Role of Mutations in clustered Protocadherin genes in the altered neurocircuitry of Autism Spectrum Disorder

By:

Lars ten Brinke

S4196112

Bachelor Thesis

Faculty of Science and Engineering

University of Groningen

Supervised by:

Prof. Dr. U.L.M. Eisel

Dr. L. M. Jorna

Preface

During my time as a bachelor student, I once attended a lecture that proposed CLARITY as a revolutionary tool to change the future of neuroscience by visualizing the brain in ways not previously possible. CLARITY enables us to study the 3D brain structure in immense detail in which neurons can be made visible along with all their dendrites and neurites, without any disruptions of its complete structure. One striking example which was presented during that specific lecture was the visualization of the brain of an autistic patient. This revealed the complete neurocircuitry of the autistic brain including the connections between neurites, synapses and more. Upon closer examination of the autistic brain, some odd ladder-like structures could be observed. This suggests that something went wrong in the wiring of the brain of the ASD-symptomatic individual.

This somehow stuck with me for a very long time, as I was diagnosed with ASD when I was a child. I have always wondered how my brain would differ from others ever since. It was this lecture that inspired me to delve deeper on how neurons would form connections with one another and how aberrations in these patterns could lead to the development of ASD. I remember emailing lecturers and studying the available literature about how deficits in neurocircuitry could play a role in the development of ASD.

Approximately one year later, I found a mechanism in the literature that proposed to facilitate cell-cell recognition between neurons by so-called protocadherins. These protocadherins are thought to engage in neural recognition models by establishing a large variety of different neuronal identities. This further inspired me to write my bachelor's thesis regarding these protocadherins and their role in the development of ASD.

Not only did writing this thesis grant me new insights on autism, but I also discovered more about myself as a person. I hope with my whole heart that you, as a reader, will uncover new insights whilst reading this thesis and that the complexity of neurocircuitry will inspire you, just as it did for me.

I would like to especially thank Professor Eisel, for he has supported me writing this thesis, helping me grasp the complex mechanism of neurocircuitry establishment in the development of ASD and for his support on this specific topic.

Summary

The human brain is composed by a complex network with billions of neurons, each forming synaptic connections with high specificity to ensure correct signal transduction. Deficits in this high specificity and neurocircuitry are frequently associated with several neurodevelopmental disorders, including autism spectrum disorder (ASD). ASD has been a focal point of research and knowledge regarding its development and knowing how deficits in neurocircuitry lead to ASD could potentially pave the way for new therapeutic strategies.

In a recent CLARITY study, which visualized the brain of an ASD patient in great detail, ladder-like structures were observed in the neurites in which the neurites formed connections with themselves, indicating that the mechanism ensuring correct cell-cell recognition is impaired in ASD. Correct cell-cell recognition is normally mediated by both clustered (cPCDHs) and non-clustered (ncPCDHs) protocadherins which establish unique neural identities and engage in cell-cell interactions to either promote or terminate binding between neurons.

This study aimed to review the exact functioning of these PCDHs in establishing such identities and cell-cell recognition, as well as how mutations within these PCDHs or their regulators could contribute to the abnormal neurocircuitry found in ASD. Indeed, we found that some mutations, especially in epigenetic regulation were associated with ASD etiology. However, the exact mode of how this is mediated is not entirely understood and requires further research.

Keywords: Autism Spectrum Disorder (ASD), Cadherins (CDHs), Clustered Protocadherins (cPCDHs), non-clustered Protocadherins (ncPCDHs), iso-neural avoidance, CLARITY, Epigenetic Regulation, Mutations

Table of Contents

Introduction

The human brain is composed of billions of neurons, contributing to a network of numerous branches or neurites forming trillions of synaptic connections (Wu & Jia, 2020) (Flaherty & Maniatis, 2020). These highly branched neurites engage in synaptic connections by extending and innervating large fields, whilst still maintaining high specificity (Südhof, 2017). In these circuits, correct assembly of dendrites and axons between neurons is deemed crucial in appropriate cognitive functioning. Impairments in this assembly are frequently reported to be associated with several neurodevelopmental disorders, such as epilepsy, schizophrenia (SZ), Fragile-X syndrome (FXS), bipolar disorder (BD), autism spectrum disorder (ASD) and many more (Harrison et al., 2020) (Mancini et al., 2020) (Hirabayashi & Yagi., 2013).

Specifically, ASD is a neurodevelopmental disorder characterized by restricted and abnormal social interactions, as well as restricted interests, repetitive behaviors, irritability, hyperactivity and attention deficits (Tsai & Huber, 2017) (NIMH, 2023). It severely impacts daily functioning in affected individuals and can prove to make life very challenging. To this date, there exists no complete pharmaceutical cure for ASD and intervention is highly based on therapies targeting self-awareness and behavioral control. This is proven to be highly beneficial when ASD is diagnosed early in life and if therapy starts at a relatively young age (12-18 months).

Nowadays, diagnosis is mainly based on clinical evaluation and on assessment of social and communicative behavior. Conversely, it has no universal standardization leading many undiagnosed individuals (Chen & Maniatis, 2013). This has negative implications on education, socioeconomic status, and mental health of those affected. Increasing the efficiency and specificity of a clinical diagnosis therefore could lead to substantially increased diagnosis and an earlier start of therapy, improving the life quality of those affected.

Although several theories have been proposed to unfold the complex development of ASD, the exact mechanism of its development remains unknown (NIMH, 2023) (Girault & Piven, 2020). It presumably involves both genetic and non-genetic factors and is considered to be a multifactorial neurodevelopmental disorder (Tsai & Huber, 2017) (Kumar et al., 2019). This consideration subsequently resulted in the broadening and revision of the diagnostic criteria for autism in the DSM-V, where ASD became an umbrella term for different forms of autism including Asperger syndrome, Pervasive Developmental Disorder (PDD) and childhood disintegrative disorder (Takumi et al., 2020).

Aside from this umbrella term, ASD patients show key characteristics in which the impaired social interactions and restricted or repetitive behaviors are likely to originate in deficits of neurocircuitry (Kumar et al., 2019). However, the exact connection between (non-)genetic factors and deficits in this circuitry leading to abnormal behavior is not entirely understood, resulting in huge limitations in the development of therapeutic strategies (Takumi et al., 2020).

Interestingly the prevalence of ASD is a 4-fold higher in males opposed to females and can primarily be determined by genetic factors (>50%) (Weiss et al., 2009) (Tsai & Huber, 2017) (NIMH, 2023). This puts emphasis on the necessity to study the impacted mechanisms underlying the formation between these neural connections and its genetic organization (Wu & Jia, 2020) (Mancini et al., 2020).

Research on these mechanisms has shed light upon several candidates of the cadherin superfamily, which play crucial roles within neurocircuit assembly and is known to mediate a wide variety of processes of Ca^{2+} dependent cell adhesion and cell-cell recognition through transmembrane interactions of extracellular cadherin repeats (ECs) (Harrison et al., 2020) (Mancini et al., 2020) (Tsai & Huber, 2017). The superfamily can be further subdivided into classical cadherins (type I and II) and protocadherins (PCDHs) (Wu & Jia, 2020). Of which the latter are considered to be vital players in

ensuring correct cell-cell recognition and iso-neural (self) avoidance, by providing unique cell-specific identities through an enormous diversity of expression patterns (Flaherty & Maniatis, 2020).

In a recent study, immunohistological visualization was performed on the brain of an autistic patient through a CLARITY analysis shown in Figure 1 (Chung, et al., 2013). Unlike other visualization methods, CLARITY preserves the continuity of the brain structure, which allows both the tracing of neurites as well as information about the three-dimensional structures of neurons. In this analysis, CLARITY revealed ladder-like structures between neurites of the same neuron (iso-neural) and between other neurons (hetero-neural), especially in the deep layers of the brain.

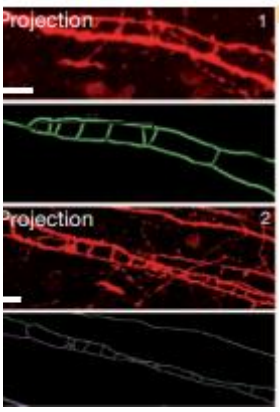

Figure 1: ladder-like structures found in the neurocircuitry of an ASD patient in a CLARITY study (Chung, et al., 2013).

These ladder-like structures were found to be associated with mutations in protocadherin genes, indicating that this cadherin subfamily may play a role in establishing proper neurocircuitry in preventing the formation of iso-neural synapses or autapses (Chung, et al., 2013) (Phillips et al., 2017). How this exactly is mediated however, is not exactly understood and the identification of these mutations and their corresponding role in establishing neurocircuitry could pave the way for developing new therapeutic strategies (Wu & Jia, 2020) (Flaherty & Maniatis, 2020). Therefore, the following question was posed:

"How and to what extent do mutations in genes encoding clustered protocadherins (cPCDHs) affect neurocircuitry, specifically in the neurodevelopment of Autism Spectrum Disorder (ASD)?"

To answer this question, we must first explore the current views and hypotheses on the development of ASD. Next, we must understand the structural and genomic organization of PCDHs and how this differs from the classical cadherins. Subsequently, we have to investigate the role of protocadherins in establishing proper neurocircuitry, specifically in cell-cell recognition and iso-neural avoidance. Furthermore, we have to focus on the intracellular cascade taking place in neurons engaging in PCDH signaling. Finally, we aim to integrate how mutations may affect the development of ASD by such mechanisms and how this may prove to be helpful in establishing diagnostic tools.

ASD is associated with a multitude of factors that could contribute to its development

As stated in the introduction, ASD is considered to be a multifactorial neurodevelopmental disorder in which multiple different pathways could contribute to its development (Kumar et al., 2019). Several of these mechanisms are proposed to mediate the development of ASD, but all involve the neurocircuitry in one way or another. Some studies suggest that an increase or decrease in the number of these connections may play a role in its development, while others propose that this problem lies within abnormal neural migration or patterning (Watts, 2018).

The latter involves a dysregulation of key signaling pathways which are thought to be crucial in neural differentiation, migration, development, brain region organization and patterning. These specific mechanisms include WNT, bone morphogenic protein (BMP), retinoic acid (RA) and sonic hedgehog (SHH) signaling. Mutations within these signaling mechanisms may in turn cause perturbations in proper neurodevelopmental and post-neurodevelopmental processes resulting in ASD (Kumar et al., 2019).

Another hypothesis states that an imbalance between excitatory and inhibitory synapses could also act as a potential mechanism that could contribute to ASD, which mainly include GABA and glutamate receptors, as well as an altered calcium signaling (Watts, 2018). This dysregulation may then ultimately lead to the development of autistic behavior, as proper establishment of neurocircuitry is deemed crucial for exhibiting an ASD-like phenotype.

Furthermore, some studies also reported that alterations in the immune system might play a large role in the development of ASD (Meltzer & Van de Water, 2016). This includes changes in the gestational environment of the maternal womb upon activation of the immune system of the mother, where cytokines in a limited degree may cross the placenta. Moreover, the production of anti-brain autoantibodies by the maternal immune system can likewise cross the placenta, inducing auto-inflammation within the fetal brain leading to an altered neurodevelopment. Additionally, congenital infection of certain types of viruses (e.g. Rubella) and bacteria also correlated with an increased prevalence of ASD. Some studies suggest that this correlation can be explained by the use of certain anti-fever medications (e.g. Advil, Tylenol or Nyquil) that affect methylation patters, whereas other studies imply that the infection itself can have consequences on the neurodevelopment of the fetal brain via epigenetic regulation.

In addition to these prenatal conditions, alterations in the postnatal immune system could also been seen as a major player, as immune dysregulation is frequently found in ASD patients (Meltzer & Van de Water, 2016). Most of these dysregulations were found in natural killer (NK) cell populations and activity patterns, as well as imbalances in $CD4^+$ and $CD8^+$ T-cells. However, the exact connection with autism and this immune dysregulation is not entirely understood.

In total, more than 800 genes were identified in genome-wide studies in the etiology of ASD, including genes involved in critical processes in the development of synapses, including chromatin remodeling, transcriptional and translational control, and overall synapse functioning (Tsai & Huber, 2017) (De Rubeis et al., 2014). A subset of these mutations can be ascribed to mutations in PCDHs and their regulators, as several human genome sequencing studies reported a higher prevalence of ASD upon identification of mutations across the 5q31 loci encoding for cPCDHs (Flaherty & Maniatis, 2020) (Ing-Esteves et al., 2018) (Hirabayashi & Yagi., 2013) (Lefebvre et al., 2012) (Peek et al., 2017) (Jia & Wu, 2020).

These findings initiated extensive research on PCDH genes in association with ASD (Flaherty & Maniatis, 2020). As a result, numerous mutations in both clustered and non-clustered PCDHs and its related

transcriptional enhancers have been found ever since (Hirabayashi & Yagi., 2013). But how these mutations contribute to its development is not exactly understood and we will therefore aim to explain how mutations could affect the mechanism of cell-cell recognition in this study. Due to the multitude of different factors, however, it is important to keep in mind that protocadherins do not fully explain the development of ASD and that it is likely that other factors might interfere with one another in which the interplay could also have a profound role on the etiology of ASD (Kumar et al., 2019) (Takumi et al., 2020).

History of exploring neural cell identities

In 1940, it was hypothesized that neurons express individual identification tags on their plasma membranes that specified synaptic connections through the so-called chemo-affinity hypothesis (Wu & Jia, 2020). Although the exact nature and mechanisms of these identification tags remained to be an unsolved puzzle for a long time, extensive efforts have been made ever since to uncover the key players in neuronal circuit assembly.

This led to the discovery that in *Drosophila melanogaster,* neural cell identities were established to discriminate self from non-self through combinational expression of different isoforms of the Dscam1 (Down syndrome cell adhesion molecule 1) molecule generated through alternative splicing, resulting in an unique repertoire of different neural identities (Wu & Jia, 2020) (Thu et al., 2014). These Dscam1 isoforms are expressed on the cell membrane and are known to engage in trans interactions of Dscam1 isoforms of other cells. Whenever these isoforms come into contact, they will repel one another if they are identical. However, the Dscam1 isoforms will not repel one another when 2 neurites connect from different neurons. Whenever this mechanism of Dscam1 contains mutations, the Drosophila nervous system revealed the same ladder-like structures as in the CLARITY analysis of the autistic brain (Chung, et al., 2013).

In vertebrates however, neural cell identities are thought to be established through combinational stochastic expression protocadherins of their α, β and γ isoforms (Wu & Jia, 2020) (Flaherty & Maniatis, 2020) (Hirabayashi & Yagi., 2013). These isoforms are encoded by gene clusters containing multiple unusually large variable exons, which are subject to alternative splicing (Harrison et al., 2020) (Mancini et al., 2020). Each cluster along with its exons encodes for variable extracellular, transmembrane and cytosolic domains. In total, there are 15 PCDH-α, 16 PCDH-β and 22 PCDH-γ different genes which make up for the enormous diversity of combinatorial expression. To understand how these PCDHs work and how they may play a role in the development of ASD, we must first explore their structure and how this structure is involved in the establishment of neural cell identities.

Structural and genomic organization of protocadherins is fundamentally different from classical cadherins

Protocadherins account for the largest subgroup of the cadherin (CDH) superfamily, which consists of classical cadherins (type I and II), desmosomal cadherins and protocadherins (Mancini et al., 2020). All members of the CDH superfamily are single-pass transmembrane proteins containing extracellular cadherins (EC) domains at the N-terminus as well as a cytoplasmic domain at the C-terminus (Figure A (Shibata-Seki et al., 2020). The EC domains are highly conserved molecules which contain ~110 cadherin motif repeats, which consists of Ca^{2+} adhering sequences accounting for the proper orientation of the protein (Weis, 1995) (Tsukasaki et al., 2014) (Shapiro et al., 2009).

Each EC domain consists of 7 β-sheets in a Greek-key topology with the N- and C- termini on opposite ends, promoting efficient stacking (Shapiro et al., 2009). These domains determine homophilic binding specificity and bind by incorporating Ca^{2+} at the Ca^{2+} adhering sequences between the different ECs (Mancini et al., 2020) (Shibata-Seki et al., 2020). The domains are numbered according to the ordered position from the N-terminus (e.g. EC1 is the closest domain to the N-terminus), where especially the EC1 domains engage in establishing cellular contact by interchanging β sheets in the classical CDHs (Harrison et al., 2020). The additional cytoplasmic domain links the cadherin to the actin cytoskeleton by having binding sites for p120 and β-catenin and this linkage is essential in controlling cell adhesion and avoidance in classical and desmosomal cadherins.

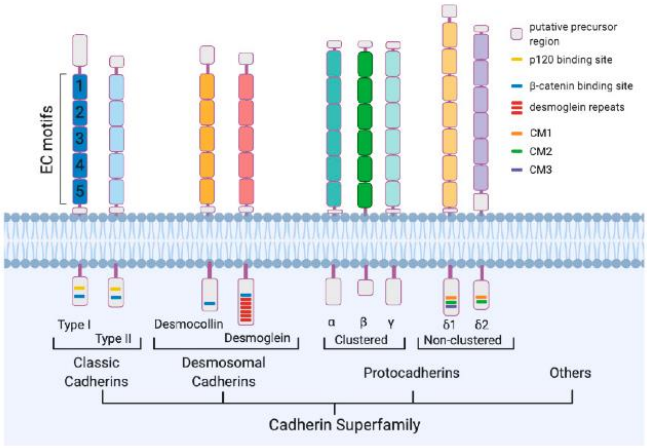

Figure 2: protein structures of members of the cadherin superfamily. Classical cadherins (type I and II) both contain 5 EC motifs in addition to a cytoplasmic domain with p120- and β-catenin binding sites but show variations in their putative precursor regions. Like classical cadherins, desmosomal cadherins also contain 5 EC motifs, but have varying cytoplasmic domains. Protocadherins on the other hand, are structurally very different from classical by having 6 or more EC motifs and structurally different cytoplasmic domains. They can be subdivided into clustered (α, β and γ) and non-clustered (δ1 and δ2) protocadherins.

Interestingly, genome-wide studies revealed some single nucleotide polymorphisms (SNPs) and other mutations within the genes encoding for these classical CDHs that correlate with ASD, affecting the expression of CDH5, CDH9, CDH10, CDH11 and CDH15 (Hawi et al., 2017). This again puts emphasis on ASDs multifactorial character and indicates that regular cell-cell adhesion could also potentially play a role in the development of ASD.

Although there is a high similarity between the different members of the superfamily, there are some key differences giving rise to the enormous functional diversity between them (Mancini et al., 2020). One striking difference among these members is the number of ECs. The classical (type I and II) and desmosomal cadherins contain 5 EC domains, whereas protocadherins have 6, 7 or even more EC domains. Furthermore, the cytoplasmic domain is structurally and functionally different in protocadherins as these lack catenin binding sites for p120 and β-catenin (Pancho et al., 2020). Lastly, PCDHs do not engage in strand-swapping in adhering to one another and can only engage in weak interactions, suggesting that its functionally very different from classical cadherins (Shapiro et al., 2009).

Even more intriguing is that there also appears to be also a high amount of variety within the protocadherin subgroup, since protocadherins can be categorized into 3 subgroups of clustered PCDHs

(α, β and γ), non-clustered PCDHs (ncPCDHs), also referred to as δ-PCDHs (δ1 and δ2) and solitary εPCDHs, which are a subgroup of ncPCDHs containing a higher or lower number of ECs compared to δPCDHs (Flaherty & Maniatis, 2020) (Kim et al., 2010) (Hirabayashi & Yagi., 2013) (Pancho et al., 2020). These PCDH clusters can be subsequently subdivided into different isoforms, as shown in table 1 and vary in their number of ECs and in their cytoplasmic domains. These different isoforms are stochastically expressed on the neural membranes by complex genetic organization, explained in the next section. Additionally, two other atypical PCDH subgroups have been identified, including seven-transmembrane PCDHs and giant fat PCDHs, but will not be further reviewed as there is no association found between these types of PCDHs and ASD (Mancini et al., 2020).

Table 1: Different Protocadherin subgroups along with their corresponding isoforms (Harrison et al., 2020)(Kim et al., 2010) (Mancini et al., 2020) (Kim et al., 2011b).

Protocadherin subgroups	Isoforms	Total
αPCDHs	PCDHα1-PCDHα13, PCDHαC1, PCDHαC2	15
βPCDHs	PCDHβ1-PCDHβ16	16
γPCDHs	PCDHγA1- PCDHγA12, PCDHγB1- PCDHγB7, PCDHγC3-PCDHγC5	22
δ1-PCDHs	δ1-PCDHs: PCDH1, PCDH7 PCDH9, PCDH11, PCDH12 (alternative: NO CM motifs)	5
δ2-PCDHs	δ2-PCDHs: PCDH8, PCDH10, PCDH17, PCDH18 and PCDH19, PCDH20 (alternative: NO CM motifs)	6
εPCDHs	PCDH12, PCDH15, PCDH20, PCDH21	4

Clustered PCDHs establish a 'Zip-code'-like neural cell identity through complex regulation of gene expression which is susceptible to mutations

The genomic organization of PCDHs genes is most peculiar and shows high similarity to that of the immunoglobin and T-cell receptor genes (Wu & Jia, 2020) (Chen & Maniatis, 2013). As the name implies, the genomic organization of clustered PCDHs (cPCDHs) is clustered on the chromosomal region 5q31 on clusters α, β and γ respectively (Mancini et al., 2020) (Hirabayashi & Yagi., 2013). These clusters are encoded by unusually large and variable exons encoding for different ECs and transmembrane domains, which show high similarity between one another (Flaherty & Maniatis, 2020). Non-clustered δ-PCDH (ncPCDHs) genes, on the other hand, are scattered across the genome. Even though ncPCDHs are produced through alternative splicing, they do not contain variable exon domains, which results in having relatively small variations between their isoforms (Harrison et al., 2020).

The variable exons of the cPCDHs can be further subdivided into alternate and constant (C-type) exons, in which the constant exons are present in every neuron and encode for the cytoplasmic domain (Wu & Jia, 2020) (Chen & Maniatis, 2013). Three of these constant (C-type) exons are located downstream of the multiple variable exons of the αPCDH and γPCDH gene clusters. Conversely, the βPCDH cluster does not encode constant exon regions encoding for a cytoplasmic domain and therefore it is thought that βPCDHs play a functionally different role opposed to the other cPCDHs (Mancini et al., 2020).

Each exon of a cluster precedes with its own promotor, which enables them to generate a huge variety of isoforms (Flaherty & Maniatis, 2020). In cPCDHs, stochastic expression of different isoforms is achieved by regulating promotor choice through 3 different mechanisms including long-range DNA looping between individual promotors, methylation of CpG sites and through transcriptional enhancer DNase I hypersensitive sites (HS) in αPCDHs (Mancini et al., 2020) (Chen & Maniatis, 2013) (Guo et al., 2012).

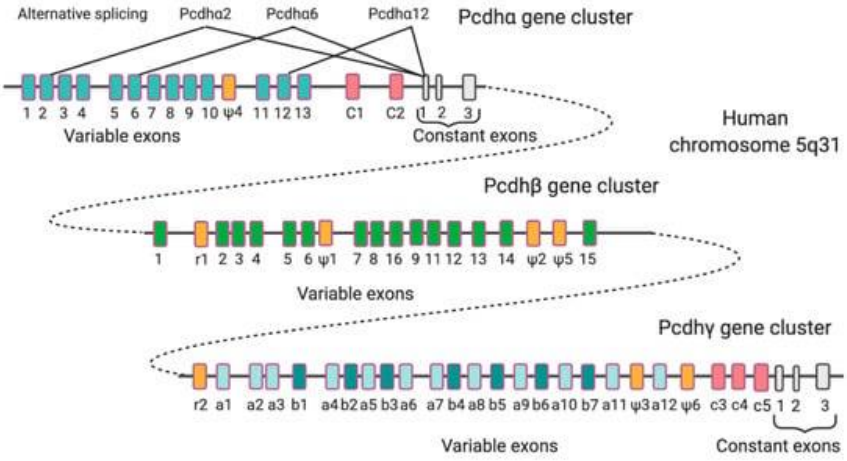

Figure 3: Genomic organization of clustered protocadherins. The α-, β- and γPCDHs are arranged in tandem on chromosome 5q31 in which both variable exons encoding for the variety of EC domains and constant exons encoding for the cytoplasmic tail are shown. Note how only the α- and γPCDH contain constant exons that encode for the cytoplasmic tail. The orange sequences are either relic or pseudogenes. In addition, an example is shown of how the PCDHα2, PCDHα6 and PCDHα12 isoforms are expressed (Mancini et al., 2020).

These transcriptional enhancers are required for efficient transcription of PCDH expression. They contain CCCTC binding factor (CTCF) binding sites can recruit the cohesin complex (Chen & Maniatis, 2013) (Guo et al., 2012). Recruitment of cohesin by CTCF, which is also present in the variable exons, facilitates conformational changes in chromatin to form DNA loops which promotes a direct interaction between the transcriptional enhancers and exon promotors by binding to a conserved sequence element (CSE) upstream to all promotor regions and the CTCF within each exon (Hirabayashi & Yagi., 2013). Two of these transcriptional enhancers were found in αPCDHs, namely HS-7, located in between the last 2 constant exons and HS5-1, located downstream the very last constant exon. Loss of these HS elements in the αPCDH cluster in knockout mice, result in a strong downregulation of the αPCDHs and mutations within these elements are associated with changes in neurocircuitry found in bipolar disorder.

β- and γPCDH gene clusters, however, do not contain these CTCF binding sites within their exons, and rely more on the CTCF binding sites within the CSEs (Chen & Maniatis, 2013) (Guo et al., 2012). Methylation of CpG sites within these CSEs decreases the binding of the cohesin complex, leading to reduced promotor activation. Additionally, methylation of promotor regions of PCDHs themselves was also shown to reduce promotor activity. Methylation of these regions occurs through the Methyl-CpG-binding Protein 2 (MCP2) (Hirabayashi & Yagi., 2013). It does so by incorporating 5-methylcytosine groups at CpG sites. These mechanisms of methylation by epigenetic regulation have thus a large impact on promotor choice and lead to a mosaic of both hypermethylated (silenced) and hypomethylated (expressed) isoforms. Whereas constant exons are always hypomethylated in α- and γPCDH gene clusters (Mountoufaris et al, 2018). This form of regulation is critical for correct PCDHs expression and perturbations in these regulators have been frequently associated with ASD, as will be explained in the further sections (Jia & Wu, 2020).

Eventually, each neuron expresses 2 PCDH-α genes, 4 PCDH-β genes and 4 PCDH-γ genes in a mono-allelically and stochastical manner, in addition to the expression of all 5 C-type PCDHs genes (PCDHαC1, PCDHαC2 and PCDHγC3- PCDHγC5) (Wu & Jia, 2020) (Flaherty & Maniatis, 2020) (Harrison et al., 2020) (Hirabayashi & Yagi., 2013) (Phillips et al., 2017) (Pancho et al., 2020). Forming a multimeric complex of αPCDH, βPCDH and γPCDH isoforms shown in figure 4. These isoforms can then engage to in total 1.443.381.182.464 (13^2 x 16^4 x 19^4) different combinations, resulting in a huge repertoire of unique neuronal identities. The stochastic expression of exons shows some resemblance with the generation of zip-codes, which are required to help address homes and are generated by a random combination of digits and letters. The variable exons can then be seen as the digits which generates the largest amount of variety, the letters would then be the c-type exons as they appear to be more constant. The combination between these variable and c-type exons would then generate the unique neural identity that is required to ensure that other neurons can properly address to this identity.

Dysregulation of cPCDH genes and mutations within its cluster have been frequently associated with malfunctions in neurocircuitry, including the formation of iso-neural synapses and reduced axonal branching and we will further discuss these malfunctions in the next sections (Jia & Wu, 2020).

To conclude, the regulation of the stochastical expression of these exons is orchestrated by complex forms of (epi)genetic regulation, which includes DNA methylation and long-range DNA looping of PCDHs promotors upstream of the variable exons, along with transcriptional enhancers downstream of the PCDH gene clusters (Flaherty & Maniatis, 2020) (Chen & Maniatis, 2013) (Guo et al., 2012). All these mechanisms result in neurons expressing unique, specific 'zip-code'-like identities. These Zip-code like identities can then help to address the right dendrite to the right axon, leading to a proper recognition model which will be explained in the next section.

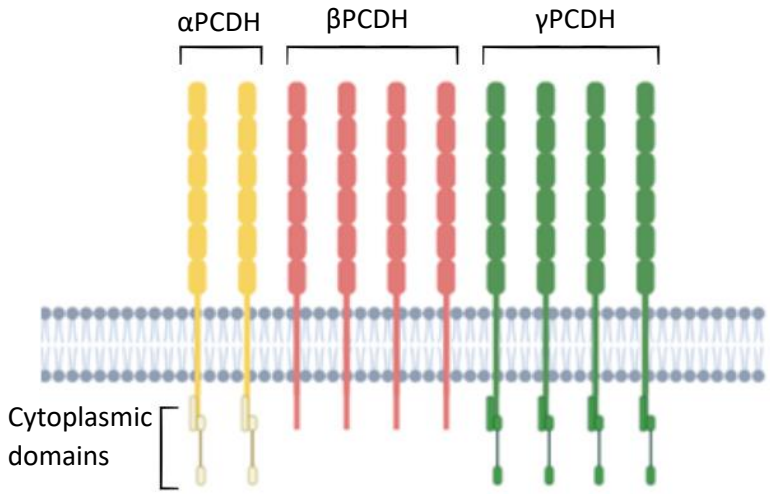

Figure 4: Schematic example of a multimeric complex consisting of 2 αPCDH, 4 βPCDH and 4 γPCDH variable isoforms which generates a huge repertoire of unique zip-code like identities. Also note that βPCDHs lack a cytoplasmic domain because of lacking the C-type exons. These multimeric complexes of PCDHs interact with one another to engage in cell-cell recognition.

Neurocircuitry assembly is mediated by both nc- and cPCDHs and mutations in these regions are associated with abnormal neurocircuitry

Proper brain functioning is largely relying on proper wiring of neurocircuitry. For specific and proper wiring to take place, the establishment and recognition of neural identities by PCDHs is essential (Wu & Jia, 2020) (Mancini et al., 2020) (Pancho et al., 2020) (Thu et al., 2014). As stated in the previous section, cPCDHs generate unique and cell-specific zip-code like neural identities through a large repertoire of combinatorial PCDH gene expression and acts as a recognition unit. However, aside from the assignment of neural identities, PCDHs also need to establish or abolish wiring by sorting out different neurons. This discrimination between self- and non-self is crucial in the proper establishment of neurocircuitry and plays a vital role during the embryonic stage and in later neurodevelopmental stages.

These PCDHs are thought to do so by probing other PCDHs on different neurites, which in turn leads to intracellular signaling via the cytoplasmic domains (Mancini et al., 2020) (Hirabayashi & Yagi., 2013) (Pancho et al., 2020). In contrast to classical cadherins in which primarily the EC1 domains bind, PCDHs engage in highly specific isoform-binding by interacting with their EC1-EC4 domains in an antiparallel orientation (Harrison et al., 2020) (Rubinstein et al., 2015) (Rubenstein et al., 2015). This interaction facilitates trans (cell-cell) dimer formation in cPCDHs, like that of Dscam1 isoforms, eventually leading to iso-neural avoidance. To enhance this specificity, correct matching of all cPCDHs is necessary to facilitate binding of the 2 isoforms, as even a single mismatch of a single cPCDH isoform can prevent dimer formation and eventually repulsion or hetero-neural crossing (Thu et al., 2014). It is this termination of binding that determines the discrimination between self- and non-self (Rubenstein et al., 2015). Hetero-neural adhesion is facilitated if PCDHs of different neurites bind (Figure A) and iso-neural avoidance is facilitated whenever PCDHs of the same neurites bind (Figure A).

It may therefore seem likely, that cPCDH dysregulation and mutations may ultimately lead to the failure of correct cell-cell matching between neurites of the same neuron, resulting in the inability to terminate binding between these neurites leading to the formation of the ladder-like structures found during the CLARITY analysis shown in figure 1 (Flaherty & Maniatis, 2020) (Jia & Wu, 2020) (Chung, et al., 2013).

Matching between PCDH isoforms also happens for ncPCDHs (Bisogni et al., 2018). However, these are thought to be more involved in the fine-tuning and modification of the adhesive properties of the cPCDHs. Additionally, the EC5 and EC6 domains of cPCDHs can also engage in heterophilic cis (same cell) formations, in contrast to ncPCDHs (Kim et al., 2010) (Mancini et al., 2020). These cis-interactions between the EC6 domains are shown to be important to enhance the cell-surface availability of the cPCDHs (Phillips et al., 2017) (Thu et al., 2014). Important to note is that PCDHs are not only involved in the processes mentioned above. PCDHs can also mediate axon-axon interactions in which a follower neuron binds to an axonal scaffold of another neuron to help it migrate into the appropriate direction in the brain (Flaherty & Maniatis, 2020). The mechanism via which iso-neural avoidance by PCDHs lead to the termination of binding is thought to be mediated through certain interactions of their cytoplasmic domains (Hirabayashi & Yagi., 2013) (Phillips et al., 2017). Subsequently leading to the

activation of intracellular cascades. The exact mechanism of how this is exactly mediated will be explained in the next section.

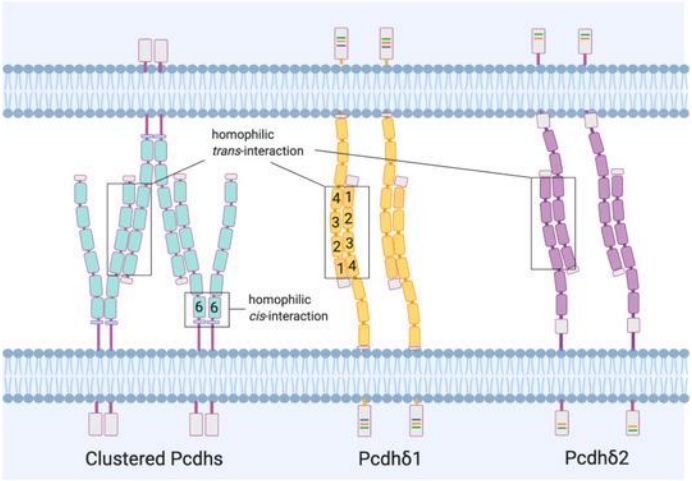

Figure 5: interactions between cells mediated by PCDHs on different neurites. Both cPCDHs and ncPCDHs adhere to one another through homophilic trans-interactions of their EC1-4 domains in an anti-parallel fashion. In addition, cPCDHs can also engage in cis-interactions via their EC6 domain, which is thought to be important in enhancing cell-surface expression (Mancini et al., 2020).

Intracellular signaling by PCDHs results in the ability to terminate or facilitate synapse formation

While the cytoplasmic domains in classical cadherins (type I and II) act directly on the cytoskeleton via its catenin binding sites which give rise to its adhesive abilities, the cytoplasmic domains of PCDHs are structurally different and engage in direct intracellular signaling pathways (Phillips et al., 2017) (Peek et al., 2017) (Jia & Wu, 2020). The cytoplasmic domains (CM1, CM2 & CM3) of both α- and γPCDHs (βPCDHs lack a cytoplasmic domain) associate with two tyrosine kinases: a proline-rich tyrosine kinase 2 (PYK2) and a focal adhesion kinase (FAK) (Hirabayashi & Yagi., 2013).

Interactions between PCDHs is necessary, as these interactions inhibit the activity of these kinases (Hirabayashi & Yagi., 2013) (Jia & Wu, 2020). Inhibition of the PYK2 pathway promotes cell survival, as an overactivation of PYK2 results in apoptosis. Inhibition of the FAK pathway results in increased branching of neurites. Therefore, it seems that the presence of both α- and γPCDHs is crucial in maintaining cell survival and promoting neurite arborization and an absence of such may promote apoptosis, preventing survival of neurons that cannot engage in cell-cell recognition by lacking these PCDHs. How these pathways exactly mediate these intracellular cascades, however, is not fully understood.

Additionally, all cPCDHs (including βPCDHs) contain highly conserved VCD motif at their cytoplasmic side, which is found to be critical of endosomal trafficking (Phillips et al., 2017). One theory suggests that the matching of PCDHs triggers endocytosis of the PCDHs via the trafficking signals of the VCD motifs, where the endocytosis might also cause internalization of other adhesion molecules into vesicles, leading to the termination of the adhesion between matching neurites. These vesicles are then transported to the late or recycling endosomal complex, where they can be either redirected to degradation in the lysosome or to be recycled to the membrane to re-engage in the PCDH recognition process.

ncPCDHs have functionally and structurally very different cytoplasmic domains. δ1-PCDHs contain 7 EC domains in addition to CM1 and CM2 motifs in their cytoplasmic tails and contains an additional CM3 protein phosphatase-1α (PP1α) domain (Harrison et al., 2020) (Kim et al., 2010) (Mancini et al., 2020) (Kim et al., 2011b). δ2-PCDHs, on the other hand, consist of 6 EC domains along with a CM1 and CM2 motif in their cytoplasmic tails, lacking the third CM3 PP1α binding domain. The function of these CM motifs and their corresponding structure, however, also remains to be investigated and it could also be very possible that mutations within these mechanisms could also play a role in the development of ASD, potentially through WNT, BMP, RA or SHH signaling, stated in the first section. However, this still would require extensive research.

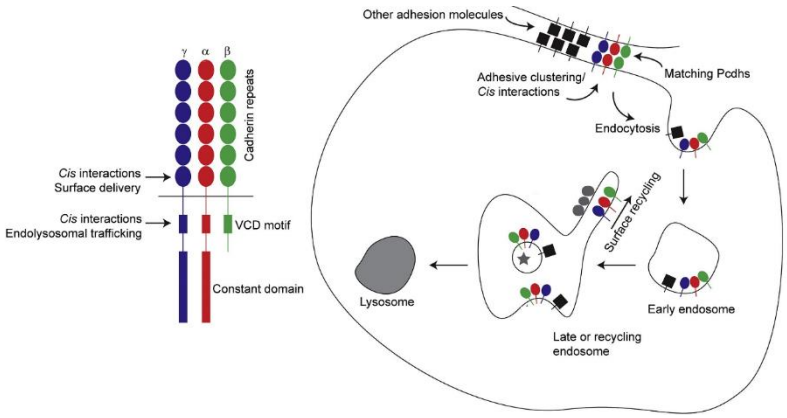

Figure 6: schematic model of cPCDHs trafficking. On the left, the protein structure of cPCDHs is shown, containing VCD motifs in each cPCDHs Whenever PCDHs of different neurites match, these VCD motifs engage in cis interactions with one another to generate a trafficking signal, which subsequently triggers endocytosis along with other adhesion molecules via early endosomes. These endosomes are transported to the late or recycling endosome, which can either redirect the compartments to the lysosome or back to the cell membrane (Phillips et al., 2017).

Mutations in PCDHs and their regulators are associated with neurocircuitry impairments and ASD

As to this date, there exist no direct evidence of a causal relationship between mutations in PCDH genes and ASD (Mountoufaris et al, 2018). Therefore, one can only speculate of how such mutations could lead to aberrant neurocircuitry evoking the ASD-like phenotype. We aim to do so by connecting the role of PCDHs in neurocircuitry establishment and the finding of the ladder-like structures found in the CLARITY analysis of an ASD-brain.

Consistent with the findings that PCDHs are involved in cell-cell recognition and iso-neural avoidance, knockout studies of both nc- and cPCDHs have revealed the relevance of these genes (Hirabayashi & Yagi., 2013) (Peek et al., 2017) (Jia & Wu, 2020). In these studies, it was shown that knockouts of the α- and βPCDHs had severe deficits in their neurocircuitry and showed alterations in their behavior. Amongst these deficits is that olfactory neurons of αPCDH mutant mice which lack the constant exon and thus the cytoplasmic domain had altered axon projections as they failed to project into a single glomerulus and instead projected into multiple small glomeruli (Figure 7C) (Peek et al., 2017). Additionally, serotonergic neurons are also impacted by a loss of PCDHαC2 as their neurites failed to form extensive axonal branches upon approaching other neurites (Figure 7D), which has also been found in some cases of ASD. Loss of other αPCDHs however did not result in the same effects, indicating that only PCDHαC2 is required for serotonergic wiring.

Mice lacking the C-type γPCDHs (PCDH-γ c3-c5) and specifically PCDH-γc4 however, were unviable due to massive apoptosis of neurons within the spinal cord (Figure 7E) (Hirabayashi & Yagi., 2013) (Peek et al., 2017) (Jia & Wu, 2020). Interestingly this massive apoptosis was not found in other types of neurons and the knockout even led to increased synapse formation in cortical neurons (Mancini et al., 2020). In other neural cell types such as Starbust amacrine cells (Figure 7A) and Purkinje cells (Figure 7B), the loss of γPCDH also coincided with a dendritic collapse in which the neurites aggregate with one another (Figure 7B).

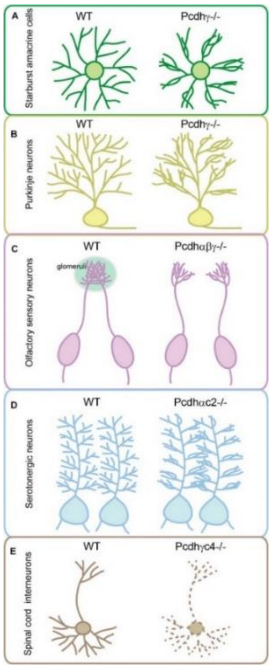

Figure 7: effect of different cPCDHs knockouts on neurites of different types of neurons. 7A shows the effect of γPCDH knockout on Starbust amacrine cells. 7B shows the effect of γPCDH knockout Purkinje cells. 7C shows the impact of either α, β or γPCDH knockouts on olfactory sensory neurons, where the neurons project their axons into multiple small glomeruli, instead of a single glomerulus. 7D shows the effect of a PCDHαC2 knockout on serotonergic neurons. 7E shows how a PCDHγC4 knockout leads to absence of neurites and apoptosis of spinal cord interneurons (Jia & Wu, 2020).

Interestingly, the wiring of neurocircuitry in these cPCDH knockouts shows high similarity with the ladder-like structures found during the CLARITY analysis of the autistic brain (Figure 7A, 7B, 7D and Figure 1) (Chung, et al., 2013). Indicating that mutations in cPCDHs indeed could be the responsible factor resulting in a loss of correct cell-cell recognition and iso-neural avoidance, ultimately leading to the development of the ladder-like structures found in ASD patients.

Failure of correct cell-cell recognition might be a result of the inability to establish proper cell-specific identities normally established by the PCDHs. As stated in the previous sections, these are generated by a complex genomic organization in which the generation of PCDH isoforms is like that of the generation zip-codes. In this generation, correct methylation of the CSE elements by MeCP2 is crucial for the stochastic expression of cPCDH isoforms by promotor choice and even small perturbations can lead to deficiencies in establishing proper neural identities that are associated with ASD (Flaherty & Maniatis, 2020) (Jia & Wu, 2020) (Mountoufaris et al, 2018).

Indeed, perturbations in the regulation of MeCP2 and mutations in its gene is the major cause of Rett's syndrome, which was included under the ASD umbrella for a long time but was considered as an individual disorder since the discovery of a mutation within the MeCP2 protein (Hirabayashi & Yagi., 2013) (Peek et al., 2017) (Shah & Bird, 2017) (Chahrour et al., 2008) (Golan-Mashiach et al., 2011).

Rett's syndrome is characterized by autism, language dysfunctioning, ataxic movements, altered growth of extremities including the head. As MeCP2 is an important regulator of PCDH promotor choice, mutations in its gene are directly associated with dysregulations of cPCHDs in which especially a massive upregulation of PCDH7 and PCDHβ1 is frequently found (Jia & Wu, 2020).Interestingly, another study has shown that the loss or increase of MeCP2 protein levels also directly correlated with changes in the number of glutamatergic synapses (Monteggia & Kavalali, 2009). Changes in the number of these glutamatergic synapses could lead to an imbalance between excitatory and inhibitory synapses, which was also one of the proposed theories of the development of ASD (Watts, 2018). It is therefore tempting to think that the dysregulation of the PCDHs could mediate the role of the altered glutamatergic synapse establishment, but further research is required to make this connection.

Mutations in other regulators that alter the chromatin structure, such as CTCF binding elements, cohesins, SETDB1 and Wiz are also associated with a wide variety of neurodevelopmental disorders (Jia & Wu, 2020). It is very possible that these disorders are a result of cPCDH dysregulation and there is growing evidence of such possibility. In CTCF knockouts for example, a downregulation of nearly all cPCDHs have been observed and analysis of the corresponding neurocircuitry revealed a reduction of axonal branching like that of the cPCDHs knockouts (Hirabayashi & Yagi., 2013). Additionally, several SNP mutations of the CTCF gene have been identified as possible risk factors for schizophrenia. Similarly, mutations in cohesins and its regulators are also associated with mental retardation and intellectual disability and a knockdown of this genes again significantly downregulate cPCDH expression.

Furthermore, a deletion of SETDB1 (SET domain bifurcated 1), which prevents excessive CTCF binding in the cPCDH locus, led to a 500-fold increase of cPCDH expression compared to the rest of the genes within the genome (Jia & Wu, 2020). Mutations in this gene are also frequently observed in schizophrenia, Huntington's disease, and ASD.

Lastly, the genomic regulator WIZ (widely interspaced zinc finger-containing protein) is also an important protein in cPCDH regulation. WIZ normally plays a role in chromatin looping and shows some overlap with the CTCF binding sites (Jia & Wu, 2020). WIZ mutant mice show dysregulations of the βPCDH cluster and display a remarkable amount of anxiety, which may likely be a result of the abnormal wiring of neurocircuitry mediated by cPCDHs.

Although it is not exactly understood how these dysregulations exactly lead to a disruption of cell-cell recognition, it is likely that it alters the repertoire of stochastical expression and therefore may alter the establishment of proper neural identities. The inability to establish proper neural identities may subsequently lead to failure of proper cell-cell recognition. and the altered neurocircuitry which could in turn lead to altered neural processing resulting in the ASD-like phenotype.

Mutations within the cPCDHs cluster itself are also closely associated with ASD (Anitha et al., 2013). In total, 5 SNPs (single nucleotide polymorphs) were discovered in the αPCDH cluster, including rs251379, rs1119032, rs17119271, rs155806 and rs17119346, in which especially rs1119042 had a strong correlation. These SNPs are all located in intron regions and are thought to play a role in splicing regulation or could act as miRNA or transcription binding sites. Additionally, an analysis of the βPCDHs revealed that mutations of PCDH-β4 (D555H) and PCDH-β15 genes are potential risk factors for developing ASD (Hirabayashi & Yagi., 2013) (Peek et al., 2017) (Jia & Wu, 2020). Again, little is known about the relationship between specific cPCDHs mutations and ASD and this thus requires further investigation.

Additionally, mutations in ncPCDHs have also been frequently found to associate with ASD, including PCDH9, PCDH10 and PCDH19 (Hirabayashi & Yagi., 2013) (Tsai & Huber, 2017) (Peek et al., 2017). Especially mutations within the PCDH10 gene have a strong correlation with ASD. PCDH10 are both highly expressed within the amygdala, which is an important region in social and communicative behaviors. PCDH10 acts on the neurocircuitry within this region through refinement or elimination of excitatory synapses via the mechanisms described in prior sections, eventually preventing neurocircuit hyperexcitability. Their exact function, however, remains to be further investigated (Kim et al., 2011).

Discussion

The etiology of Autism spectrum disorder (ASD) has been a focal point of research for decades and is currently considered to be a multifactorial neurodevelopmental disorder given the numerous uncovered mechanisms and factors (both genetic and non-genetic) associated with its development (Tsai & Huber, 2017) (NIMH, 2023) (Klein-Tasman & Mervis, 2018) (Stathopoulos et al, 2020) (Watts, 2018). Most theories suggest that abnormal types of neurocircuitry give rise to the ASD-like phenotype and include a proposed imbalance between excitatory and inhibitory synapses, abnormal neural migration and patterning as well as incorrect cell-cell recognition. Recently, the brain along with its neurocircuitry can be studied in immense detail by CLARITY, by which it is able to visualize the brain in a 3D structure without causing disruptions in the continuity of the brain (Chung, et al., 2013).

Interestingly, a study in which a CLARITY analysis was performed on post-mortem brains of ASD individuals verified the proposed theories that state that ASD coincides with deficits in their corresponding neurocircuitry (Chung, et al., 2013). These deficits included odd ladder-like structures between neurites of the same neurons. Under normal conditions, these ladder-like structures are thought to be prevented via cell-cell recognition mechanisms which makes it able for a neuron to discriminate self from non-self. The finding of these ladder-like structures therefore suggests that a failure of correct cell-cell recognition might be a key factor in the development of abnormal neurocircuitry. To verify this thought, this study aimed to review how wiring of neurocircuitry by proper cell-cell recognition is mediated by PCDHs and investigated the possibility whether mutations in these genes may contribute to the neurodevelopment of ASD (Mancini et al., 2020).

Proper cell-cell recognition in neurocircuitry establishment is primarily mediated by cPCDHs, of which stochastic combinations of their variable isoforms (2 αPCDH, 4 βPCDH and 4γPCDH) generate a large repertoire of multimeric PCDH recognition units, expressed on the cell membrane (Wu & Jia, 2020) (Flaherty & Maniatis, 2020) (Harrison et al., 2020) (Hirabayashi & Yagi., 2013) (Pancho et al., 2020). These multimeric recognition units act as zip-code like identities which help to address neurites to other neurites to facilitate synapse formation. It does so by probing other PCDH recognition units, where the EC domains strictly engage in homophilic trans-interactions, meaning that they could only engage in interactions if the recognition units are identical (Mancini et al., 2020) (Hirabayashi & Yagi., 2013). Matching of the EC domains by different PCDHs recognition units therefore indicates that the neurite binds to a neurite of the same neuron and this would lead subsequently to iso-neural avoidance via intracellular cascades. These intracellular cascades eventually triggers endocytosis of both PCDHs and other adhesion molecules to terminate the binding between neurites (Phillips et al., 2017).

The stochastic expression PCDH isoforms that account for these unique zip-code like identities is orchestrated by regulating promotor choice of exons within the cPCDH clusters (Flaherty & Maniatis, 2020) (Chen & Maniatis, 2013) (Guo et al., 2012). This Regulation happens via methylation of CpG sites within CSE elements induced by MeCP2, eventually resulting in the expression of 2 αPCDH, 4 βPCDH and 4γPCDH variable exons, in which additionally 2 constant α- and 3 γPCDH exons are also

expressed (Mountoufaris et al, 2018) (Hirabayashi & Yagi., 2013). This would conclude that epigenetic regulation is crucial in proper cPCDH expression and in turn correct neural identity establishment and many studies suggest that a dysregulation of cPCDH gene by incorrect promotor choice may be the responsible factor for incorrect cell-cell recognition leading to the failure of iso-neural avoidance (Peek et al., 2017) (Shah & Bird, 2017) (Chahrour et al., 2008) (Golan-Mashiach et al., 2011).

Indeed, we found that many mutations occurring in epigenetic regulators, such as MeCP2, CTCF binding elements, cohesins, SETDB1 and WIZ were associated with ASD. (Hirabayashi & Yagi., 2013) (Peek et al., 2017) (Shah & Bird, 2017) (Chahrour et al., 2008) (Golan-Mashiach et al., 2011). These mutations may potentially result into aberrant methylation of CSEs and CTCFs in the complex genomic organization of cPCDHs.

Although, current knowledge regarding these mechanisms and the exact genetic implications is very limited and still under investigation, we suggest that a dysregulation of cPCDH genes could reduce the pool of unique neural identities, as an upregulation or downregulation of certain cPCDH genes would lead to less stochasticity. This would ultimately lead to more neurons expressing the same set of cPCDH isoforms, resulting in increased synapse termination due to the increased matching of cPCDH recognition units from different neurons. Future research should therefore aim to study the stochasticity of PCDH isoforms and its impact on neurocircuitry in individuals with mutations in the epigenetic regulators stated above.

Additionally, we hypothesized that mutations in the cPCDH cluster itself, may result in the inability of affected isoforms to engage in the matching of the strictly homophilic trans-interactions between neurites, leading to failure of cell-cell recognition between affected PCDH recognition units and subsequently failure of activating the iso-neural avoidance mechanism. Serving as a possible explanation of how cPCDH mutations may lead to the ladder-like structures observed in the CLARITY analysis.

Indeed, we found that many mutations encoding for both cPCDHs and ncPCDHs and their corresponding regulators seemed to correlate with the development of ASD, but the exact causal relationship remains to be further investigated as there exists no direct evidence of this relationship to this date (Mountoufaris et al, 2018) (Anitha et al., 2013) (Hirabayashi & Yagi., 2013) (Peek et al., 2017). Additionally, knockout studies of cPCDH genes revealed neurocircuitry similar to that of the ladder-like structures found in the neurocircuitry of the ASD brain (Chung, et al., 2013) (Jia & Wu, 2020).

Lastly, we found that mutations within ncPCDHs genes also have a strong correlation with ASD, especially in the PCDH10 gene (Hirabayashi & Yagi., 2013) (Tsai & Huber, 2017) (Peek et al., 2017). While the role and exact function of ncPCDHs is still under debate, it is possible that they are involved in the fine-tuning of cPCDHs and are able to modify the adhesion between neurons. As PCDH10 is heavily expressed in the amygdala, it could be that mutations alter the adhesive properties between neuron and impact the neurocircuitry of emotional regulation, leading to the altered behavior found in ASD. However, this is merely a speculation and further research need to point out such possibility.

To conclude, we found that mutations in genes encoding for PCDH genes and in its regulatory elements may eventually lead to the disruption of proper cell-cell recognition and therefore the establishment of aberrant neurocircuitry (Flaherty & Maniatis, 2020) (Ing-Esteves et al., 2018) (Hirabayashi & Yagi., 2013) (Lefebvre et al., 2012) (Peek et al., 2017) (Jia & Wu, 2020). However, it should be noted that the exact function of PCDHs and how mutations could act on cell-cell recognition mechanisms is still not entirely understood and that it would require further extensive research to establish a causal

relationship between PCDH mutations and ASD development. Especially the role of ncPCDHs and how these may interact with cPCDHs is not well comprehended (Kim et al., 2011). It is likely that an interplay between these proteins and other mechanisms occurs in the development of ASD, which should be taken into consideration when studying ASD development. Further research on the role of PCDHs and how these interacts with other mechanisms to establish proper neurocircuitry and how such mutations could lead to the development of ASD is therefore necessary to develop new therapeutic strategies. That being said, understanding the role of PCDH in the complex mechanism of neurocircuitry establishment might open a window for a new approach to neurodevelopmental disorders, such as ASD.

Bibliography

Anitha, A., Thanseem, I., Nakamura, K., Yamada, K., Iwayama, Y., Toyota, T., Iwata, Y., Suzuki, K., Sugiyama, T., Tsujii, M., Yoshikawa, T., & Mori, N. (2013). Protocadherin α (PCDHA) as a novel susceptibility gene for autism. *Journal of Psychiatry & Neuroscience*, *38*(3), 192–198. https://doi.org/10.1503/jpn.120058

Bisogni, A. J., Ghazanfar, S., Williams, E., Marsh, H. M., Yang, P., & Lin, D. (2018). Tuning of delta-protocadherin adhesion through combinatorial diversity. *eLife*, *7*. https://doi.org/10.7554/elife.41050

Chahrour, M. H., Jung, S., Shaw, C. A., Zhou, X., Wong, S. T. C., Qin, J., & Zoghbi, H. Y. (2008). MeCP2, a Key Contributor to Neurological Disease, Activates and Represses Transcription. *Science*, *320*(5880), 1224–1229. https://doi.org/10.1126/science.1153252

Chen, W., & Maniatis, T. (2013). Clustered protocadherins. *Development*, *140*(16), 3297–3302. https://doi.org/10.1242/dev.090621

Chung, K., Wallace, J., Kim, S., Kalyanasundaram, S., Andalman, A. S., Davidson, T., Mirzabekov, J. J., Zalocusky, K. A., Mattis, J., Denisin, A. K., Pak, S., Bernstein, H. L., Ramakrishnan, C., Grosenick, L., Gradinaru, V., & Deisseroth, K. (2013). Structural and molecular interrogation of intact biological systems. *Nature*, *497*(7449), 332–337. https://doi.org/10.1038/nature12107

De Rubeis, S., He, X., Goldberg, A. P., Poultney, C. S., Samocha, K., Cicek, A. E., Kou, Y., Liu, L., Fromer, M., Walker, S., Singh, T., Klei, L., Kosmicki, J., Fu, S., Aleksic, B., Biscaldi, M., Bolton, P. F., Brownfeld, J. M., Cai, J., Buxbaum, J. D. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*, *515*(7526), 209–215. https://doi.org/10.1038/nature13772

Flaherty, E., & Maniatis, T. (2020). The role of clustered protocadherins in neurodevelopment and neuropsychiatric diseases. *Current Opinion in Genetics & Development*, *65*, 144–150. https://doi.org/10.1016/j.gde.2020.05.041

Girault, J. B., & Piven, J. (2020). The Neurodevelopment of Autism from Infancy Through Toddlerhood. *Neuroimaging Clinics of North America*, *30*(1), 97–114. https://doi.org/10.1016/j.nic.2019.09.009

Golan-Mashiach, M., Grunspan, M., Emmanuel, R., Gibbs-Bar, L., Dikstein, R., & Shapiro, E. (2011). Identification of CTCF as a master regulator of the clustered protocadherin genes. *Nucleic Acids Research*, *40*(8), 3378–3391. https://doi.org/10.1093/nar/gkr1260

Guo, Y., Monahan, K. D., Wu, H., Gertz, J., Varley, K. E., Li, W., Myers, R. M., Maniatis, T., & Wu, Q. (2012). CTCF/cohesin-mediated DNA looping is required for protocadherin α promoter choice. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(51), 21081–21086. https://doi.org/10.1073/pnas.1219280110

Harrison, O. J., Brasch, J., Katsamba, P. S., Ahlsen, G., Noble, A. J., Dan, H., Sampogna, R. V., Potter, C. S., Carragher, B., Honig, B., & Shapiro, L. (2020). Family-wide Structural and Biophysical Analysis of Binding Interactions among Non-clustered δ-Protocadherins. *Cell Reports*, *30*(8), 2655-2671.e7. https://doi.org/10.1016/j.celrep.2020.02.003

Hawi, Z., Tong, J., Dark, C., Yates, H., Johnson, B. P., & Bellgrove, M. A. (2017). The role of cadherin genes in five major psychiatric disorders: A literature update. *American Journal of Medical Genetics - Neuropsychiatric Genetics*, *177*(2), 168–180. https://doi.org/10.1002/ajmg.b.32592

Hirabayashi, T., & Yagi, T. (2013). Chapter 13: Protocadherins in Neurological Diseases – Cell Adhesion Molecules, *Springer* (pp. 293–314). https://doi.org/10.1007/978-1-4614-8090-7_13

Ing-Esteves, S., Kostadinov, D., Marocha, J., Sing, A. D., Joseph, K. S., Laboulaye, M. A., Sanes, J. R., & Lefebvre, J. S. (2018). Combinatorial Effects of Alpha- and Gamma-Protocadherins on Neuronal Survival and Dendritic Self-Avoidance. *The Journal of Neuroscience*, *38*(11), 2713–2729. https://doi.org/10.1523/jneurosci.3035-17.2018

Jia, Z., & Wu, Q. (2020). Clustered Protocadherins Emerge as Novel Susceptibility Loci for Mental Disorders. *Frontiers in Neuroscience*, *14*. https://doi.org/10.3389/fnins.2020.587819

Kim, S. H., Yasuda, S., Tanaka, H., Yamagata, K., & Kim, H. J. (2011b). Non-clustered protocadherin. *Cell Adhesion & Migration*, *5*(2), 97–105. https://doi.org/10.4161/cam.5.2.14374

Klein-Tasman, B. P., & Mervis, C. B. (2018). Autism Spectrum Symptomatology Among Children with Duplication 7q11.23 Syndrome. *Journal of Autism and Developmental Disorders*, *48*(6), 1982–1994. https://doi.org/10.1007/s10803-017-3439-z

Kumar, S., Reynolds, K., Ji, Y., Gu, R., Rai, S., & Zhou, C. J. (2019). Impaired neurodevelopmental pathways in autism spectrum disorder: a review of signaling mechanisms and crosstalk. *Journal of Neurodevelopmental Disorders*, *11*(1). https://doi.org/10.1186/s11689-019-9268-y

Lefebvre, J. S., Kostadinov, D., Chen, W., Maniatis, T., & Sanes, J. R. (2012). Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature*, *488*(7412), 517–521. https://doi.org/10.1038/nature11305

Mancini, M. C., Bassani, S., & Passafaro, M. (2020). Right Place at the Right Time: How Changes in Protocadherins Affect Synaptic Connections Contributing to the Etiology of Neurodevelopmental Disorders. *Cells*, *9*(12), 2711. https://doi.org/10.3390/cells9122711

Meltzer, A., & Van De Water, J. A. (2016). The Role of the Immune System in Autism Spectrum Disorder. *Neuropsychopharmacology*, *42*(1), 284–298. https://doi.org/10.1038/npp.2016.158

Monteggia, L. M., & Kavalali, E. T. (2009). Rett Syndrome and the impact of MECP2 associated transcriptional mechanisms on neurotransmission. *Biological Psychiatry*, *65*(3), 204–210. https://doi.org/10.1016/j.biopsych.2008.10.036

Mountoufaris, G., Canzio, D., Nwakeze, C. L., Chen, W., & Maniatis, T. (2018). Writing, Reading, and Translating the Clustered Protocadherin Cell Surface Recognition Code for Neural Circuit Assembly. *Annual Review of Cell and Developmental Biology*, *34*(1), 471–493. https://doi.org/10.1146/annurev-cellbio-100616-060701

Autism Spectrum Disorder. (2023). National Institute of Mental Health (NIMH). https://www.nimh.nih.gov/health/topics/autism-spectrum-disorders-asd

Pancho, A., Aerts, T., Mitsogiannis, M., & Seuntjens, E. (2020). Protocadherins at the Crossroad of Signaling Pathways. *Frontiers in Molecular Neuroscience*, *13*. https://doi.org/10.3389/fnmol.2020.00117

Peek, S. L., Mah, K. M., & Weiner, J. A. (2017). Regulation of neural circuit formation by protocadherins. *Cellular and Molecular Life Sciences*, *74*(22), 4133–4157. https://doi.org/10.1007/s00018-017-2572-3

Phillips, G. R., LaMassa, N., & Nie, Y. (2017). Clustered protocadherin trafficking. *Seminars in Cell & Developmental Biology*, *69*, 131–139. https://doi.org/10.1016/j.semcdb.2017.05.001

Rubinstein, R., Thu, C. A., Goodman, K. M., Wolcott, H. N., Bahna, F., Mannepalli, S., Ahlsen, G., Chevée, M., Halim, A., Clausen, H., Maniatis, T., Shapiro, L., & Honig, B. (2015b). Molecular Logic of Neuronal Self-Recognition through Protocadherin Domain Interactions. *Cell*, *163*(3), 629–642. https://doi.org/10.1016/j.cell.2015.09.026

Shah, R., & Bird, A. (2017). MeCP2 mutations: progress towards understanding and treating Rett syndrome. *Genome Medicine*, *9*(1). https://doi.org/10.1186/s13073-017-0411-7

Shapiro, L., & Weis, W. I. (2009). Structure and Biochemistry of Cadherins and Catenins. *Cold Spring Harbor Perspectives in Biology*, *1*(3), a003053. https://doi.org/10.1101/cshperspect.a003053

Shibata-Seki, T., Nagaoka, M., Goto, M., Kobatake, E., & Akaike, T. (2020). Direct visualization of the extracellular binding structure of E-cadherins in liquid. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-72517-2

Stathopoulos, S., Gaujoux, R., Lindeque, J. Z., Mahony, C., Van Der Colff, R., Van Der Westhuizen, F. H., & O'Ryan, C. (2020). DNA Methylation Associated with Mitochondrial Dysfunction in a South African Autism Spectrum Disorder Cohort. *Autism Research*, *13*(7), 1079–1093. https://doi.org/10.1002/aur.2310

Südhof, T. C. (2017). Synaptic Neurexin Complexes: A Molecular Code for the Logic of Neural Circuits. *Cell*, *171*(4), 745–769. https://doi.org/10.1016/j.cell.2017.10.024

Takumi, T., Tamada, K., Hatanaka, F., Nakai, N., & Bolton, P. (2020). Behavioral neuroscience of autism. *Neuroscience & Biobehavioral Reviews*, *110*, 60–76. https://doi.org/10.1016/j.neubiorev.2019.04.012

Thu, C. A., Chen, W., Rubinstein, R., Chevée, M., Wolcott, H. N., Felsovalyi, K., Tapia, J. A., Shapiro, L., Honig, B., & Maniatis, T. (2014). Single-Cell Identity Generated by Combinatorial Homophilic Interactions between α, β, and γ Protocadherins. *Cell*, *158*(5), 1045–1059. https://doi.org/10.1016/j.cell.2014.07.012

Tsai, N. P., & Huber, K. M. (2017). Protocadherins and the Social Brain. *Biological Psychiatry*, *81*(3), 173–174. https://doi.org/10.1016/j.biopsych.2016.10.025

Tsukasaki, Y., Miyazaki, N., Matsumoto, A., Nagae, S., Yonemura, S., Tanoue, T., Iwasaki, K., & Takeichi, M. (2014). Giant cadherins Fat and Dachsous self-bend to organize properly spaced intercellular junctions. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(45), 16011–16016. https://doi.org/10.1073/pnas.1418990111

Watts, T. J. (2008). The Pathogenesis of Autism. *Clinical Medicine*, *1*, CPath.S1143. https://doi.org/10.4137/cpath.s1143

Weis, W. I. (1995). Cadherin structure: a revealing zipper. *Structure*, *3*(5), 425–427. https://doi.org/10.1016/s0969-2126(01)00174-5

Weiss, L. M., Arking, D. E., Daly, M. J., Chakravarti, A., Brune, C. W., West, K., O'Connor, A., Hilton, G., Tomlinson, R. L., West, A. A., Cook, E. H., Green, T., Chang, S., Gabriel, S., Gates, C., Hanson, E., Kirby, A., Korn, J. M., Kuruvilla, F. G., Paterson, A. D. (2009). A genome-wide linkage and association scan reveals novel loci for autism. *Nature*, *461*(7265), 802–808. https://doi.org/10.1038/nature08490

Wu, Q., & Jia, Z. (2020). Wiring the Brain by Clustered Protocadherin Neural Codes. *Neuroscience Bulletin*, *37*(1), 117–131. https://doi.org/10.1007/s12264-020-00578-4 .