and the mutational signature are significantly different between CS and control neurons. SNVs in both control and CS neurons are predominantly C>T transitions, presumably reflecting the particular susceptibility of cytosine to deamination to uracil, which is amplified to thymidine. On the other hand, CS neurons also have a large increase in the frequency of T>C transitions compared to controls (Figure 3.9). This difference is consistent with previous work showing that
oxidative DNA damage repaired by NER primarily causes C>T and T>C transitions (Xu et al., 2015).

Figure 3.9: Somatic SNV substitution types in CS and control neurons

Distribution of substitution types for somatic SNVs in control and CS neurons. CS neurons have increased T>C transitions and decreased C>T transitions compared to controls.

“Mutational signature” analysis (Alexandrov et al., 2013) showed a more striking difference between control and CS SNVs (Figure 3.10). Mutational signature describes the immediate DNA context surrounding a SNV by measuring the frequency of each three-base combination around a mutated base; different types of cancer and different mutagens (such as UV light) produce distinct and reproducible mutational signatures (Alexandrov et al., 2013). Mutational signature analysis replicated our observation that CS somatic SNVs have more T>C substitutions than controls (Figures 3.10A, B). The
most striking difference, however, is the appearance of specific enrichment for three signatures within C>T transitions: GCA, GCC, and GCT (Figure 3.10C). Indeed, if the signature of C>T transitions found in control neurons is subtracted from CS neurons, the remaining difference shows a strong enrichment for GCN signatures (Figure 3.10D). This comparison of signatures between CS and normal suggests that CS neurons contain the kinds of DNA lesions and mutations that characterize normal neurons, but that, superimposed on top of this, they have additional SNVs with different sequence characteristics that presumably reflect defective nucleotide excision repair.

In order to try to determine the mechanism that might account for the specifically increased mutational signatures seen in CS, we created a “net” CS mutational signature by subtracting the normal SNV pattern from the CS pattern, and then compared that CS pattern to other known mutational signatures. The COSMIC database (http://cancer.sanger.ac.uk/cosmic) includes 30 specific mutational signatures identified in different tumors. Visual review of these signatures reveals that the “net” CS SNV signature is quite similar to a signature found in tumors caused by defects in mismatch repair (MMR) (Figure 3.10E). In fact, some previous reports have shown that NER is defective in cells with MMR deficiency suggesting there is a link between these two types of DNA repair (Kobayashi et al., 2005). Substitution type and mutational signature analysis of SNVs from CS neurons reveal further difference between CS and normal neuron SNVs and suggest specific mutagens and other DNA repair processes that drive increased somatic mutation rate in CS neurons.
Figure 3.10: Mutational signatures of SNVs from CS and control neurons

(A) Mutational signature (Alexandrov et al., 2013) of somatic SNVs in an example neuron from normal brain. This signature is dominated by C>T transitions. (B) Mutational signature of somatic SNVs in an example neuron from CS brain. This signature contains more T>C transitions than control and a different pattern of C>T transitions. (C) Average mutational signatures of C>T transitions in normal and CS neurons. (D) “Net” mutational signature of C>T transitions in CS neurons generated by subtracting signature of control neurons. (E) Mutational signature of tumor with MMR deficiency (Alexandrov et al., 2013).
Figure 3.10 (Continued)

A Mutational signature from example normal neuron

B Mutational signature from example CS neuron

C C>T mutational signature

D “Net” CS

E MMR signature
Deleterious somatic mutations in CS neurons

Our data show that cortical neurons from CS patients have about 4 times more somatic SNVs than neurons from normal individuals. We hypothesize that this increase in somatic mutation rate may contribute to neurodegeneration in CS by impairing the function of genes that are essential for neuron function and survival. There were 52 truncating mutations in protein coding genes in the 15 sequenced CS neurons, including several hits in functionally important neuronal genes (Table 3.5). Truncating mutations were present in TNIK, TRAF2 and NCK interacting kinase, a schizophrenia risk gene (Potkin et al., 2009), and CNTN4, contactin 4, a gene deleted in autism (Roohi et al., 2009). Truncations were also found in two genes for recessive inborn errors of metabolism with neurologic defects, SMPD1, acid sphingomyelinsase (Niemann Pick disease, Ledesma et al., 2011) and ARSB, arylsulfatase B (Maroteaux-Lamy syndrome, Debette and Germain, 2014). Other notable genes include CNGB1, cyclic nucleotide gated channel beta 1, mutated in retinitis pigmentosum (Bareil et al., 2001), Myc binding protein 2, Gli family member GLIS3, and neuronal leucine rich repeat 3. CS neurons harbor loss of function mutations essential for neuron function and survival supporting our hypothesis that somatic mutations lead to neuron dysfunction and cell death in CS.
Table 3.5 CS neuron truncating somatic mutations in disease genes

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Protein name</th>
<th>Human disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMPD1</td>
<td>Acid sphingomyelinase</td>
<td>Niemann Pick disease</td>
</tr>
<tr>
<td>TNIK</td>
<td>TRAF2 and NCK interacting kinase</td>
<td>Schizophrenia risk gene</td>
</tr>
<tr>
<td>ARSB</td>
<td>Arylsulfatase B</td>
<td>Maroteaux-Lamy syndrome</td>
</tr>
<tr>
<td>CNGB1</td>
<td>Cyclic nucleotide gated channel beta 1</td>
<td>Retinitis pigmentosum</td>
</tr>
<tr>
<td>MYCBP2</td>
<td>Myc binding protein 2</td>
<td></td>
</tr>
<tr>
<td>CNTN4</td>
<td>Contactin 4</td>
<td>Autism</td>
</tr>
<tr>
<td>GLIS3</td>
<td>GLI 3 family protein</td>
<td></td>
</tr>
<tr>
<td>LRRN3</td>
<td>Leucine rich repeat, neuronal 3</td>
<td></td>
</tr>
</tbody>
</table>

Model for neurodegeneration

In this chapter, we report the first genome-wide measurement of somatic mutation in neurons from a neurodegenerative disease. Together, our results suggest a model for neurodegeneration in CS. In the absence of transcription-coupled repair, DNA lesions accumulate, and some become fixed somatic mutations. Unrepaired DNA damage, present as single strand lesions, may eventually lead to cell death through global transcription failure and increased somatic mutations may lead to cell death through deleterious mutations in genes for neuron function and survival. Neuronal cell death causes brain atrophy and neurodegeneration. This model for DNA damage driven degeneration may be generally applicable to other neurodegenerative diseases and to dementia and cerebral atrophy associated with old age.
3.4 Discussion

Validation for MDA based single neuron WGS

Over the past two years, our lab and several others have published greatly varying estimates for the somatic mutation rate based on identification of single cell somatic SNVs (Behjati et al., 2014; Hazen et al., 2016; Lodato et al., 2015; Rouhani et al., 2016). Our lab and others find similar rates of CNVs and retrotransposon insertions (Evrony et al., 2012; Cai et al., 2014; Hazen et al., 2016); this makes the difference in SNV rate even more interesting. Our observation that somatic mutation rate is increased in CS neurons compared to normal agrees with previous reports that measured mutation rate in cultured CS cells exposed to DNA damaging agents (Vreeswijk et al., 1998; Wijnhoven et al., 2001). Importantly, this result provides evidence that our method for single cell WGS, which amplifies the DNA from a single cell using multiple displacement amplification (MDA), accurately identifies at least some somatic SNVs. In particular, the increased somatic SNV rate in CS neurons and the distinct mutational signature of these SNVs provides evidence that the SNVs we identify, while likely including some artifacts of amplification and sequencing, contain enough specificity that the technology can be used to investigate important mechanistic processes. If our method primarily identified MDA artifacts, we would not observe a difference in somatic SNV rate and mutational signature between CS and normal neurons.

The 4-fold increase in somatic SNV rate in CS neurons compared to normal is central to our argument that MDA genome amplification followed by WGS is an accurate measure of somatic mutations in neurons. We are currently working on several follow up
experiments to support the validity of this result, including reducing coverage differences between CS and normal neurons and using alternative genome amplification methods.

In some brains, the number of triple called SNVs is correlated with the MAPD (median absolute pairwise deviation) score (Cai et al., 2014), a measurement of coverage and evenness of amplification following MDA. It is possible that because the average MAPD score of the CS brains is 1 and the average MAPD score of the control brains is 0.6, part of the increase in number of SNVs is a result of uneven amplification in CS neurons as reflected by a higher MAPD score. In order to address this possibility, we are using low coverage WGS (~0.1X) to calculate the MAPD score for about 100 additional isolated nuclei from each brain used in this study. We will then perform high coverage WGS and somatic SNV calling on cells with a range of MAPD scores from each brain to determine the relationship between SNV rate and MAPD score. We can then use this rate to correct the measured SNV rate for each sequenced neuron and then directly compare the corrected number of SNVs in normal and CS neurons.

MDA is one of several methods for whole genome amplification. We chose it because it generates the most complete coverage for single cell DNA input. A different technology called GenomePlex (Sigma-Aldrich) amplifies less of the genome from a single cell but has more even coverage of amplified regions. GenomePlex uses a different polymerase and so will produce different amplification artifacts compared to MDA. We will amplify single neuronal nuclei from CS and normal brains with GenomePlex and then perform WGS and somatic SNV calling. If we find a similar sized increase in somatic SNV rate as we found with MDA, that will suggest that our current method is detecting real somatic variants, not MDA artifacts.
Somatic mutation rate varies by species and cell type

Increased somatic SNVs in CS neurons supports the accuracy of our method for somatic mutation identification and thus also supports the accuracy of our original estimate of about 1500 SNVs per adult cortical neuron (Lodato et al., 2015). 1500 SNVs per neuron is substantially higher than 86 SNVs per mouse mitral and tufted neuron reported by Kristin Baldwin’s lab (Hazen et al., 2016), 391 SNVs per human endothelial progenitor cell reported by Allen Bradley’s lab (Rouhani et al., 2016) and 601 SNVs per mouse intestinal stem cell reported by Michael Stratton’s lab (Behjati et al., 2014). Though some of these differences probably reflect species differences, and while technical artifacts of amplification may increase the rates in single cell studies somewhat, the difference between ours and other groups’ estimated mutation rate likely reflects, at least in part, biological differences between these experiments. Mouse mitral and tufted neurons and intestinal stem cells may have few somatic SNVs because DNA damage repair in mice is more efficient than in humans. Indeed, mouse models of CS and XP require loss of two genes in the NER pathway to faithfully model the human disorder caused by loss of a single gene in the pathway (Cleaver et al., 2009), suggesting species differences in NER. In humans, cortical neurons may contain many more SNVs per cell than endothelial progenitors because neurons are very metabolically active and so may experience more DNA damage from reactive oxygen species than endothelial progenitors.
Advantage of MDA: detection of DNA lesions

In addition to biological differences, our estimate for somatic SNVs rate is much higher than Baldwin and Bradley’s because it likely includes variants generated from single-stranded DNA lesions. It is estimated that a typical cell will suffer up to $10^6$ single-stranded DNA lesions per day, >99% of which are destined to be successfully repaired without creating a permanent, fixed mutation (Madabhushi et al., 2014). When DNA damage occurs, a single stranded lesion, such as a cyclo-purine or pyrimidine dimer, is created. Genome amplification via proliferation of cultured cells would not detect DNA lesions because DNA damage repair pathways are intact and the cell repairs the lesion, and because a clone of cells, even though they share clonal somatic mutations, would not be expected to share single stranded DNA lesions. WGS of cultured cells will only detect fixed somatic SNVs that are present on both strands of DNA. In contrast, during genome amplification by MDA after cell lysis, there is no repair machinery and so single stranded lesions are likely to be amplified by phi29 polymerase and generate a somatic SNV. Thus, WGS following MDA is expected on theoretical grounds to detect both DNA lesions and fixed somatic SNVs. These two types of variants appear to be distinguishable to a certain extent because a DNA lesion that is present on 1 DNA strand will generate an SNV with alternate allele frequency (AAF) equal to 0.25, while a fixed somatic SNV present on 2 DNA strands will generate an SNV with AAF equal to 0.5.

DNA damage is increased in cells from CS patients (Vreeswijk et al., 1998); as such, we expected to detect more DNA lesions in CS neurons compared to normal. This is what we observed. On average, CS neurons have 1,780 DNA lesions (AAF = 0.25) while normal neurons have 296, a 6-fold increase. This increase is larger than for fixed
mutations (2.6-fold increase) consistent with massive accumulation of DNA damage in
the absence of NER. This finding provides evidence that our method can detect DNA
lesions, which provides a more complete characterization of DNA damage and somatic
mutation than cell culture methods than can only detect fixed SNVs.

Although SNVs with AAF = 0.25 may provide a surrogate measurement for DNA
damage, this interpretation will require further supporting evidence, and so we are now
measuring DNA damage chemically to confirm that CS neurons have more damage than
normal. Mass spectrometry of hydrolyzed DNA can identify normal and modified bases
and so can measure DNA damage. To determine if CS neurons have increased DNA
damage we will measure the frequency of two modified bases, cyclo-adenosine and
cyclo-guanosine, that are generated by reactive oxygen species and are detected and
repaired by NER.

**Loss of strand bias in CS supports the accuracy of MDA based WGS**

CS proteins are required for transcription-coupled DNA repair. It has been
demonstrated in many experiments that transcription-coupled repair preferentially repairs
DNA lesions on the transcribed strand (Vreeswijk et al., 1998). This is likely because
stalled RNA polymerase II is a signal to recruit CS and XP proteins for NER (Hu et al.,
2015). A recent study expanded this finding genome-wide by sequencing DNA attached
to the NER complex following UV treatment. Confirming earlier work, they found that
NER preferentially occurs on the transcribed strand and has a modest bias towards
transcribed regions (Hu et al., 2015). In addition, they showed that bias towards repair on
the transcribed strand and bias towards transcribed regions were eliminated in CS
fibroblasts. We measured strand bias in our somatic SNVs and also found depletion of
SNVs on the transcribed strand in control neurons that was absent in CS neurons. Our
ability to detect this mutational difference that is consistent with CS pathology provides
further support for the accuracy of our SNV calling pipeline.

**Mutational signature**

Analysis of the substitution types and mutational signatures of somatic SNVs
from CS neurons highlights possible explanations for endogenous mutations in cortical
neurons and additional DNA repair pathways that interact with NER. 80 – 90% of
somatic SNVs in CS neurons are T>C (A>G) and C>T (G>A) transitions. This pattern
matches the pattern of substitution types caused by a class of oxidative DNA lesions
called cyclo-purines that are repaired by NER (Xu et al., 2015). Free radicals and
reactive oxygen species from cellular respiration and fatty acid metabolism generate
endogenous cyclo-purines (Brooks, 2008). Thus, the somatic SNVs we detected in CS
neurons are consistent with endogenous oxidative damage as the driver of somatic
mutation in CS and perhaps also in normal neurons.

We identified a distinct mutational signature for CS SNVs compared to normal
SNVs, and the CS-specific features of this signature closely match the signature of
somatic mutations in tumors with defective mismatch repair (Alexandrov et al., 2013).
There is evidence that NER and MMR are connected pathways for repair because cells
from some tumors with MMR deficiency also have defective NER (Kobayashi et al.,
2005). One potential explanation for the similarity between CS and MMR mutational
signatures is that both repair processes detect distortions in the DNA helix structure. In
NER, cyclo-purines, pyrimidine dimers, and bulky adducts bend the DNA helix out of shape, while in MMR, mismatched base pairing such as G/T or A/C also bend the DNA helix. Based on our results and previous work (Alexandrov et al., 2013), C>T transitions in CS and MMR occur much more frequency in the context of GCN. Perhaps this motif exaggerates helical distortions and is more easily detected by NER and MMR machinery.

**Increased somatic mutation is a potential contributor to CS neurodegeneration**

Our results support a model for CS neurodegeneration in which DNA damage accumulates in CS neurons, increases the somatic mutation rate, generates deleterious mutations in essential neuronal genes, and results in cell death, atrophy, and degeneration. Previous work measuring endogenous somatic mutation rate in CS only detected an increase in lymphocytes when apoptosis was inhibited by removing P53 (Wijnhoven et al., 2000). We likely detected more somatic SNVs in CS neurons because they are post-mitotic and so can accumulate mutations for a longer time-period compared to lymphocytes. Based on this model, and the general hypothesis that somatic mutations accumulate in post-mitotic cells over the course of an individual’s lifetime, we might expect older CS patients to have more somatic SNVs than younger patients. Although we did not observe this effect, the interpretation of the results are complicated by several factors. The neurons we sequenced were from 4-, 9-, and 33-year-old patients, yet the somatic SNV rates were similar between all individuals and there was no correlation between patient age and SNV rate. One explanation for this is that neurons can only tolerate 5000 – 6000 somatic SNVs before too many deleterious mutations accumulate and lead to cell death. Indeed, we found one outlier neuron with over 40,000 variants
likely representing a cell undergoing apoptosis and DNA fragmentation. Another possibility is that the 33 year-old patient survived to that age (which is unusual for CS patients) because he had a milder version of the disease. Thus our results suggest that increased somatic mutation rate causes neurodegeneration in CS by activating cell death.

**Alternative explanations for CS and XP neurodegeneration**

Although our finding that CS neurons have more somatic SNVs than normal supports the hypothesis that increased somatic mutations are associated with cell death and neurodegeneration in CS, it is possible that increased rate of DNA lesions in CS neurons also contributes to disease through inhibition of transcription. Previous work shows that recovery of transcription following DNA mutagenesis in impaired in CS cells (Vreeswijk et al., 1998). This is consistent with results showing that helix-distorting lesions (detected and repaired by NER) cause RNA polymerase II arrest and inhibit transcription. We can measure transcription arrest in two ways in CS neurons. First, we can perform RNA-sequencing of bulk cortex RNA and determine if CS brains have more reads at the 5’ end of transcripts compared to the 3’ end, which would indicate RNA polymerase stalling. Second, we can try to measure RNA pol II stalling directly with RNA pol II ChIP-seq to measure RNA pol II occupancy on DNA (Le Martelot et al., 2012). If these assays do not work well on frozen brain tissue, we can use genome editing to generate CS iPSCs and differentiate these cells into neurons. If we detect increased transcription arrest in CS brains or differentiated neurons it will suggest that this mechanism contributes to degeneration in addition to accumulation of deleterious somatic mutations.
Further exploration of CS pathology

Our results show that CS neurons have about 4 times as many somatic SNVs as normal neurons. To better understand somatic variation in NER, we are planning experiments to explore additional aspects of CS and to examine XP.

Poor genome coverage in several CS and XP brains

Although we set out to measure somatic mutation in CS and XP neurons we were unable to include the XP brains and several CS brains in our analysis because of low coverage following MDA genome amplification. This is likely a result of accumulated DNA lesions that cause phi29 polymerase to stall and fall off, resulting in large regions of poor amplification. In order to use the WGS data from these cells, we are currently identifying regions of genome that are well covered and then calling somatic SNVs in these regions. We will then be able to calculate a somatic SNV rate by normalizing the number of SNVs to the percent genome coverage in each cell.

Somatic mutation in oligodendrocytes

White matter degeneration is a common finding in CS patients (Cleaver et al., 2009) and oligodendrocytes produce myelin in the CNS. We hypothesize that white matter degeneration could be caused by dysfunction and death of oligodendrocytes as a result of accumulated somatic mutations. A post-doc in the Walsh lab is testing this hypothesis. He has developed a protocol to sort single oligodendrocyte nuclei, based on the protocol I used to sort neuronal nuclei, and will perform WGS to sequence several single nuclei. Since oligodendrocytes have particularly high expression of myelin and
because TCR is defective in CS and XP cells, we predict oligodendrocytes from these patients will accumulate mutations in *MBP* and other genes in this pathway. If we observe increased somatic mutation rate and deleterious mutations in *MBP*, it will provide an explanation for white matter damage in CS and XP and provide more support for our argument that CS and XP cause neurodegeneration through increased somatic mutation rate.

**Structural variants and copy number variants in CS**

Another important question to explore is whether other types of somatic variation besides SNVs are increased in CS. There is evidence that copy number variation (CNV) and structural variation (SV) are abnormal in CS cells (Wijnhoven et al., 2001). CNVs and SVs are an important source of genetic variation in normal and abnormal brain development. For example, duplication and deletion at 16p11.2 locus is a recurrent genetic cause of autism (Sanders et al., 2011), patients with trisomy 21 have early onset neurodegeneration (Choong et al., 2015), and recent evidence suggests that double-stranded breaks are generated in promoters of early response genes following neuronal activity (Madabhushi et al., 2015). In addition, genetic disorders of double stranded break repair, such as ataxia telangiectasia, cause neurodegeneration (Savitsky et al., 1995). Our collaborators in Peter Park’s lab are developing tools to identify CNVs and SVs in single neuron WGS datasets. This work will help us build a more comprehensive understanding of somatic mutation disruption in CS. If we find changes in CNV and SV rates in CS, it will show links between different DNA repair mechanisms in this disease.
3.5 Experimental Procedures

**Single neuron MDA**

A small piece of frozen cortex (~100mg) was manually homogenized in lysis buffer containing: 0.32M sucrose, 5mM CaCl2, 3mM MgAc2, 0.1 mM EDTA, 10mM Tris-HCl, pH 8, 1 mM DTT, 0.1% Triton X 100. Homogenized cortex was layered over a cushion buffer containing: 1.8M sucrose, 3mM MgAc2, 10mM Tris-HCl, pH 8, 1 mM DTT. Homogenized cortex was centrifuged at 13,000 rpm in a Beckman SW-28.1 rotor for 2 hours at 4 C and pelleted nuclei were resuspended in PBS with 3mM MgCl2 and then stained with NEUN (Millipore MAB377) antibody and Alexa 647 goat anti-rabbit and Alexa 488 donkey anti-mouse secondaries. Stained nuclei were then sorted into each well of a 96 well plate with a BD FACS Aria fluorescence cell sorter. Nuclei were lysed and genomic DNA was amplified with a customized phi-29 based MDA genome amplification protocol (Evrony et al., 2012).

**Bulk DNA isolation**

Bulk DNA was isolated from frozen heart from individuals 1465, 4638, and 4643, and from frozen cortex from all other individuals. Tissue was lysed in buffer ATL (Qiagen) at 55 C. DNA was isolated and purified with phenol-chloroform extraction and ethanol precipitation (D’Gama et al., 2015).

**Low coverage WGS to determine coverage**

We assessed individual neuron DNA for genome amplification coverage using two PCR based assays at multiple loci and with low coverage (0.1x) WGS (Evrony et al., 2012; Lodato et al., 2015). After selecting cells with the most complete and uniform
amplification from each brain we performed whole genome sequencing (40-45x target coverage).

**Library preparation and whole genome sequencing**

Libraries from single neuron MDA genomic DNA or from bulk unamplified genomic DNA were prepared using commercial kits. All libraries were paired-end and barcoded. Individual 1465 bulk heart and single neuron libraries were prepared using the NEXTflex DNA sequencing kit (Bioo scientific). Individual 4638 and 4643 bulk heart and single neuron libraries were prepared using the TruSeq Nano LT sample preparation kit (Illumina). Individual 1762, 5105, 5316, and 5416 bulk cortex and 2 single neuron libraries were prepared using the TruSeq Nano kit. All other libraries were prepared using the TruSeq PCR Free kit (Illumina).

1465 samples were sequenced with 100bp x 2 or 101bp x 2 paired-end reads on a HiSeq 2000 sequencer (Illumina) at Harvard Biopolymers facility and Axeq. 4638 and 4643 samples were sequenced with 150bp x 2 paired-end reads on a HiSeq X sequencer at New York Genome Center. 1762, 5105, 5316, and 5416 bulk cortex and 2 single neuron libraries were sequenced with 150bp x 2 paired-end reads on a HiSeq X at New York Genome Center. All remaining samples were sequenced with 150bp x 2 paired-end reads on a HiSeq X at Macrogen.

**WGS read mapping**

Paired-end reads were aligned with Burrows-Wheeler Aligner (BWA) with default parameters (Li and Durbin, 2009). For 1465, BWA version was 0.6.2-r126 and reference genome was GRCh37 + decoy. For 4638 and 4643 BWA version was 0.7.8-r455 and reference genome was GRCh37. For all other samples BMA version was 0.7.8-
r455 and reference genome was GRCh37 + decoy. MarkDuplicate of Picard tools was used to identify duplicate reads (http://picard.sourceforge.net). Genome Analysis Toolkit (GATK) IndelRealigner was used for local realignment around indels and base quality score (McKenna et al., 2010).

**Allele and locus dropout analysis**

Allele and locus dropout rates were calculated by identifying loci in single cell genomes missing one or both bulk alleles. Heterozygous bulk SNPs with quality score ≥100 were identified and then these SNPs were examined in each single cell. If one allele of the SNP was missing from a single cell it was labeled “allele dropout.” If both alleles were missing, and no genotype was called at that base, it was labeled “locus dropout.”

**Bulk tissue sequencing to define germline SNVs**

In order to distinguish somatic SNVs from germline SNVs we also sequenced unamplified bulk DNA from each individual (WGS, 40-45x target coverage). In the three normal individuals, heart tissue was used as the bulk control because it is derived from endoderm in contrast to cortical tissue, which is derived from ectoderm. Other tissue besides brain was generally unavailable for CS and XP cases, so we used cortical tissue as bulk control. Because bulk and single cells were taken from the same tissue in the CS and XP individuals, it is possible that some somatic variants present widely in cortex would be filtered out as being germline and so the SNV estimate could be slightly lower than if heart was used as the bulk control. However, previous work from our lab showed that SNVs present in 15% or more neurons in bulk cortex are also present in
heart, suggesting that SNVs called heterozygous in bulk cortex would also be called heterozygous in bulk heart (Lodato et al., 2015).

**Somatic SNV calling workflow**

Somatic SNVs were called in single neurons using a SNV calling algorithm that overlaps calls from three independent SNV callers developed in Lodato et al., 2015. Three sets of SNVs calls were intersected to generate a list of high confidence triple-called variants. Lodato et al. showed that these SNVs are validated by PCR and Sanger sequencing of DNA from the single neuron at a rate of 90%.  

1. SNVs were called with GATK (McKenna et al., 2010) by joint calling using either the Unified Genotyper module (brains 1465, 4638, 4643, GATK version 2.8) or the Haplotype Caller module (all other brains, GATK version 3.5).  
2. Somatic SNVs were called with MuTect (version 1.17, Cibulskis et al., 2013)  
3. Somatic SNVs were called with VarScan2 (version v2.3.8, Koboldt et al., 2012). These three lists of SNVs were intersected to create a list of triple-called SNVs for each neuron. Finally, any somatic SNV that was called at a base with locus dropout in another cell from that same brain was discarded.
Chapter 4: Discussion
4.1 Summary

My dissertation research, presented in Chapters 2 and 3, focused on understanding genetic mechanisms active at the beginning and end of a neuron’s life. Chapter 2 focused on neurodevelopment: I presented evidence that CHMP1A, a microcephaly and cerebellar hypoplasia protein, is essential for brain development because it is required for exosome-mediated secretion of sonic hedgehog. Here, I will highlight the promise of future exploration of exosome biology in the brain.

Chapter 3 focused on neurodegeneration: I showed that the rate of single neuron somatic SNVs is increased in patients with Cockayne syndrome (CS), a genetic syndrome of defective DNA damage repair and neurodegeneration. These results suggest that DNA damage and somatic mutation are potential mechanisms of neurodegeneration, and in this chapter, I will discuss broader implications of this work including the promise of characterizing somatic variation in other forms of degenerative neurologic disease.
4.2 CHMP1A is required for exosome-mediated SHH secretion

**CHMP1A illustrates the importance of exosome mediated cell communication during development**

In the first part of my dissertation work, I studied the beginning of a neuron’s life by examining the function of CHMP1A, a human microcephaly and cerebellar hypoplasia protein. I presented evidence in Chapter 2 that CHMP1A is essential for brain development because it is required for exosome release and exosome-mediated secretion of SHH in several developmental contexts. Exosomes contain protein cargo such as growth factors and other bioactive molecules including mRNAs and miRNAs that enable cell-to-cell communication (Cocucci and Meldolesi, 2015). Defects caused by the absence of CHMP1A, including, microcephaly, cerebellar hypoplasia, and short stature, highlight the importance and widespread function of exosome-mediated cellular communication during development.

**Future directions: exosome biology in the vertebrate CNS**

Our results provide the first *in vivo* evidence for exosome function in vertebrate brain development and the *Chmp1a* null mouse provides a tool to further explore exosome biology in the brain. Based on *in vitro* experiments and work in *Drosophila*, there are many potential CNS functions of exosome release in addition to SHH driven neural progenitor proliferation.
Mapping physical locations of MVBs and exosome release

We have shown that MVBs are present in Purkinje cell dendrites and are present throughout the dendritic tree of a cortical pyramidal neuron. This suggests exosome release from neuronal dendrites may be widespread in the CNS. One important approach for understanding *in vivo* functions of exosomes in the brain is comprehensive MVB mapping within neurons. We have started collaborating with Wei Lee in the neurobiology department to do this. Wei recently published an incredible TEM reconstruction of mouse cortex (Lee et al., 2016). He reconstructed a block of visual cortex 450 x 450 x 150 microns including cortical layers I-III. We have started to use this data set to identify MVBs in pyramidal neurons, but there are several more analyses we are eager to begin. Wei and his team have fully traced dendritic and axonal arbors of 50 pyramidal cells in the volume. We would like to begin by annotating MVBs in these neurons. This will allow us to determine how MVBs are distributed: are they more abundant in proximal or distal dendrites, do they cluster near branch points, are they enriched near synapses, and where are they located in the axon? Answering these questions will allow us to more precisely predict what nervous system functions require exosome release. In addition, Wei has used pre-fixation calcium imaging to determine the visual tuning properties (orientation selectivity) of each reconstructed pyramidal cell. Our annotation of MVBs in these neurons will allow us to determine if there is a relationship between tuning properties and exosome function. For example, it could be that highly tuned neurons have more MVBs or MVBs with a different cellular distribution, compared to poorly tuned neurons. We also hope to develop a machine-learning algorithm to automatically annotate MVBs throughout the entire reconstruction.
This would enable us to visualize MVB fusion and exosome release, a relatively rare event, and would directly show the location of exosome release within pyramidal cell dendrites and with respect to surrounding structures including other dendrites, axons, and glial processes. Mapping MVB location and locations of exosome release will provide global information about neuronal exosomes in the cortex and direct our attention to specific potential functions.

**Activity-dependent exosome release**

Recent studies have demonstrated that exosome release from neurons and glia can be triggered by membrane depolarization (Budnik et al., 2016). Another idea for further exploration is that activity dependent release of SHH and other exosome cargos is a mechanism that links neuronal activity and CNS development. There are several examples of activity-development coupling that could occur through exosome release. There is evidence that neuron depolarization and calcium signaling is required for proliferation of cortical progenitors (Uhlén et al., 2015). CHMP1A is expressed in postmitotic neurons during cortical development; perhaps, in response to depolarization by calcium waves in the cortical plate (Yuryev et al., 2015), these cells release exosomes containing mitogens for progenitor proliferation. WNT ligands could be released this way as *Wnt1* mRNA is expressed in postmitotic neurons in the cortex during neurogenesis (Allen Brain Atlas) and WNT is secreted on exosomes (Gross et al., 2012). Similarly, in the developing cerebellum, calcium waves can be detected in Purkinje cells at P4-P6, the peak of GCP proliferation (Watt et al., 2009). Purkinje cell depolarization could trigger exosome release and SHH secretion to drive GCP proliferation at this age.
Later, during cortical synaptogenesis, SHH expression in dendrites is required for synapse formation between pyramidal cells (Harwell et al., 2012). Exosome release following depolarization could release SHH from dendrites of active synapses to attract more incoming axons or to activate pruning of weak synapses. In fact, there is *in vitro* evidence that exosomes from neuronal PC-12 cells activate complement expression in microglia and increase neurite pruning via complement-mediated microglia engulfment (Bahrini et al., 2015). Activity-dependent exosome release from strong synapses could be the punishment signal that activates microglial pruning. Activity driven exosome release is of great interest because it can provide one mechanism for coupling between neuronal activity and development.

**Exosomes mediate retrograde signaling**

In addition to activating synapse pruning via microglia, SHH positive exosomes released at a cortical pyramidal cell synapse may bind to BOC on incoming axons and provide a retrograde signal for synapse formation. At the *Drosophila* NMJ, exosome-mediated secretion of Syt4 activates retrograde signaling. Vivian Budnik’s lab showed depolarization of the postsynaptic muscle cell drives ghost bouton formation in the presynaptic motorneuron and that Syt4 is a retrograde signal first released by the presynaptic neuron on exosomes and then required in the postsynaptic muscle for ghost bouton formation (Korkut et al., 2013). In the mouse cortex, Corey Harwell showed that both SHH expression in dendrites of layer V pyramidal cells and expression of a SHH receptor, BOC, in incoming axons of callosal and local projection neurons are required for synapse formation (Harwell et al., 2012). A plausible model is that exosome-bound
SHH released from dendrites binds to BOC on incoming axons and activates a retrograde signaling cascade similar to the cascade activated by SHH-BOC binding during commissural axon migration in the developing spinal cord (Okada et al., 2006). Additionally, because exosome release can be triggered by depolarization, SHH secretion may increase after initial electrical contact to strengthen that synapse and recruit more axons to nearby layer V dendrites. If this were the case, it would suggest that exosome driven retrograde signaling is a key mechanism underlying activity-dependent axon guidance and synapse formation in developing brain.

**Exosome also carry miRNAs**

In addition to growth factors such as SHH and WNT, exosomes transport RNA cargo. In particular, cultured mouse cortical neurons release exosomes containing MIR124A (Morel et al., 2013). Release of these exosomes alters expression of EAAT2, an amino acid transporter on astrocytes, suggesting that exosome-mediated secretion of miRNA can regulate functional properties of receiving cells. In fact, a recent paper describing mice with mutant *Chmp2b* may have unknowingly discovered this same mechanism. Fen-Biao Gao’s lab reported that transgenic mice expressing a dominant mutation in *Chmp2b* (from a patient with frontotemporal dementia) have decreased MIR124A levels in cortex and social behavior deficits (Gascon et al., 2014). In their discussion they write that the molecular link between *Chmp2b* mutation and decreased MIR124A is unknown; however, a reasonable explanation is that *Chmp2b* mutation causes a defect in exosome release, which then impairs MIR124A secretion. Indeed, they found altered AMPA receptor subunit expression in mutant mice, which fits very well
with the work showing MIR124A exosomes regulate EAAT2 expression in astrocytes. These papers not only show that miRNA secretion via exosomes plays an important role in brain function, but also illustrate the idea that exosome biology in the brain is ubiquitous, and often overlooked. Greater awareness of exosome biology will encourage a more comprehensive understanding of neurodevelopment and neurobiology. In particular, exosome-mediated secretion of miRNAs highlights the idea that bioactive molecules previously thought to be unstable in the extracellular environment or not secreted can use exosomes for secretion and can have important functions in cell-to-cell communication. Further, because neurons can take up RNA from exosomes, exosomes can provide a delivery vehicle for RNA based therapeutics, such as siRNAs targeting overactive genes in cancer or neurodegeneration.

**Exosome-mediated secretion of neurotoxic proteins**

Related to this idea, a new hypothesis suggests that neurodegenerative diseases can spread through the brain via exosome-mediated seeding of previously healthy neurons. Exosomes were first connected to neurodegeneration when prion protein was discovered in exosomes isolated from ovine CSF and it was demonstrated that injection of these exosomes into wild type mice induced prion disease (Vella et al., 2008). Since that discovery, exosomes secreted by both neurons and astrocytes have been found to contain phosphorylated tau and beta amyloid in models of Alzheimer’s disease (AD), alpha-synuclein in Parkinson’s disease (PD), and mutant SOD1 in amyotrophic lateral sclerosis (ALS) (Basso and Bonetto, 2016). The seeding hypothesis of neurodegeneration developed because these proteins can act like prions and induce
misfolding of wild type protein and because exosomes containing these proteins can physically spread disease. For example, spread of tau protein from the entorhinal cortex to the hippocampus in a mouse model of AD was found to require exosome secretion from microglia (Asai et al., 2015). These discoveries not only revealed a pathologic type of cell-to-cell communication, but also provided a new pathway, exosome secretion, for therapeutic development. Exosomes are active during brain development, normal function, and in degeneration, and provide a therapeutic target for restoration of normal brain function.

**Exosome-mediated secretion and cell-to-cell communication**

My work and the work of others shows that exosomes are ubiquitous in the brain and are involved in several aspects of brain development and function. An important question is why are exosomes involved in so many neural processes, ranging from progenitor proliferation, to synaptic formation and plasticity, to neurodegeneration? One feature connects all of these examples of exosome function: cell-to-cell communication mediated by insoluble or unstable signaling molecules. Exosomes mediate secretion of hydrophobic SHH, WNT and Syt4, insoluble tau and prion protein, and unstable miRNAs. Rather than requiring a different secretion mechanism for each of these molecules, exosomes – structures that are compatible with hydrophobic proteins and provide protection from the extracellular environment – offer a single mechanism that can be used for many different types of cargo. Exosomes enable greater complexity in CNS growth and function because they expand the library of available molecules cells
can use to communicate. It is for this reason that exosomes are essential for normal growth and function of the brain.

**CHMP1A project conclusion**

Our discovery that exosome-mediated SHH secretion is required for normal growth of cortex and cerebellum illustrates the role of exosome biology in neurodevelopment and highlights the intersection between exosomes and neurobiology. As I have illustrated above, promising preliminary results support several additional functions for exosomes in the brain, functions that can be better studied with models such as the *Chmp1a* null mouse we generated to study neurodevelopment and the *Chmp2b* mutant mouse generated to study neurodegeneration. My work provides an example of the unexpected mechanisms that can be revealed by studying human genetics of brain malformations and shows how this line of inquiry continues to drive exciting basic neurobiology.
4.3 Neuronal somatic mutation is increased in disorders of DNA damage repair

Neuronal somatic mutation in Cockayne syndrome

In the second part of my dissertation work, I examined Cockayne syndrome, an inherited disorder of neurodegeneration, to understand genetic mechanisms that regulate the end of a neuron’s life. We performed whole genome sequencing (WGS) of single cortical neurons from patients with Cockayne syndrome (CS) and our results, while preliminary, show a 4-fold increase in the rate of somatic single nucleotide variants (SNVs) compared to neurons from unaffected individuals. SNVs in CS neurons are more common at a range of alternate allele frequencies, but are especially increased for SNVs with 25% AAF, suggesting that CS neurons have increases in both fixed mutations and single-stranded DNA lesions. CS SNVs have a distinct mutational signature from SNVs in normal neurons suggesting that a disease-specific process drives increased somatic SNV rate. In addition, among CS neuron SNVs, we found several truncating mutations in genes important for neurologic disease and neuron function. Together, our results provide the first genome-wide measurement of somatic mutation in a neurodegenerative disease and suggest a model for neurodegeneration in which both DNA damage and somatic mutations in essential genes accumulate and eventually cause neuron cell death and neurodegeneration.

Although we still have experiments underway to better understand the role of somatic mutation in neurodegeneration of CS, our preliminary finding that somatic SNV rate is dramatically increased in neurons from these patients provides a starting point to
think more broadly about somatic mutation measurement, somatic mutation in other types of neurodegeneration, and somatic mutation as a tool of evolution.

**Cell culture free method to measure neuronal somatic mutations**

Our sequencing of single cortical neurons from patients with CS provides additional evidence that single cell WGS following multiple displacement genome amplification (MDA) can accurately measure somatic mutation. This is important because many cells of interest cannot be amplified through tissue culture. Investigating somatic mutation in human brain almost always requires post-mortem tissue, because obtaining fresh resected brain tissue is difficult and rare. In addition, human somatic cell nuclear transfer to establish a stem cell line using a neuron’s nucleus requires human oocytes and is legally infeasible. In short, tissue culture is not feasible for somatic mutation analysis in human neurons, and so my work, which provides additional evidence for the accuracy of somatic mutation identification following MDA, opens the door for exploration of many new questions. Single neuron WGS following MDA can address several current questions in the field including whether somatic mutation rate is increased in neurons of geriatric individuals, patients with neurodegenerative diseases (AD, PD, ALS), patients with autism spectrum disorder, and patients with psychiatric disease (schizophrenia, bipolar disorder, depression). In parallel, answering these same questions in glial cells, including astrocytes, oligodendrocytes, and microglia, is also a high priority.
Role of somatic mutation in neurodegeneration

Increased DNA damage and decreased DNA repair is hypothesized to underlie aging and neurodegeneration. Our investigation of somatic mutation in CS neurons can provide insights in how somatic mutations resulting from impaired DNA repair cause degeneration. Although somatic mutation rate has not been measured in other degenerative diseases, there is evidence for many types of DNA damage in AD, PD, and ALS including oxidative damage, DNA adducts, strand breaks, cytogenetic abnormalities, and mitochondrial DNA mutations (Coppede and Migliore, 2015). Our finding that somatic SNV mutation rate is dramatically increased in patients with neurodegeneration caused by defective nucleotide excision repair in CS suggests that degeneration in other diseases may also be caused by somatic mutations resulting from increased DNA damage. If this is the case, single neuron sequencing in aged brains and brains from patients with AD, PD, and ALS can provide valuable information for new therapeutics by identifying pathways with recurrent deleterious mutations. This finding also suggests treatments that increase DNA repair and decrease DNA damage are important directions for therapeutic development in neurodegeneration.

What is the role of somatic mutation in the brain?

Larger questions raised by these results include, what role does somatic mutation play in evolution, and how is the balance defined between too little somatic mutation and too much? Introduction of genetic variants in each generation enables evolution because new traits that result from these variants can be selected for or against in subsequent generations to adapt to a changing environment. However, loss of DNA repair
mechanisms causes catastrophic disorders including CS and XP that are associated with early lethality. Thus, for a species’ long-term survival, there must be a balance between too few and too many DNA variants. An open question in genetics is whether this argument can also be applied at the cellular level. Previous work from our lab has shown that adult human neurons contain about 1,500 somatic SNVs and that over 99% of these are unique to each neuron. On the other hand, neurons from patients with CS contain about 5,500 somatic SNVs per cell and these patients experience rapid neurodegeneration. One explanation for the large number of somatic SNVs in neurons of normal individuals is simply that neurons have constantly high levels of transcription, which results in accumulation of somatic SNVs from transcription-associated DNA damage. An alternate explanation is that somatic SNVs serve a purpose in neuronal function. Somatic SNVs can change expression of neuron identity genes and give each neuron a unique molecular identity to differentiate self and other or to provide that neuron with a unique set of functional capacities. It is not yet known whether somatic mutations play a role in neuronal function or if they are simply collateral damage from transcription. Exploring this question further will increase our understanding of normal brain function.
4.4 Conclusion

In this dissertation, I have presented a story about neurodevelopment and beginning of a neuron’s life: sonic hedgehog is secreted on exosomes in the developing brain and this process requires the microcephaly protein, CHMP1A, and a story about neurodegeneration and the end of a neuron’s life: single neurons from patients with Cockayne syndrome have a much higher burden of somatic mutation than normal individuals. My study of CHMP1A reveals a novel mechanism for SHH secretion in vertebrates and highlights the importance of exosome biology in brain development and function. My work identifying somatic SNVs in CS neurons shows that accumulated DNA damage and increased somatic mutation rate can lead to neurodegeneration and illustrates the importance of studying somatic mutation in common types of neurodegeneration. Together, these two studies show that studying inherited human neurologic disorders can reveal unexpected biological mechanisms the shape the brain and highlight ideas broadly relevant to brain development, function, and degeneration.
References


D’Gama, A.M., Pochareddy, S., Li, M., Jamuar, S.S., Reiff, R.E., Lam, A.-T.N., Sestan,


Kim, S., Lehtinen, M.K., Sessa, A., Zappaterra, M.W., Cho, S.-H., Gonzalez, D., Boggan,


McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A.,


