4

What Happens When It Goes Wrong?

Using Human Genetics to Understand Human Brain Development and Evolution

Michael E. Coulter and Christopher A. Walsh

Abstract

This chapter explores what happens when the development of the cerebral cortex goes awry. It presents results on work with *CHMP1A* mutations, which highlight the importance of specialized cell-to-cell communication via extracellular vesicles in cortical development and function. It reviews genetic causes of microcephaly, with an emphasis on centrosomal proteins, and presents novel insights about cortical evolution shown using a ferret model of microcephaly caused by *ASPM* loss of function. It reviews recent work to identify noncoding mutations that cause brain malformations, which has expanded understanding of cortical development beyond protein-coding genes. These three examples illustrate general principles of cortical growth and function (cellular communication and synaptic plasticity, evolution, and utilization of large data sets), made possible by recent advances in DNA sequencing technology.

Introduction

The human genetics of cerebral cortical malformations and developmental disorders has provided a powerful tool to identify genes essential for cortical development. Since the Dahlem Workshop some thirty years ago, and rapidly accelerating since 2001 with completion of the Human Genome Project, dozens of cortical development genes with diverse functions have been identified using this approach (Zhang et al. 2014). Over the last decade, it has been appreciated that those genes essential for human brain development represent a rich source of genes modified during the evolution of humans

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

to provide the human brain with its large size and unique network properties. A handful of human genes show evidence of having been added during the human lineage, whereas others show evolutionary selection at the level of the protein-coding sequence, where changes in amino acid sequence can result in changes in biochemistry. Only 2% of the human genome represents protein-coding genes (Gregory 2005). Larger numbers of genes show evidence of evolutionary changes that occurred to their noncoding sequences, which have the potential to alter patterns or timing of gene expression (Reilly et al. 2015). Recently, knowledge of sequence variation and presumptive functions of the noncoding genome has permitted initial insights into how changes in the noncoding portions of the genome might relate to human disease or to evolutionary change.

Genetics of Developmental Disabilities of the Brain

Identifying the genetic causes of human cortical malformation disorders is a powerful tool for revealing critical mechanisms of cortex development and function. With a total human population of over 7.5 billion, it is likely that every gene has mutated multiple times in humans, and that every gene has mutated at least once across all cells in any individual (Bernards and Gusella 1994; Walsh 1999; Brenner 2003; Walsh and Engle 2010). As a result, individuals with disorders of cortical development represent an unbiased screen for genes that are essential to that process. Catalyzed by publication of the human genome sequence 18 years ago (Lander et al. 2001) and by widespread availability of high-throughput DNA sequencing technology about ten years ago, dozens of genes that cause such developmental disorders when mutated have been identified through the sequencing of single disease-linked genes, whole exome sequencing (which sequences all protein-coding genes), and whole genome sequencing (which sequences all DNA, both coding and noncoding).

The following molecular mechanisms of cortical development have been discovered by identifying genes mutated in cortical malformations:

- The role of centrosomes and mitotic spindles in cortical neurogenesis: *ASPM*, *CDK5RAP2*, *WDR62*, *NDE1*, KATNB1, *CEP63* (Bond et al. 2002, 2005; Shen et al. 2005; Hassan et al. 2007; Nicholas et al. 2009, 2010; Bilgüvar et al. 2010; Yu et al. 2010; Alkuraya et al. 2011; Bakircioglu et al. 2011; Sir et al. 2011; Hu et al. 2014b; Mishra-Gorur et al. 2014)
- The role of extracellular matrix proteins and G-protein coupled receptors in maintaining integrity of the pial surface: *POMT1*, *FKTN*, *FKRP*, *GTDC2*, *POMK*, *GPR56* (Beltrán-Valero de Bernabé et al. 2002; Silan et al. 2003; Beltrán-Valero de Bernabe 2004; Piao 2004; Manzini et al. 2012; Bae et al. 2014; Di Costanzo et al. 2014)

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

- Regulation of microtubule dynamics during neuron migration: *DCX*, *LIS1*, *TUBA1A*, *TUBB3*, *TUBB5*, *DYNC1H1*, *KIF5C*, *KIF2A* (Reiner et al. 1993; Lo Nigro et al. 1997; des Portes et al. 1998; Gleeson et al. 1998; Keays et al. 2007; Poirier et al. 2010, 2013; Breuss et al. 2012)
- The role of amino acid synthesis and metabolism: *QARS*, *AMT* (Yu et al. 2013; Zhang et al. 2014)
- The role of DNA damage repair: *NBN*, *PNKP* (Varon et al. 1998; Shen et al. 2010)
- The role of transcriptional regulation: *MECP2*, *ZNF335* (Amir et al. 1999; Yang et al. 2012)

These mechanisms span the cell types of cortical development from undifferentiated progenitors to radial glial cells, to committed neural progenitors, to postmitotic neurons. They also span cellular processes from progenitor proliferation, to neuron differentiation, to neuron migration. This diverse and widespread list of affected processes highlights the unbiased and saturating nature of cortical malformation human genetics, illustrating the extent to which human genetics can systematically identify mechanisms underlying key steps of normal development.

Extracellular Vesicles in Cortex Formation and Function

An example of the surprising novel mechanisms that can be identified by human genetic screens involves the recent analysis of *CHMP1A* mutations that cause microcephaly and cerebellar hypoplasia (Mochida et al. 2012), which unexpectedly implicates small extracellular vesicles (EVs) in cortical and cerebellar development. EVs are small membrane-bound vesicles that are released by many cell types for specialized cell-to-cell communication through transfer of unstable molecules, such as RNA (Tietje et al. 2014), or hydrophobic proteins, such as transmembrane proteins and some growth factors (Korkut et al. 2009; Budnik et al. 2016). EVs are released by neurons and glia and may have many roles in the nervous system (Amir et al. 1999; Lachenal et al. 2011; Frühbeis et al. 2013). EVs, for example, have been implicated in wingless secretion during neuromuscular junction synapse growth (Koles et al. 2012) and Synaptotagmin 4 secretion in retrograde signaling (Korkut et al. 2009) through *in vivo* experiments in *Drosophila* as well as in synaptic strength modulation (Lachenal et al. 2011), and prion-like protein and Tau secretion through experiments in cultured mammalian neurons (Asai et al. 2015). Indeed, newly published work directly links EVs to synaptic plasticity through *Arc*, a master regulator of activity-dependent glutamate receptor trafficking (Pastuzyn et al. 2018). In *Drosophila*, EVs enable Arc1 protein transfer between neurons and muscle cells during neuromuscular junction synapse maturation; in mammalian neurons, EVs enable transfer of *Arc* mRNA between neurons allowing

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

localized Arc translation in the recipient cell that modulates synaptic strength (Ashley et al. 2018; Pastuzyn et al. 2018).

Recently, we have identified a new role for EVs in cortex and cerebellum development by creating a mouse model of a human microcephaly gene, *CHMP1A* (Coulter et al. 2018). Human loss-of-function (LOF) mutations in *CHMP1A* cause recessive microcephaly with severe cerebellar hypoplasia (Mochida et al. 2012), and we found that a *Chmp1a* null mouse model recapitulated this phenotype. Investigating the mechanism of microcephaly in *Chmp1a* null mice, we found that secretion of the hydrophobic growth factor sonic hedgehog (*SHH*) is substantially reduced in the embryonic cerebral spinal fluid (CSF). CHMP1A is a member of the ESCRT protein complex. Since one of its functions involves the formation of multivesicular bodies (MVBs) and EV secretion, we examined MVBs in *Chmp1a* null mice. We found that MVBs were abundant in choroid plexus epithelial cells and in cerebellar Purkinje cells, two sources of SHH during brain development, and that in the absence of *Chmp1a*, each MVB had fewer luminal vesicles. To test the hypothesis that *CHMP1A* regulates SHH secretion via EVs, we turned to *in vitro* experiments and found that in the absence of *CHMP1A*, SHH-positive EV secretion was impaired. We characterized these EVs using protein mass spectrometry and found they are a new subtype of EVs, which we call ART-EVs, that carry SHH protein. Intriguingly, we found that SHH-positive ART-EVs exist in adult human CSF and that MVBs are abundant in adult mouse cortical pyramidal cells; together, this suggests a continued role of EVs in adult cortex function. Serial TEM reconstruction showed that MVBs are often located near synapses in the dendrites of pyramidal neurons (Figure 4.1); together with recently published work showing that *Arc* mRNA and protein is transferred between neurons via EVs (Ashley et al. 2018; Pastuzyn et al. 2018), these findings highlight the likely importance of EVs in mechanisms

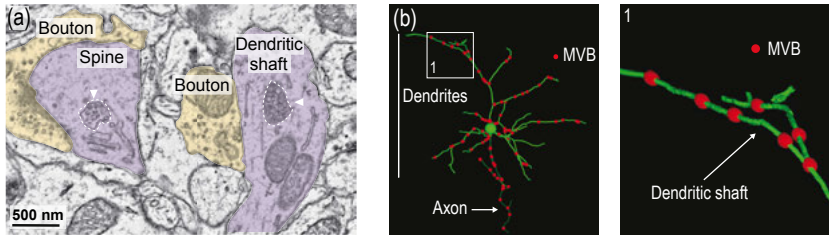

Figure 4.1 Multivesicular bodies (MVB) near synapses in dendritic tree of mouse cortical pyramidal neuron: (a) High magnification EM image of mouse cortex neuropil showing MVBs (white arrows) near synapses. Two synapses are shown with the presynaptic bouton (yellow) and the postsynaptic dendrite (purple). (b) Three-dimensional reconstruction of a single pyramidal cell from mouse cortex serial EM showing >80 MVBs (red) distributed throughout the dendritic tree and the axon. Inset 1 shows an enlarged section of the dendritic tree from the reconstructed pyramidal cell with MVB locations labeled in red. Data in figure from Lee et al. (2016).

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

of cortical processing such as LTP or LTD. Further, during late cortical development, SHH protein secretion by layer V corticofugal projection neuron dendrites is required for local synapse formation with callosal projection neurons (Harwell et al. 2012), thus suggesting a potential additional role for EV-mediated SHH secretion in spatially restricted neuronal connectivity. The *Chmp1a* null mouse provides a tool for future experiments to explore more comprehensively the role of EVs and this form of cellular communication in development and function of the cortex.

Mitotic Spindle and Centrosome Mutations Are a Common Cause of Microcephaly

Two of the earliest genes cloned in recessive microcephaly, *APSM* and *CDK5RAP2*, encode proteins that localize to the centrosome and are required for mitotic spindle organization in cortical progenitors (Bond et al. 2002, 2005; Hassan et al. 2007; Nicholas et al. 2009; Pagnamenta et al. 2012). Since then, several additional microcephaly genes have been shown to localize to the mitotic spindle or to the centrosome and centrioles, showing that these structures have a central role in the development of the cortex (Hu et al. 2014a). The centrosome is a microtubule-organizing protein complex that is present in cells during interphase (reviewed in Fu et al. 2015). It is composed of two smaller structures called centrioles; one centriole, the mother, has distal and subdistal appendages that enable it to form the basal body of the primary cilium during G1. During the DNA replication (S) phase of mitosis, the centrosome is duplicated (creating 4 centrioles); then, during metaphase, the two centrosomes migrate to opposite ends of the cell and form the poles of the mitotic spindle. As spindle poles, they recruit and organize microtubules to become the spindle. Finally, following cytokinesis, each daughter cell receives one of the two spindle poles, which become centrosomes again in the new cells (O'Connell et al. 2001; Balestra et al. 2013). Based on these roles, it is no surprise that disruptions of centrosome number, structure, or function impair cell division and neurogenesis. In fact, microcephaly proteins have been discovered that impair each step in the centrosome cycle.

KATNB1, encoding a microtubule-severing protein in which LOF mutations cause severe microcephaly, controls centrosome duplication and when *KATNB1* is absent, cell division is impaired on account of supernumerary centrosomes and disordered mitotic spindles (Hu et al. 2014b; Mishra-Gorur et al. 2014). *WDR62*, in which LOF mutations cause microcephaly, encodes a protein that localizes specifically to the mother centriole and is required for centriole and centrosome duplication. In the absence of *WDR62*, centrosomes fail to duplicate during S-phase and mitosis is subsequently impaired (Bhat et al. 2011; Jayaraman et al. 2016). Interestingly, milder, partial LOF mutations in *WDR62* cause other cortical malformations, without microcephaly, suggesting

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

centrioles also play an important role in cortex development outside of progenitor proliferation (Murdock et al. 2011). *ASPM* human mutations cause severe primary microcephaly (where brain size is decreased but body size is normal), and in the absence of *ASPM* there is partial loss of centriole duplication (Bond et al. 2002; Jayaraman et al. 2016). *CDK5RAP2* also encodes a centrosomal protein and is found mutated in patients with microcephaly. During mitosis of cortical progenitors, the loss of *CDK5RAP2* creates extra spindle poles that disrupt progenitor cell division and lead to abundant cell death (Lizarraga et al. 2010; Pagnamenta et al. 2012).

Recent work reveals physical and genetic interactions between many microcephaly-related centriole proteins. For example, *WDR62*/*ASPM* double knockout (KO) mice show more severe centriole duplication defects than single KO of either gene alone (Jayaraman et al. 2016). In addition, losing a single *WDR62* allele on an *ASPM* KO background produced an intermediate phenotype. ASPM and WDR62 proteins interact physically, and WDR62 is required to recruit ASPM to the centrosome; these two proteins form part of a larger complex that includes CDK5RAP2, CENPJ, and CEP63. Together, these findings suggest a model of centrosome protein recruitment that occurs in a specific order, with WRD62 recruited before ASPM (Kodani et al. 2015; Jayaraman et al. 2016).

Evolutionary Mechanisms from Cortical Development Disorders

One of the most striking features of the cerebral cortex is the enormous expansion in size as well as regional and cellular complexity throughout the course of mammalian evolution. The cortex has increased relative to body size from mice to humans with a particular increase in the frontal cortex (Rakic 2009). Interestingly, evolution of microcephaly and developmental disorder genes contribute to the genetic mechanisms driving these changes. For example, *FOXP2*, a highly expressed transcription factor in human cortex, was mutated in a British family with severe language impairment (Lai et al. 2001). *FOXP2* shows evidence of human-specific evolution because the few amino acid changes between mouse, primate, and human dramatically alter the array of *FOXP2* transcriptional targets (Konopka et al. 2012), and because mice expressing humanized *FOXP2* exhibit accelerated learning and increased vocalizations (Enard et al. 2002; Fujita et al. 2008). This suggests that the evolutionary changes in *FOXP2* contribute critically to language development, a function unique to human cortex.

There is evidence of positive selection across mammals and an association with increased brain mass in additional cortical development disease genes, including *CDK5RAP2* and *ASPM* (Zhang 2003; Kouprina et al. 2004; Montgomery and Mundy 2014). *ASPM* is an interesting example because both its coding sequence and protein length have increased consistently over evolution. ASPM protein is composed of two N-terminal CH domains and a variable

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

number of C-terminal IQ domains. *Caenorhabditis elegans ASPM* has 2 IQ domains, *Drosophila* has 24, mouse has 55, and human has 63 (Bond et al. 2002; Johnson et al. 2018). Although the shorter sequence of ASPM in mice compared to humans was originally thought to represent a length increase from mice to humans, direct analysis has shown that in fact rodents are outlier mammals, with an unusually short ASPM protein (Zhang 2003): this suggests a potential contribution of this ASPM shortening to the unusually small cortex which characterizes rodents. Interestingly, *ASPM* null mice have very mild microcephaly and thus model the human phenotype only very poorly (Pulvers et al. 2010; Fujimori et al. 2014; Capecchi and Pozner 2015; Williams et al. 2015; Jayaraman et al. 2016). This hypothesis raises the question of whether other mammalian *ASPM* models, whose ASPM sequence more closely resembles that of humans, would exhibit a greater degree of microcephaly in the absence of *ASPM*, and hence provide a better model system.

Aspm Knockout in Ferrets Recapitulates Human Primary Microcephaly

A recent test of the hypothesis that animals with a larger cerebral cortex might better model human microcephaly came by generating *Aspm* KO ferrets through gene editing technology. The ferret cortex is larger than the mouse and, unlike the mouse, is gyrified, like the human cortex. In addition, ferret *Aspm* has 64 IQ domains, similar to 63 in human, and more than the 55 in mouse (Johnson et al. 2018). Moreover, the fetal ferret brain shows a broader diversity of progenitor types than mice, with abundant outer SVZ basal radial glial cells (radial glia lacking an apical process) unlike mice, which virtually lack this progenitor type (Hansen et al. 2010; Reillo et al. 2011). *Aspm* KO ferrets show robust microcephaly with an up to 40% reduction in brain weight and no change in body weight as well as decreased cortical surface area and volume (Johnson et al. 2018; Figure 4.2). *Aspm* KO in ferret provides a much more accurate model of human *ASPM* LOF than mouse and likely reflects the active evolution of *ASPM* in the mammalian lineage that increases the gene length from mice to ferrets to humans.

Ferret and human cortical neurogenesis is driven both by progenitors at the ventricular surface (apical progenitors) and by progenitors above the SVZ (basal progenitors or outer radial glia), whereas outer radial glia are rare or absent in mice (Fietz et al. 2010; Reillo et al. 2011; Johnson et al. 2015). Interestingly, *Aspm* KO ferrets showed an increased number of proliferating cells (Ki67+) in the basal SVZ and intermediate zone compared to wild type, the location of outer radial glia (Johnson et al. 2018). In *Aspm* KO cortex, excess basal proliferative cells formed discontinuous clusters accompanied by reduced thickness of the corresponding ventricular zone, suggesting that premature withdrawal of progenitors from VZ into the oSVZ is the cellular mechanism driving microcephaly (Johnson et al. 2018). oSVZ progenitors are

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

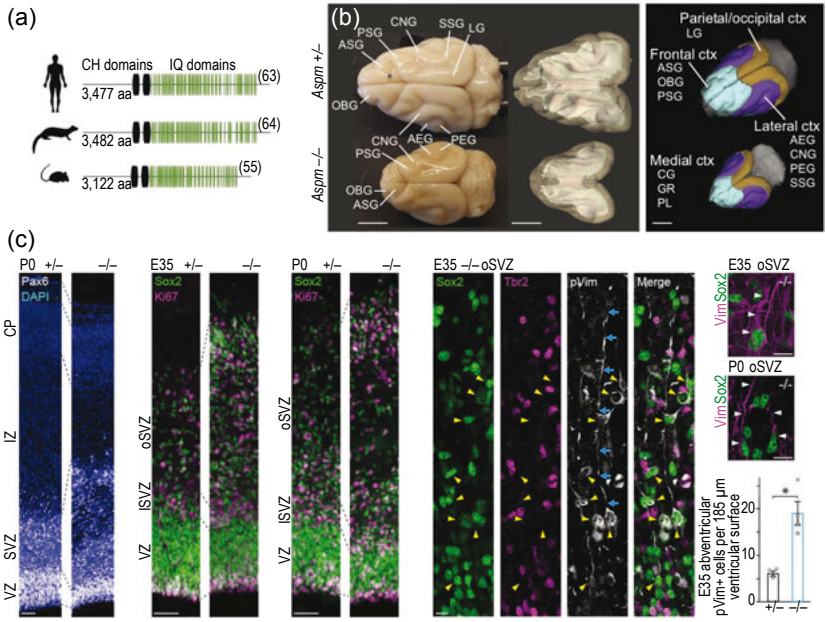

Figure 4.2 *Aspm* KO causes microcephaly in ferret: (a) Protein structure of ASPM in mouse, ferret, and human shows that the number of IQ repeats increases farther in the evolutionary tree. (b) Photograph (left) and MRI reconstruction (right) reveals decreased ferret brain size in *Aspm* –/– ferret compared to *Aspm* +/– control. Gyrification pattern is preserved. Labeling of cortical regions (far right): frontal cortex shows greatest reduction in size of *Aspm* –/– ferret. (c) The outer subventricular zone (oSVZ) is expanded in *Aspm* –/– ferret at birth and in late gestation. Pax6 and Sox2/KI67 stained layer above ventricular zone (VZ) is larger in *Aspm* –/–. In *Aspm* –/– ferret, oSVZ has outer radial glial cells defined as Sox2/Vimentin+ and Tbr2– cells (yellow arrowheads), one of which has a clear process extending toward pial surface (blue arrows). Quantification of increase in Vimentin+ cells in oSVZ in *Aspm* –/– ferret. Adapted from Johnson et al. (2018).

predicted to give rise to fewer postmitotic neurons than apical progenitors; thus, a premature switch to outer radial glia will result in decreased overall neurogenesis.

The essential role of *Aspm* in regulating the switch from apical RG to outer RG provides an evolutionary mechanism that might dynamically control cerebral cortical size, in which subtle variation in *Aspm* sequence or structure might alter the number of apical RG, which in turn would dictate cerebral cortical surface area. This key cellular feature of human cortex development (i.e., neurogenesis via outer RG) is not modeled in mice where this progenitor type is so rare. Hence, ferrets may become a powerful model in the future

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

to understand mechanisms of many human cortical disorders that are poorly modeled in mice.

Coding and Noncoding Mutations in Human Disease

Although 98% of the human genome is noncoding DNA (i.e., DNA that does not encode a protein), recent work has started to identify human cortical disease mutations in noncoding DNA. Noncoding DNA includes introns, unique regulatory elements, transposable elements, and repetitive DNA not related to transposable elements (Gregory 2005). Noncoding cis-regulatory sequences such as promoters and enhancers have been known for several years, and they have been shown to modulate gene expression through transcription factor binding sites as well as steric interactions and DNA folding (Nobrega 2003; Pennacchio et al. 2006). Although noncoding DNA does not directly create protein products, it regulates gene expression and encodes active RNA molecules (Reilly et al. 2015). Indeed, it has been hypothesized that tissue-specific expression and different expression levels dictated by noncoding DNA greatly increases the complexity of the human transcriptome and proteome, even with a relatively small number of coding genes (~19,000), and that this complexity is a key feature of human evolution (Geschwind and Rakic 2013; Kellis et al. 2014; Reilly et al. 2015). Mutations in noncoding DNA have now been identified in diseases of cortical development.

Noncoding Mutations Mimic Phenotype of Loss of Function Mutations

Heterozygous LOF mutations in the growth factor *SHH* cause holoprosencephaly in humans, a syndrome of incomplete separation of the two cerebral cortical hemispheres that results in a single lateral ventricle and craniofacial anomalies (Roessler et al. 1996). Recent work identified a heterozygous mutation in a conserved noncoding element 460 kilobases upstream of *SHH* called Shh brain enhancer-2 (SBE2) (Jeong et al. 2008). Jeong et al. found a single base substitution in a 10 basepair (bp) sequence of SBE2 highly conserved across species, and then showed that this sequence binds the transcription factors (TF) *Six3* and *Six6* and that TF binding was largely abolished by the mutation. In addition, expression of lacZ in developing mouse embryo, driven by wild-type or mutant SBE2, showed that expression in the developing brain was reduced with the patient mutation (Jeong et al. 2008). These findings illustrate how a noncoding DNA mutation can cause neurodevelopmental disease by reducing expression of an essential, dosage-sensitive growth factor in the brain.

Some Noncoding Mutations also Highlight Evolutionary Mechanisms

LOF mutations in the G-protein coupled receptor, *GPR56*, cause a brain malformation syndrome called polymicrogyria (PMG), in which the cortical

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

surface is covered in numerous small gyri (Piao 2004). Recently, a family was identified with a recessively inherited variant form of PMG in which the cortex surrounding the Sylvian fissure was strongly affected with PMG, but the rest of the cortex had normal gyrification (Bae et al. 2014; Figure 4.3). The syndrome showed strong linkage to *GPR56* locus; however, no mutations were identified in the *GPR56* coding sequence. Instead, sequencing of 38 conserved noncoding elements upstream of the first exon revealed a homozygous 15-bp deletion, which segregated with disease. This deletion is located about 150 bps upstream of a noncoding alternative start exon for *GPR56* (e1m). Remarkably, when a large (23 kb) region containing these upstream elements was expressed

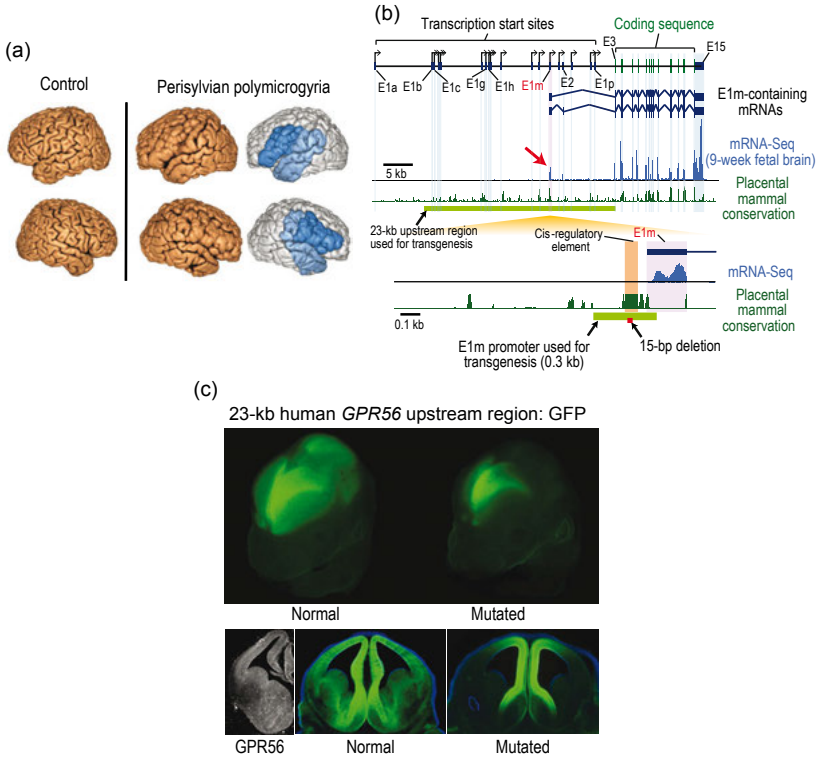

Figure 4.3 Human *GPR56* noncoding deletion causes perisylvian polymicrogyria (PMG): (a) Three-dimensional reconstruction of MRI from patient with GPR56 noncoding deletion shows perisylvian PMG (highlighted in blue). Left image shows reconstruction of a normal brain MRI. (b) 15-bp patient deletion is found near a 5′ UTR noncoding exon in human *GPR56*. This exon, labeled E1m, is highly expressed in human fetal brain (blue trace) and has a highly conserved cis-regulatory DNA element just upstream (lower panel, green trace and highlighted in orange) which contains the 15-bp deletion. (c) Transgenic green fluorescent protein (GFP) expression of 23-kb human *GPR56* upstream region containing E1m promoter [light green bar in (b)] in developing mouse brain, eliminated lateral cortical expression of *GPR56* while preserving medial expression. Adapted from Bae et al. (2014).

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

transgenically in developing mouse brain, the 15-bp deletion eliminated lateral cortical expression of *GPR56* while medial expression was preserved. This finding suggests that this noncoding conserved element is required for expression of *GPR56* in the perisylvian cortex via cis-acting activation of e1m transcription. In this case, spatially localized expression of a cerebral cortex gene is regulated by a noncoding DNA element. This illustrates the importance of gene expression refinement by noncoding DNA and shows how mutation of such DNA can produce a spatially restricted cortical malformation syndrome.

Interestingly, this *GPR56* mutation also illustrates an important concept: in addition to protein-level evolution, as discussed above in *ASPM*, evolution also changes noncoding DNA. Noncoding DNA in the *GPR56* locus is actively evolving and has greatly expanded between mouse and human. In particular, a number of new untranslated exons, alternative promoters, and other noncoding regulatory elements are found in *GPR56* only in the primate linage, including humans (Bae et al. 2015). These additional noncoding elements enable *GPR56* to be expressed with more regional and temporal precision in humans and this increased repertoire of expression may drive the greater complexity and capabilities of the human cortex.

Discovery and Analysis of Human Accelerated Regions

There is a class of noncoding DNA elements defined by human-specific evolution, called human accelerated regions (HARs). HARs are regions of DNA that are highly conserved in most mammals but which show strong, specific sequence divergence in humans (Pollard et al. 2006a). The genetic differences in HARs in humans suggest that HARs are under recent evolutionary selection (thus, the name "accelerated") and that they represent essential functional sequences whose precise function may have changed between nonhumans and humans. Evidence suggests HARs have varied functions, including expression of RNA (Pollard et al. 2006b) and as transcriptional enhancers through physical interaction with promoter DNA (Capra et al. 2013). Epigenetic signatures suggest ~30% of HARs are active during embryonic development in the limbs, heart, and brain (Capra et al. 2013).

Recent work examining the genetic causes of developmental disorders has found that HARs are essential for normal brain development. Enrichment analysis suggests HARs may have roles in neurologic disease. HARs are enriched near haploinsufficient genes, raising the possibility that gene expression changes resulting from mutations within HARs may cause disease. Contributing to this hypothesis is a recent study, which reported that several HARs are within linkage regions for schizophrenia identified by genome-wide association studies (Xu et al. 2015). Of particular interest in neurodevelopment, HARs are enriched near genes associated with autism spectrum disorder (ASD), suggesting that they may play a role in ASD pathogenesis (Doan et al.

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

2016). Indeed, single nucleotide variants found in HARs in ASD patients highlight the importance of noncoding mutations in neurodevelopmental disease. Doan et al. studied a cohort of ASD patients with no coding mutation or copy number variant identified through whole genome sequencing and examined the HAR sequences in each patient. Compared to neurologically normal controls, they showed that in ASD patients there was a significant enrichment of biallelic point mutations in HARs. Several mutated HARs were shown to interact with brain-expressed genes, in particular *MEF2C*, *CUX1*, *TMEM161B*, *PTBP2*, *GPC4*, *CDKL5*, *USP32*, and *DAB2* (Doan et al. 2016). These HARs have predicted enhancer activity, and transgenic expression of the mutated HAR in developing mouse embryos showed changes in expression of the target gene.

In one example, a homozygous point mutation was identified in HAR426, located 200 kb upstream of the *CUX1* promoter (Figure 4.4). This mutation was

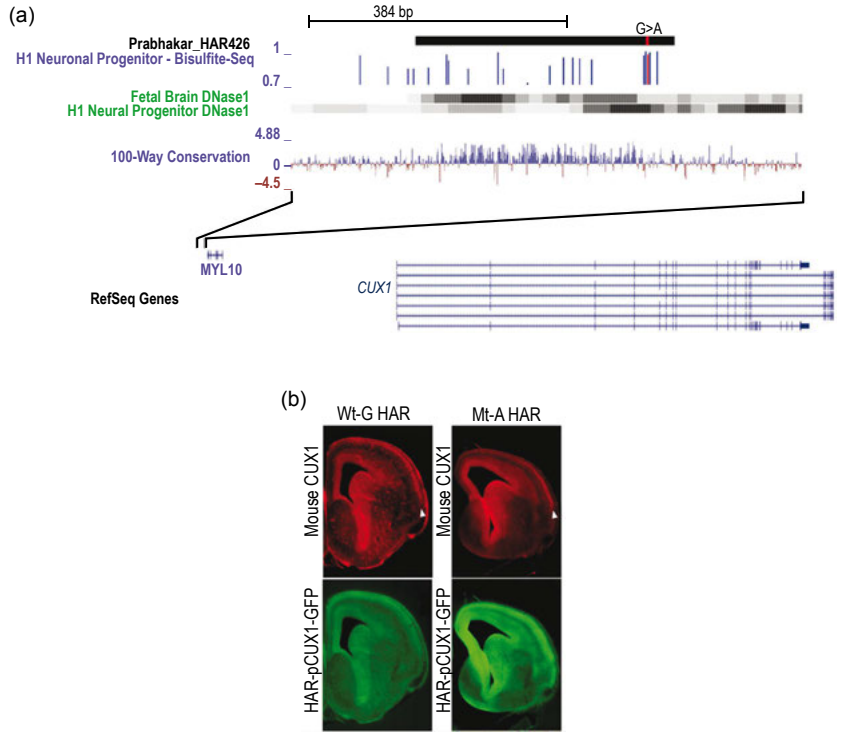

Figure 4.4 Human mutation in autism spectrum disorder (ASD) in HAR426 alters *CUX1* expression: (a) Annotated HAR426 (black bar) with G>A point mutation found in a patient with ASD. This HAR has multiple bisulfite peaks in DNA from neuronal progenitors and shows DNase1 activity in fetal brain and progenitors. This HAR sequence is highly conserved among 100 species. Bottom panel shows that this HAR is upstream of *CUX1*. (b) Wild-type (Wt) and mutant HAR426 driving green fluorescent protein (GFP) expression in developing mouse cortex shows increased transcriptional activity of mutant HAR426. Adapted from Doan et al. (2016).

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

found in three individuals from two families with ASD and intellectual disability. The mutation was predicted to create a new transcription factor binding site, and expression of the mutated HAR in mouse embryos showed increased expression compared to expression of the wild-type HAR. Overexpression of *Cux1* in cultured cortical neurons showed increased spine density (Cubelos et al. 2010), suggesting that HAR426 mutation may interfere with normal spine refinement (Doan et al. 2016).

In a second example, a 5-bp homozygous insertion/deletion in HAR169 near *PTBP2* was identified in two brothers with ASD and intellectual disabilities. *PTBP2* encodes a brain-specific splicing protein that regulates neuron differentiation (Licatalosi et al. 2012; Li et al. 2014). Chromatin interaction measurement showed that HAR169 binds the *PTBP2* promoter. The patient mutation is predicted to disrupt TF binding: a luciferase expression assay showed that the mutation decreased the enhancer activity of HAR169 by 50% in cells in a neuron-like state, and an expression analysis showed it decreased enhancer activity by 40% in primary mouse neurospheres (Doan et al. 2016). Together, HAR mutations in ASD patients illustrate how noncoding mutations can disrupt normal brain development by altering expression of neuronal genes, and, more importantly, how the essential functions of these evolutionarily important sequences can be analyzed through larger-scale application of human genetics of neurological disorders.

Decreased Sequencing Cost Will Accelerate Identification of Noncoding Mutations

The previous three examples of noncoding mutations causing neurodevelopmental disease were reported over the last ten years (in 2008, 2014, and 2016, respectively) and illustrate the power of recent advances in DNA sequencing technology. The amount of DNA sequenced increased across these studies, from targeted sequencing of a 1,000 bp enhancer in 2008, to a collection of 38 noncoding elements totaling 5,000 bp in 2014, to extraction of HAR sequences from 3,000,000,000 bp of whole genome sequencing in 2016. In parallel, across this same time span, the cost of sequencing a single human genome decreased 1,000-fold, according to the National Human Genome Research Institute: from $1,300,000 ($15/Mb) in 2008 to about $1,000 ($0.01/Mb) in 2017. These studies illustrate how the decreasing cost of DNA sequencing, through introduction of new technology for high-throughput sequencing, has enabled ever more complete examination of noncoding DNA, which comprises 98% of the human genome. Over the next few years, sequencing costs are predicted to continue to fall, which will make it feasible to perform whole genome sequencing on an increasing number of patients. This, in turn, will further expand our ability to identify patients with mutations in noncoding DNA and add to the rich diversity of genetic causes of neurodevelopmental disease. In addition, greater

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3

understanding of noncoding variants in disease will increase our understanding of noncoding DNA in normal brain development. Noncoding DNA includes gene enhancers and repressors, noncoding RNAs, micro RNAs, inserted retrotransposons, and additional regulatory sequences; each of these categories has increasingly appreciated functions in brain development.

Conclusion

In this chapter we have presented three lessons that can be learned by studying the genetic causes of cortical development that has gone awry. *CHMP1A* mutation highlighted the function of EVs and specialized cellular communication in cortical development and adult function. Microcephaly in *Aspm* KO ferrets illustrated the active role that evolution plays in cortical development and the advantages of higher-order model organisms. Finally, noncoding mutations in *GPR56* and HARs demonstrated new appreciation for genetic regulation of cortical development beyond protein-coding genes. These three discoveries were made possible by recent advances in DNA sequencing technology, and we hope they will raise new questions that drive us to continue advancing our understanding of the cerebral cortex in the years to come.

Acknowledgments

M. E. C. was supported by F30 MH102909, a Howard Hughes Medical Institute Medical Student Fellowship, and a Nancy Lurie Marks Family Foundation Medical Student Fellowship. C. A. W. was supported by R01-NS35129 and R01-NS032457 from the NINDS, U01 MH106883 from the NIMH, and the Allen Discovery Center Program through the Paul G. Allen Frontiers Group. C. A. W. is an Investigator of the Howard Hughes Medical Institute.

From "The Neocortex," edited by W. Singer, T. J. Sejnowski and P. Rakic.
Strüngmann Forum Reports, vol. 27, J. R. Lupp, series editor.
Cambridge, MA: MIT Press. ISBN 978-0-262-04324-3