ANNUAL REVIEWS

Annual Review of Cell and Developmental Biology

Writing, Reading, and Translating the Clustered Protocadherin Cell Surface Recognition Code for Neural Circuit Assembly

George Mountoufaris,[1,2] Daniele Canzio,[1] Chiamaka L. Nwakeze,[1] Weisheng V. Chen,[1,3] and Tom Maniatis[1]

[1]Department of Biochemistry and Molecular Biophysics, Columbia University Medical School, and Zuckerman Institute, Columbia University, New York, NY 10027, USA; email: tm2472@cumc.columbia.edu

[2]Current address: Department of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125, USA

[3]Current address: Leveragen, Inc., Cambridge, Massachusetts 02139, USA

Annu. Rev. Cell Dev. Biol. 2018. 34:471–93

The *Annual Review of Cell and Developmental Biology* is online at cellbio.annualreviews.org

https://doi.org/10.1146/annurev-cellbio-100616-060701

Keywords

protocadherins, single-neuron identity, homophilic interactions, self-avoidance, tiling, stochastic expression, neural circuits, neuropsychiatric diseases

Abstract

The ability of neurites of individual neurons to distinguish between themselves and neurites from other neurons and to avoid self (self-avoidance) plays a key role in neural circuit assembly in both invertebrates and vertebrates. Similarly, when individual neurons of the same type project into receptive fields of the brain, they must avoid each other to maximize target coverage (tiling). Counterintuitively, these processes are driven by highly specific homophilic interactions between cell surface proteins that lead to neurite repulsion rather than adhesion. Among these proteins in vertebrates are the clustered protocadherins (Pcdhs), and key to their function is the generation of enormous cell surface structural diversity. Here we review

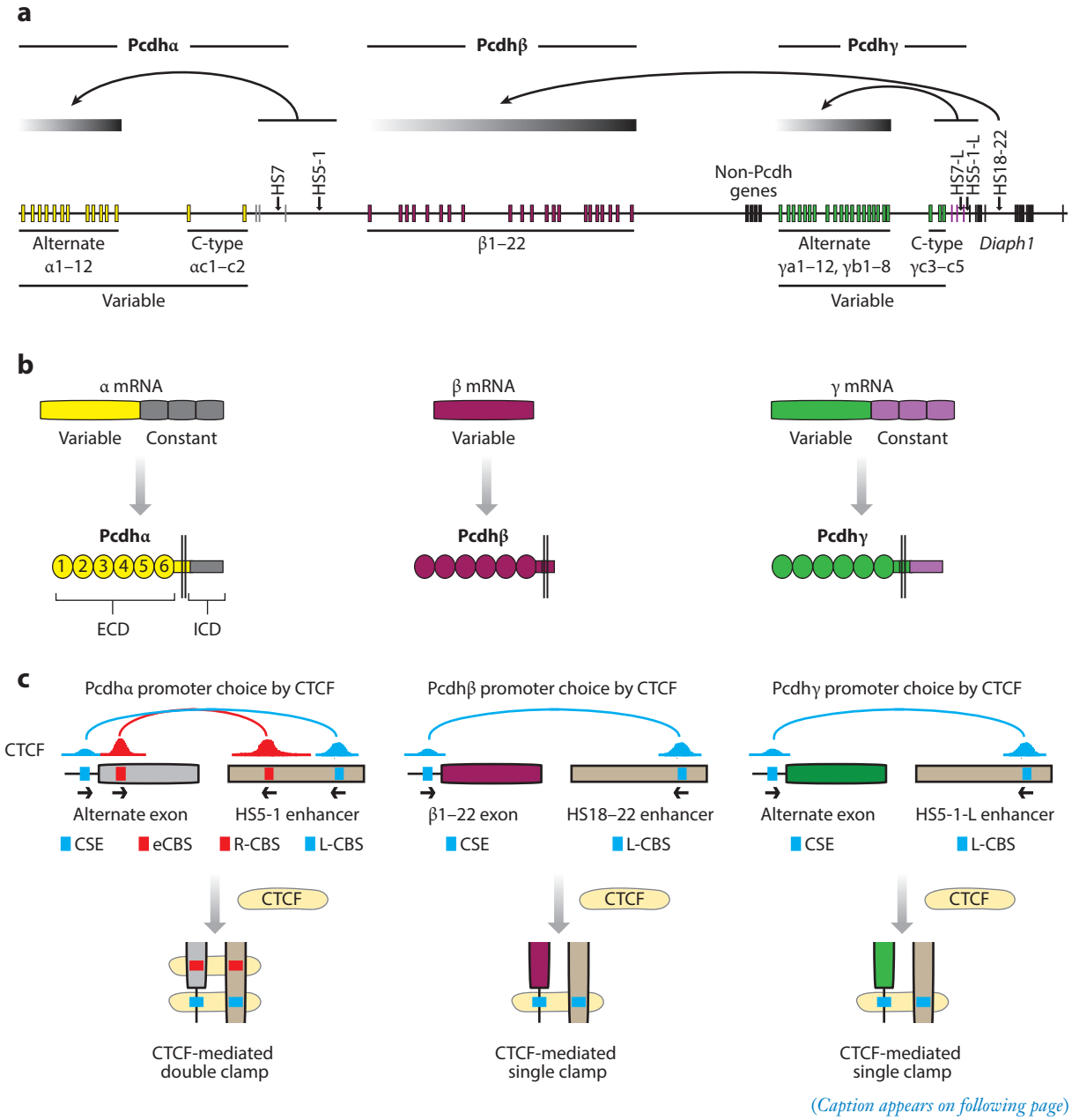
recent advances in understanding how a Pcdh cell surface code is generated by stochastic promoter choice; how this code is amplified and read by homophilic interactions between Pcdh complexes at the surface of neurons; and, finally, how the Pcdh code is translated to cellular function, which mediates self-avoidance and tiling and thus plays a central role in the development of complex neural circuits. Not surprisingly, Pcdh mutations that diminish homophilic interactions lead to wiring defects and abnormal behavior in mice, and sequence variants in the Pcdh gene cluster are associated with autism spectrum disorders in family-based genetic studies in humans.

Contents

INTRODUCTION

Neurons in the mammalian brain display remarkable diversity, as revealed by their patterns of gene expression; electrical, chemical, and morphological properties; and unique spatiotemporal connectivity maps that transform individual neurons into functional neural circuits. Recent evidence from biochemistry, structural, and functional studies points to clustered protocadherin (Pcdh) cell surface proteins as essential players in neural circuit assembly. The Pcdh gene family was first identified in a polymerase chain reaction (PCR) screen of brain cDNAs using degenerate primers designed to detect classic cadherins (Sano et al. 1993). Because of the unusually broad and early evolutionary representation of Pcdhs, as well as their cadherin-like properties, the term protocadherin—after the Greek word *proto*, meaning "first"—was used to describe these genes (Sano et al. 1993, Suzuki 1996). Relative to classic cadherins, the Pcdhs have distinct sequence features in their extracellular domains (cadherin-like repeats) and distinct intracellular domains. The genes that became known as the clustered Pcdhs were initially designated as CNRs (cadherin-related neuronal receptors) (Kohmura et al. 1998) because of their predominant expression in the nervous system and presumed association with the protein kinase Fyn on the basis of a yeast two-hybrid screen (Kai et al. 1997). The most striking feature of the CNR cDNA-encoded protein sequences at the time of their discovery was that their N-terminal extracellular domains are diverse,

while their C-terminal intracellular domains are identical. This interesting sequence organization suggested the possibility that mature CNR mRNAs are generated by alternative pre-mRNA splicing. To investigate this possibility, CNR cDNA sequences were used to align the then rapidly accumulating genomic DNA sequences deposited in GenBank as part of the Human Genome Project (Wu & Maniatis 1999). This effort, which required both computer-based and manual alignment of short regions of genomic DNA sequences, led to the assembly of a genomic region of approximately 750 kb. This effort revealed an extraordinary genomic architecture reminiscent of that of the immunoglobulin and T cell receptor genetic loci (Wu & Maniatis 1999) (**Figure 1*a***).



(Caption appears on following page)

Figure 1 (*Figure appears on preceding page*)

Genomic organization and expression of murine clustered *Pcdh* genes. (*a*) An illustration of the three *Pcdh* gene clusters: *Pcdhα*, *Pcdhβ*, and *Pcdhγ*. Tandemly arrayed exons in the *Pcdhα* and *Pcdhγ* gene clusters are divided into two categories: variable exons (indicated by the *yellow*, *maroon*, and *green rectangles*) and constant exons [indicated by the *gray* (α-cluster) and *purple* (γ-cluster) *rectangles*]. There are two types of variable exons: alternate exons, which are stochastically expressed, and C-type exons, which are independently regulated. The *Pcdhβ* cluster does not contain constant exons. The DNase I hypersensitive sites (HS) in the *Pcdhα*, *Pcdhβ*, and *Pcdhγ* regulatory regions are also shown. The gradient bar above each Pcdh cluster indicates the strength of functional interaction between individual gene clusters and their identified regulatory elements. *Diaph1* denotes Diaphanous 1. (*b*) RNA splicing joins each α or γ variable exon with the three cluster-specific constant exons to produce mature Pcdh mRNA. Each variable exon encodes the entire extracellular domain (ECD), transmembrane domain, and proximal intracellular domain (ICD), while the constant exons encode the common distal intracellular domain of the corresponding Pcdh protein. (*c*) Assembly of a promoter-enhancer complex by the CCCTC-binding factor (CTCF). Arrows indicate the orientation of conserved sequence elements and CTCF-binding sites (CBSs) in variable exons and regulatory elements (enhancers). Abbreviation: CSE, conserved sequence element.

These *Pcdh* genes were found to consist of three closely linked gene clusters (designated *Pcdhα*, *Pcdhβ*, and *Pcdhγ*) arranged in tandem, with the genes encoding the CNR mRNAs corresponding to the *Pcdhα* gene cluster (Wu & Maniatis 1999). This genomic organization and complexity led to speculation that clustered Pcdhs could provide a "molecular code required to establish complex networks of neuronal connections in the brain" (Wu & Maniatis 1999). Thus, although the association of CNR proteins with Fyn could not be validated in vivo (Yagi 2008), the characterization of the isolated cDNA clones opened a window into a fascinating multigene family of fundamental importance to neural development and neurological disorders (Chen & Maniatis 2013).

At approximately the same time that the Pcdh gene cluster was discovered and its genomic organization determined (Wu & Maniatis 1999), an equally remarkable genomic organization of the *Drosophila* Down syndrome cell adhesion molecule (*Dscam1*) gene was discovered by the Zipursky lab (Schmucker et al. 2000) (**Supplemental Figure 1**). This organization revealed, in conjunction with follow-up studies, the most striking example of alternative pre-mRNA splicing reported for any gene to date. The Zipursky lab speculated that this molecular diversity may "play an important role in neuronal connectivity and patterning in invertebrates," a speculation that was later confirmed by elegant genetic studies in *Drosophila* and by biophysical and structural studies of Dscam1 proteins (reviewed in Hattori et al. 2008, Zipursky & Grueber 2013). As discussed below, the mammalian clustered Pcdhs and the *Drosophila* Dscams provide a remarkable example of convergent evolution, whereby distinct mechanisms have evolved to generate the cell surface diversity required for neuronal self-recognition and neural circuit assembly.

Supplemental Material >

In this review, we provide a broad overview of and perspective on the mechanisms by which the clustered Pcdh code is generated, or "written," at the level of transcription and RNA splicing. We discuss how the reading of this code on the cell surface of neurons is largely due to the unique structure of clustered Pcdh protein *cis* dimers, which dramatically expands functional protein diversity. Finally, we provide examples in which the Pcdh code is translated into cellular functions essential for neural circuit assembly in mice. We refer the reader to a number of recent reviews that explore individual aspects of the clustered Pcdh gene family in greater detail (Hirayama & Yagi 2017, Lefebvre 2017, Mah & Weiner 2017, Peek et al. 2017, Rubinstein et al. 2017).

WRITING THE PCDH CELL SURFACE RECOGNITION CODE

The three Pcdh gene clusters (α, β, and γ) are located on human chromosome 5 (Wu & Maniatis 1999) and on mouse chromosome 18 (Wu et al. 2001). Together, the α, β, and γ Pcdh gene clusters encode 52 Pcdh protein isoforms in humans (15α, 15β, and 22γ) and 58 isoforms in mice (14α, 22β, and 22γ). DNA sequence analyses revealed a high degree of evolutionary conservation

across vertebrate species (Wu et al. 2001; Yu et al. 2007, 2008) (**Figure 1*a***). The extracellular domain, the transmembrane domain, and the variable intracellular domain of each Pcdh protein are encoded by unusually large (between 2 and 3 kb in length) variable exons, arrayed in tandem, in each gene cluster (Wu & Maniatis 1999). The distinct intracellular domains of Pcdhα and Pcdhγ proteins are encoded by three constant exons located at the distal end of each gene cluster (Wu & Maniatis 1999). By contrast, the *Pcdhβ* gene cluster lacks constant exons. Consequently, the Pcdhβ proteins do not contain a common intracellular domain. Among the variable exons, there are two and three C-type exons located at the end of the *Pcdhα* and *Pcdhγ* gene clusters, respectively (for Pcdhα, the exons are designated as c1 and c2, and for Pcdhγ, the exons are designated as c3, c4, and c5) (Wu & Maniatis 1999). The C-type Pcdhs are phylogenetically more similar to each other than to the other cluster-specific isoforms, and their expression and functions are distinct (see below). All non-C-type variable exons are designated as alternate exons on the basis of their distinct single neuron expression patterns, as described below (Wu et al. 2001).

Transcription and Splicing of Pcdh RNA

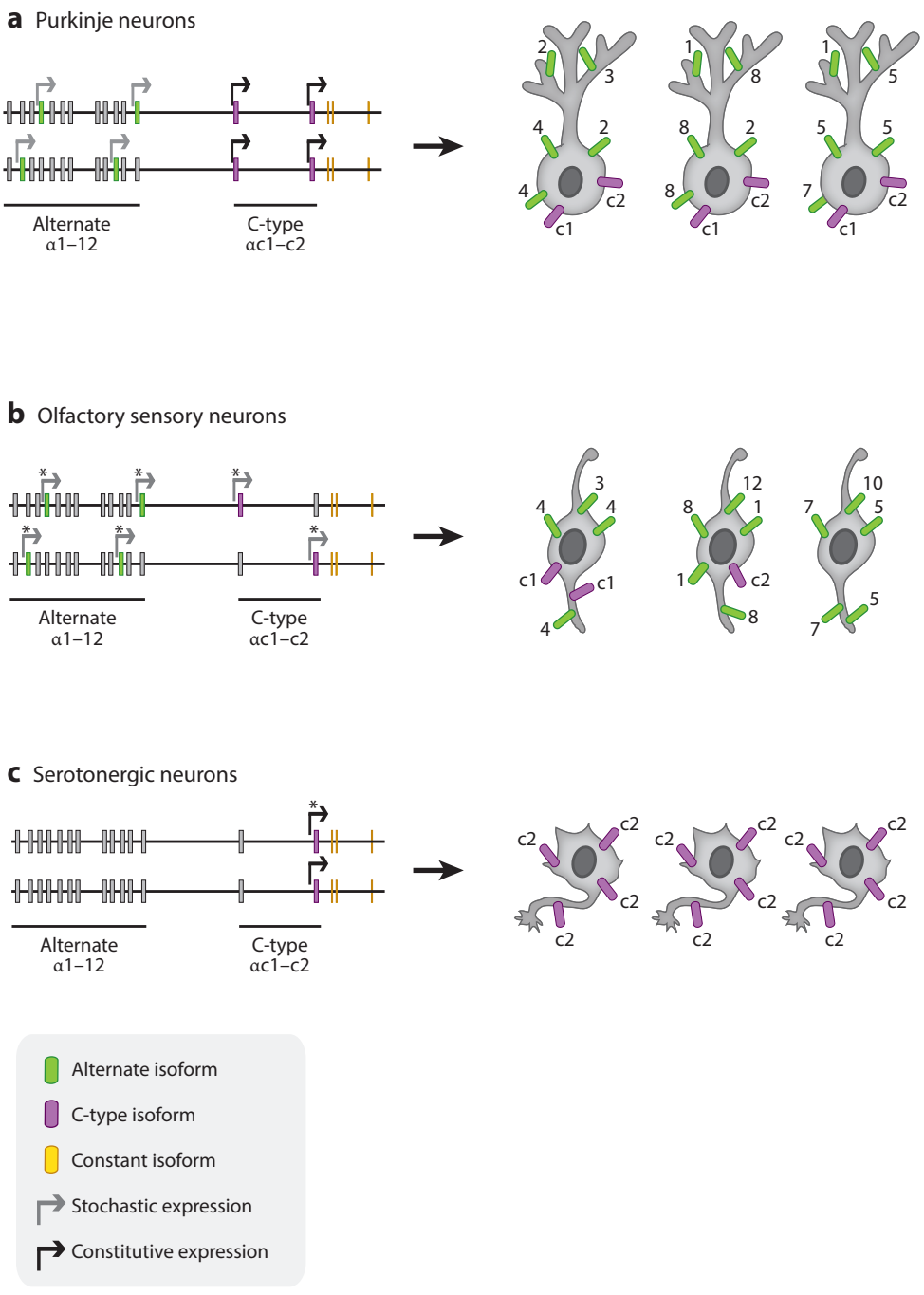
The unique genomic organization of the Pcdh gene cluster led to several models for Pcdh gene expression (Wu & Maniatis 1999). The earliest insights into the transcriptional regulation of the clustered Pcdhs were provided by studies of the *Pcdhα* (Tasic et al. 2002) and *Pcdhγ* (Wang et al. 2002a) gene clusters. These studies revealed that each Pcdh variable exon contains a highly conserved promoter sequence immediately upstream of the coding sequence of each gene, where transcription is initiated at this conserved sequence. When Pcdh exons are transcribed, the first 5′ splice site downstream from the transcriptional start site (more than 2–3 kb away) is spliced to the only known functional 3′ splice site in the gene cluster; this 3′ splice site is located immediately upstream from the first constant exon (**Figure 1*b***). The distance between the DNA sequences encoding the 5′ splice site proximal to the actively transcribed exon and the 3′ splice site can thus be as much as 200–300 kb in the *Pcdhα* and *Pcdhγ* gene clusters. Given that mammalian pre-mRNAs are transcribed at a rate of 1 to 3 kb/min, approximately 1.6 to 5 h would be required to transcribe a single Pcdh pre-mRNA bearing both the 5′ and 3′ splice sites (Maiuri et al. 2011). If such a long pre-mRNA is synthesized, all of the 5′ splice sites downstream from the first exon must be ignored by the spliceosome so that only the promoter-proximal 5′ splice site is joined to the single functional 3′ splice site in the cluster located immediately upstream of the first constant exon. Various mechanisms for this unusual splicing have been proposed, including *trans* splicing between distinct constant and variable region RNA precursors (Wu & Maniatis 1999). While efforts to unequivocally demonstrate *trans* splicing of Pcdh pre-mRNAs were unsuccessful (Tasic et al. 2002, Wang et al. 2002a), the question of whether mature Pcdh mRNAs are generated by *trans* splicing has yet to be definitively answered.

Neuronal Cell Type–Specific Expression of Clustered Pcdh Genes

Early single-cell RT-PCR (reverse transcriptase PCR) analyses, using individual Purkinje neurons from F1 generation mice derived from two distantly related mouse strains, revealed that the variable exon promoters in all three clusters are stochastically and independently activated on each of the two allelic chromosomes (Esumi et al. 2005, Hirano et al. 2012, Kaneko et al. 2006). On the basis of these studies, one or sometimes two alternate exon promoters are transcriptionally active in each of the *Pcdhα*, *Pcdhβ*, and *Pcdhγ* gene clusters of the paternal and maternal alleles. By contrast, promoters driving the expression of all of the C-type Pcdh genes on both chromosomes (αc1, αc2, γc3, γc4, and γc5) are active in all Purkinje neurons (Esumi et al. 2005, Kaneko et al. 2006) (**Figure 2*a***). Thus, it was estimated that 10 to 15 Pcdh protein isoforms are expressed in

Figure 2

Single-neuron expression of Pcdh protein isoforms in (*a*) Purkinje neurons, (*b*) olfactory sensory neurons, and (*c*) serotonergic neurons. Only the *Pcdhα* gene cluster is shown for simplicity. Note the independent stochastic expression of alternate Pcdh isoforms in Purkinje neurons and olfactory sensory neurons. The C-type Pcdh isoforms are constitutively and biallelically expressed in Purkinje cells, whereas in olfactory sensory neurons these isoforms are either absent or stochastically expressed at low levels. By contrast, individual serotonergic neurons predominantly express the Pcdhαc2 isoform. Asterisks indicate the lack of evidence for monoallelic versus biallelic expression of *Pcdh* genes in olfactory sensory neurons and serotonergic neurons.



individual Purkinje cell neurons in random combinations (Esumi et al. 2005, Hirano et al. 2012, Kaneko et al. 2006).

All of the initial single-cell PCR data were generated from Purkinje neurons, which were chosen for their unusually large size and thus significant amount of RNA per cell. However, recent single-cell RNA studies have revealed differences in Pcdh gene expression in different cell types [e.g., olfactory sensory neurons (OSNs) and serotonergic neurons] (Chen et al. 2017, Hasegawa et al. 2008, Katori et al. 2017, Mountoufaris et al. 2017). Although the pattern of *Pcdh* gene expression in OSNs is similar to that observed in Purkinje cells, the C-type Pcdhs are differentially expressed during OSN maturation (Mountoufaris et al. 2017) (**Figure 2*b***). Thus, the variable and C-type *Pcdhs* appear to be independently regulated. Remarkably, in serotonergic neurons only the *Pcdhα* and *Pcdhγ* C-type genes are expressed, and no alternate *Pcdhα*, *Pcdhβ*, or *Pcdhγ* mRNAs were detected (Chen et al. 2017) (**Figure 2*c***). These observations reveal the existence of complex regulatory mechanisms that control the temporal and neuron type–specific expression of alternate and C-type *Pcdh* genes in the brain.

The Organization of Transcriptional Enhancers in the *Pcdh* Gene Cluster

Transcriptional enhancer elements were identified and characterized in the *Pcdhα* gene cluster in an effort to gain insights into the mechanisms involved in Pcdh promoter choice (Ribich et al. 2006). The DNase I hypersensitivity assay was used to identify and functionally characterize a number of potential *cis* regulatory elements (Ribich et al. 2006). Two enhancer elements, HS5-1 and HS7 (where HS denotes hypersensitive), were first shown to be required for maximal levels of *Pcdhα* gene expression in mice (Ribich et al. 2006) (**Figure 1**). Deletion of HS5-1 decreased the expression of the *Pcdhα* 1–12 and *Pcdhαc1* genes, but not *Pcdhαc2*, whereas deletion of HS7 decreased the expression of all Pcdhα isoforms, including Pcdhαc2 (Kehayova et al. 2011). Subsequently, a cluster of enhancer elements, located downstream of the *Pcdhγ* gene cluster (designated as a cluster control region), was shown to be required for maximal expression of the *Pcdhβ* and *Pcdhγ* gene clusters (Yokota et al. 2011) (**Figure 1*a***). Although less studied, an HS5-1-like enhancer is located at the 3′ end of the *Pcdhγ* cluster between the *Pcdhγ* constant region and the Diaphanous gene (Yokota et al. 2011) (**Figure 1*a***). Finally, the *Pcdhβ* gene cluster enhancers appear to be located downstream of the *Pcdhγ* cluster, within the adjacent Diaphanous gene (Yokota et al. 2011) (**Figure 1*a***). The organization of DNA transcriptional regulatory elements in the Pcdh gene cluster was recently reviewed (Hirayama & Yagi 2017).

Distinct Mechanisms for Alternate and C-Type Pcdh Gene Expression

Differential activation of the alternate and C-type Pcdh promoters appears to be at least partly due to distinct regulatory mechanisms. Common to all Pcdh alternate promoters and the *Pcdhαc1* promoter is the presence of a conserved sequence element (CSE) (Tasic et al. 2002) and the binding of the insulator protein CCCTC-binding factor (CTCF) (Golan-Mashiach et al. 2012, Kehayova et al. 2011, Monahan et al. 2012) (**Figure 1*c***). Studies performed in a variety of mouse and human neuroblastoma cell lines that express distinct combinations of Pcdhs show that the CSE in active promoters is hypomethylated (Guo et al. 2012, Kawaguchi et al. 2008, Tasic et al. 2002) and is bound by both CTCF and Cohesin (Guo et al. 2012, Monahan et al. 2012). By contrast, the CSEs from transcriptionally silent Pcdh genes are highly methylated and are not bound by either CTCF or Cohesin (Guo et al. 2012). The functional significance of this observation was demonstrated by the observation that deletion of the CTCF gene in the mouse cortex and hippocampus leads to the downregulation of Pcdh isoforms, showing that CTCF is indeed required for Pcdh gene

expression (Hirayama et al. 2012). DNA methylation of Pcdh promoters requires the de novo DNA methyltransferase Dnmt3b, which is expressed at high levels in neuronal tissues during early embryogenesis prior to E10 (Toyoda et al. 2014). Consistent with the role of Dnmt3b in regulating promoter activation, Dnmt3b knockdown led to increased expression of Pcdh isoforms in individual Purkinje and cortical neurons (Tarusawa et al. 2016, Toyoda et al. 2014). Understanding the relationship between the timing and occurrence of DNA methylation, CTCF and Cohesin binding, and transcriptional activation during neuronal differentiation is of fundamental importance for understanding how the Pcdh code is generated.

As with many genome-wide enhancer-promoter interactions, both CTCF and Cohesin mediate DNA looping between active Pcdh promoters and cluster-specific enhancers (Guo et al. 2012, Monahan et al. 2012). In neuroblastoma cell lines in which specific *Pcdhα* variable gene promoters are active, the HS5-1 enhancer is bound to CTCF and Cohesin and is looped to the active promoters, as revealed by chromosome conformation capture experiments (Guo et al. 2012, 2015; Monahan et al. 2012) (**Figure 1*c***). Additional studies revealed that CTCF/Cohesin-mediated DNA looping between active *Pcdhα* promoters and the HS5-1 enhancer is required for transcriptional activation (Guo et al. 2012, 2015; Monahan et al. 2012) (**Figure 1**). Interestingly, in nonneuronal cells, the RE1-silencing transcription factor (REST)/NRSF complex is bound to the HS5-1 enhancer, which represses Pcdh gene expression (Kehayova et al. 2011).

A striking feature of the organization of the regulatory sequences in the *Pcdhα* gene cluster is the position and orientation of the CTCF binding sites (**Figure 1*c***). There are two CTCF binding sites in each alternate Pcdh gene—one in the promoter and a second in the downstream variable exon—and they are on average approximately 675 bp apart (Monahan et al. 2012). Similarly, there are two CTCF binding sites in the HS5-1 enhancer, with spacing similar to that in the promoters. ChIP-Seq (chromatin immunoprecipitation sequencing) experiments revealed that CTCF/Cohesin is bound to both promoter and exon sites in active promoters and in the HS5-1 enhancer (Guo et al. 2012, Monahan et al. 2012) (**Figure 1**). Thus, the active *Pcdhα* promoters and the HS5-1 enhancer appear to be brought together through DNA looping to form a double CTCF/Cohesin clamp (Guo et al. 2012, Monahan et al. 2012) (**Figure 1*c***). Surprisingly, the assembly of this functionally active Pcdh promoter-enhancer complex requires that the CTCF/Cohesin binding sites in the promoter and the exon be in opposite orientations to those in the HS5-1 enhancer. The functional significance of this organization was shown by using the CRISPR-Cas9 technologies to invert the orientation of the CTCF site in the enhancer relative to the promoter (Guo et al. 2015) (**Figure 1**). Remarkably, inversion of the HS5-1 enhancer significantly diminished the ability of the enhancer to promote DNA looping between the enhancer and the *Pcdhα* variable gene promoters (Guo et al. 2015). This enhancer inversion resulted in a significant decrease in *Pcdhα* gene cluster expression (Guo et al. 2015). Bioinformatics analyses revealed that the opposite orientation of CTCF binding sites in the Pcdh gene cluster occurs more generally throughout the genome, so this mechanism of CTCF/Cohesin-dependent DNA looping is likely used throughout the genome (Guo et al. 2015, Rao et al. 2014). The structural basis for the directionality and genome-wide conservation of CTCF binding was recently reported (Yin et al. 2017).

As mentioned above, expression of the C-type isoforms, with the exception of Pcdhαc1, is independent of the enhancers (Kehayova et al. 2011, Yokota et al. 2011). By contrast, deletion of the HS5-1 enhancer in mice reduces the expression of the *Pcdhα* alternate exons and the *Pcdhαc1* gene (Kehayova et al. 2011, Yokota et al. 2011). However, deletion of the enhancer does not affect Pcdhαc2 expression (Kehayova et al. 2011). Thus, the differential expression of the Pcdh genes in OSNs relative to, for example, serotonergic neurons is likely regulated by the differential activation of Pcdh cluster-specific enhancers in different neuronal cell types.

Higher-Order Chromatin Architecture of the Clustered Pcdh Genes

Promoter choice, transcription, and the activities of enhancer elements are essential for generating differential expression patterns in individual neurons; however, new technologies developed in the last decade have also shown that chromatin topology and genome organization in the nucleus play an important role in DNA accessibility and thus gene expression. Significantly, the reconstruction of chromosome topology often relies on assessing DNA methylation patterns and histone modifications carried out by a host of DNA and chromatin readers, writers, and erasers with specific temporal, spatial, and cell type–specific expression patterns. For example, in the Pcdh gene cluster, DNA methylation is coupled to histone H3 lysine 9 trimethylation (H3K9me3), which is mediated by the SET domain bifurcated 1 (Setdb1) protein (Jiang et al. 2017). H3K9me3 heterochromatin has been proposed to transcriptionally repress Pcdh expression by limiting CTCF binding to the Pcdh promoters and by organizing the Pcdh gene cluster into a higher-order chromatin configuration designated as topological associated domains (TADs) (Jiang et al. 2017). Such chromatin architecture of the Pcdh gene locus appears to be essential for regulating the enhancer-promoter interactions in the Pcdh gene cluster (Jiang et al. 2017). Studies of mice bearing a conditional deletion of Setdb1 in the brain revealed abnormal accumulation of CTCF at cryptic binding sites, DNA cytosine hypomethylation, histone hyperacetylation, and upregulation of gene expression in the *Pcdh* gene cluster (Jiang et al. 2017). The absence of Setdb1 leads to alteration of the clustered Pcdh TADs and in an increase in clustered Pcdh expression in cortical neurons (Jiang et al. 2017). Similarly, the Wiz protein, identified in a screen for modifiers of chromatin modifications, is involved in the regulation of *Pcdhβ* gene expression in mice (Isbel et al. 2016). In this case, Wiz haploinsufficiency decreases global gene expression, which includes the clustered Pcdhs (Isbel et al. 2016). Wiz binds to CTCF binding sites genome wide, as well as in the *Pcdhβ* gene cluster (Isbel et al. 2016).

An additional regulator of the *Pcdh* gene locus is the Smchd1 (structural maintenance of chromosome hinge domain containing 1) protein. Smchd1 contains an N-terminal ATPase domain and a C-terminal SMC (structural maintenance of chromosomes) hinge domain and belongs to the SMC family of proteins (Chen et al. 2015). Smchd1 is thought to antagonize CTCF binding at the alternate *Pcdhα* gene *cis*-regulatory regions and HS5-1, thereby silencing their expression (Chen et al. 2015). Indeed, loss of Smchd1 results in upregulation of *Pcdhα* variable region genes and the *Pcdhβ* 5′-most exons in neuronal stem cells (Chen et al. 2015). Patients with mutations in *SMCHD1* display changes in global methylation and gene expression patterns (Mason et al. 2017).

Given the unique genomic organization of the Pcdh gene cluster and the complexity of stochastic gene expression, the Pcdh locus has become a model system for studying the relationship between chromatin structure, histone and DNA modifications, and DNA looping in the regulation of gene expression. However, numerous fundamental questions regarding the mechanisms of Pcdh gene expression remain to be addressed. For example, how the differential expression of alternate and C-type Pcdh genes is regulated during neuronal differentiation is not understood. In addition, studies of the human neuroblastoma cell line SK-N-SH, which has provided deep insights into the relationship between Pcdh gene expression, DNA methylation, and CTCF/Cohesin binding, also revealed a remarkable property of the Pcdh system. That is, the promoter choice made during the generation of this cell line is stably transmitted to cellular progeny. The mechanistic basis of this remarkable example of epigenetic inheritance is not understood. However, if the underlying mechanism is in play during development of the nervous system, it is important to identify the time at which promoter choice occurs and to follow the transmission of the cell surface code through cell division and postmitotic differentiation.

READING THE PCDH CELL SURFACE RECOGNITION CODE

Supplemental Material >

Genetic studies showed that the *Drosophila Dscam1* gene provides a cell surface protein code that functions in self-recognition and neurite self-avoidance (Hattori et al. 2008) (**Supplemental Text, Section 1**). It therefore seemed reasonable to propose that the vertebrate *Dscam*1 gene homologs play a similar role. However, the remarkable genomic organization of the *Drosophila Dscam1* gene, which affords high levels of alternative pre-mRNA splicing, is not conserved in vertebrates, and functional studies of mouse *Dscam1* genes revealed that they are not required for self-avoidance (**Supplemental Text, Section 1; Supplemental Figure 1**). Thus, the Pcdhs appeared to be the only known vertebrate genes that provide high levels of cell surface diversity. However, while *Drosophila Dscam1* has the potential to generate more than 19,000 distinct extracellular isoforms, there are only 58 Pcdh isoforms in mice. How then could the clustered Pcdhs provide the diversity necessary for self-avoidance in vertebrates? The answer appears to lie in differences in the structure and nature of the homophilic interactions between the two types of cell surface proteins. Earlier studies of the Dscam1 proteins revealed that they display "exquisite isoform-specific homophilic binding" (Wojtowicz et al. 2007). Three of the immunoglobulin (Ig)-like domains (Ig2, Ig3, and Ig7) of the Dscam1 proteins engage in specific homophilic interactions in a head-to-head parallel interaction, whereby each of the Ig domains in one monomer binds to its partner in *trans* (interacting Ig domains Ig2–Ig2, Ig3–Ig3, and Ig7–Ig7) (Li et al. 2016, Wojtowicz et al. 2007) (**Figure 3*a***).

Homophilic Interactions Between Pcdh Isoforms and Their Combinations

Recently, exciting progress has been made in understanding the homophilic affinities and the structures of the clustered Pcdh proteins, which have revealed a possible mechanism for generating cell surface diversity (see Rubinstein et al. 2017 for recent review). The development of an efficient cell aggregation assay has allowed for study of the homophilic interactions between Pcdhs (Schreiner & Weiner 2010, Thu et al. 2014) and their structures and biophysical properties (Rubinstein et al. 2015). Remarkably, except for Pcdhαc1 and Pcdhγc4, the entire Pcdh repertoire (*Pcdhα*, *Pcdhβ*, and *Pcdhγ* and the C-type Pcdhs) can engage in highly specific homophilic interactions in in vitro assays (Thu et al. 2014). Initially the Pcdh complexes were proposed to exist as *cis* tetramers (Schreiner & Weiner 2010). However, systematic analytical ultracentrifugation studies revealed that Pcdhs form *cis* dimers and that they can associate in random combinations, thus dramatically increasing the level of cell surface Pcdh diversity (Rubinstein et al. 2015).

A fundamentally important observation was made in this regard upon examining homophilic interactions between distinct combinations of multiple Pcdh isoforms in a cell aggregation assay (Thu et al. 2014). In these studies, DNA plasmids expressing individual Pcdh isoforms were tagged with either a red or a green fluorescent protein marker and were separately transfected into cells grown in suspension. When the separately transfected cells were mixed, cell aggregates containing both green and red cells were observed, revealing robust and highly specific homophilic interactions (Thu et al. 2014). Similarly, cell populations expressing up to five distinct Pcdh isoforms, labeled as described above, also formed mixed red and green aggregates. However, if the two cell populations differed in the expression of only a single nonmatching Pcdh isoform out of five (i.e., in the case of a single isoform mismatch), no cell aggregates were observed (Thu et al. 2014) (**Figure 3*d***). By contrast, replacing one of the Pcdh isoforms with N-cadherin (which does not interact with clustered Pcdh protein either in *cis* or in *trans*) did not prevent cell aggregation (Thu et al. 2014). These results suggested that the presence of diverse multiclustered Pcdh isoforms creates combinatorial homophilic specificities (recognition units), which differ from the specificities of the individual Pcdh isoforms, and that the presence of a single nonmatching isoform

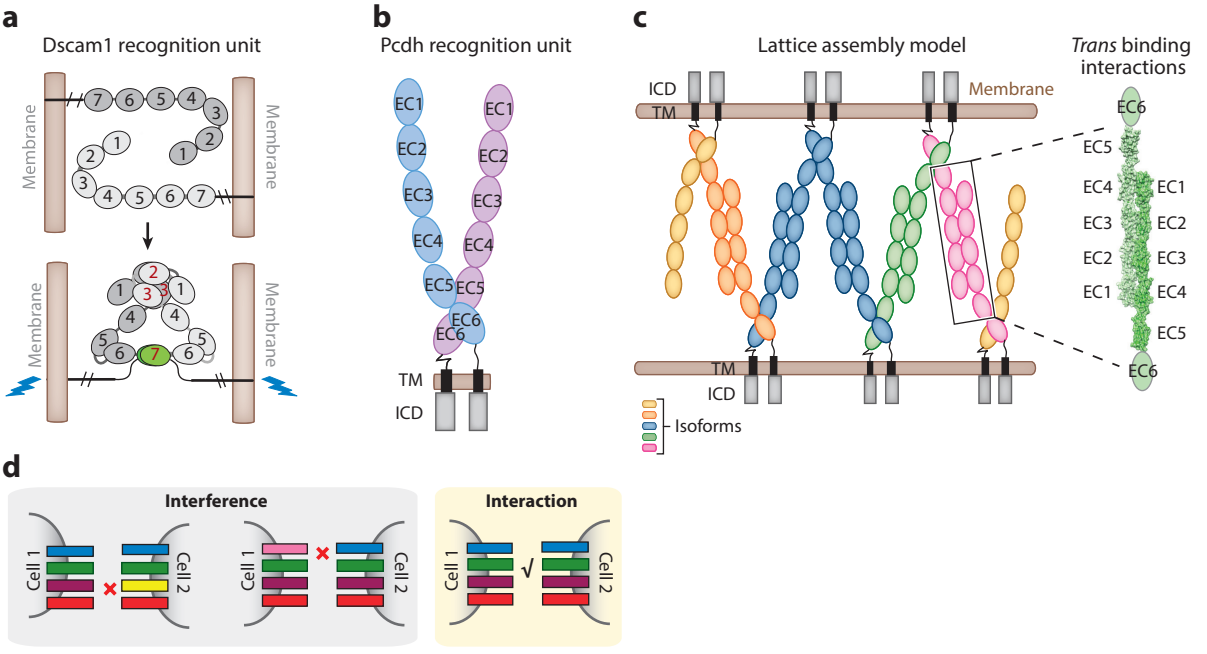


Figure 3

(*a*) Dscam monomeric recognition unit based on structural studies. Homophilic binding between the ectodomains of Dscam1 protein isoforms generates a repulsive signal between opposing membranes (Hattori et al. 2008). The homophilic binding of opposing Ig2, Ig3, and Ig7 domains is depicted by the red numbers. (*b*) Pcdh dimeric recognition unit based on biophysical, biochemical, and structural studies. (*c*) A lattice model for Pcdh-mediated cell-cell recognition based on formation of a superstructure defined by *cis* interactions (mediated by the EC5 and EC6 ectodomains) and highly specific *trans* interactions (zoom-in representation mediated by the EC1–EC4 ectodomains). The color coding represents distinct Pcdh isoforms (Rubinstein et al. 2015). (*d*) Illustration of the outcome of cell-cell interactions dictated by combinatorial homophilic specificity of four distinct Pcdh isoforms based on cell aggregation studies (Rubinstein et al. 2015, Thu et al. 2014). In this schematic diagram, the mismatch of a single Pcdh isoform can interfere with homophilic specificity. The red "x" represents the absence of homophilic interactions in the presence of a single *Pcdh* isoform mismatch between two K562 cells. When all Pcdh isoforms match, homophilic binding interactions are indicated by the black check mark. Abbreviations: ICD, intracellular domain; TM, transmembrane.

interferes with the assembly of the Pcdh recognition unit (Thu et al. 2014). This phenomenon of interference appears to play a fundamental role in the recognition of the Pcdh code (**Figure 3*d***).

In contrast to classical cadherins, in which the first ectodomain (EC1) is critical for homophilic binding specificity, the second and third ectodomains (EC2 and EC3) of clustered Pcdhs are required for their strict homophilic interactions in *trans* (Schreiner & Weiner 2010, Thu et al. 2014); although EC1 is essential for the *trans* interaction, it does not contribute to specificity (Rubinstein et al. 2015, Schreiner & Weiner 2010, Thu et al. 2014) (**Figure 3*b***). Chimeric Pcdh isoforms generated by shuffling the EC2 and EC3 domains from different Pcdh isoforms also exhibited strict homophilic binding specificity (Rubinstein et al. 2015, Schreiner & Weiner 2010). This modular binding specificity, mediated by EC2 and EC3, is similar to that mediated by Dscam variable domains (Wojtowicz et al. 2004, 2007). As expected, EC2 and EC3 are the most diverse of the six ectodomains of Pcdh proteins, as these two ectodomains have not been subject to homogenizing gene conversion like the other ectodomains (Noonan et al. 2004). That is, diversity is conserved.

Cis Dimerization Between Pcdh Isoforms

Pcdhγ protein isoforms are found in large ~480- to 1,280-kDa complexes in membrane extracts from whole brain (Han et al. 2010). Mass spectrometry of purified complexes demonstrated that they contain both Pcdhα and Pcdhβ proteins. Coimmunoprecipitation (Co-IP) experiments support this finding; Pcdhα and Pcdhγ proteins coimmunoprecipitate from whole-brain extracts (Murata et al. 2004) and in overexpression experiments with the CAD neuroblastoma cell line (Schalm et al. 2010). These initial findings led to the hypothesis that Pcdhs from all three clusters oligomerize to form complexes. However, such biochemical assays usually fail to distinguish between *cis* interactions that occur within a cell and *trans* interactions that occur between cells. Evidence supporting the occurrence of *cis* interactions between Pcdh isoforms was provided by the finding that cell surface delivery of Pcdhα isoforms can be mediated by any Pcdh single isoform (β, γ, or C type) (Thu et al. 2014), presumably as a protein dimer (**Figure 3*b***). As mentioned above, these *cis* interactions are functionally important since they can create new homophilic specificities that differ from those conferred by individual Pcdh isoforms. Cell aggregation assays using truncated Pcdhs together with biophysical and computational analyses revealed the critical role of EC5 and EC6 domains in *cis* interactions, even in the absence of *trans* binding (Rubinstein et al. 2015, Thu et al. 2014). Analytical ultracentrifugation studies of Pcdhs showed that the EC6 domain mediates promiscuous Pcdh *cis* dimerization, with comparable or even stronger affinities than the *trans* interactions (Rubinstein et al. 2015, 2017). However, recent crystal structures of *cis*-dimeric Pcdhγb7 isoforms demonstrated the asymmetric nature of Pcdh *cis* dimers, which in the case of Pcdhγc4 and all Pcdhα isoforms can be achieved through the heterodimerization with carrier Pcdh isoforms in vitro (Goodman et al. 2017). Thus, the clustered Pcdh isoforms exist on cell surfaces as combinatorial *cis* dimers (**Figure 3*b***) formed by semipromiscuous pairing of different isoforms from all three Pcdh clusters as well the C-type isoforms (Rubinstein et al. 2015).

Three-Dimensional Structure of Clustered Pcdhs

Although the crystal structure of a full-length Pcdh isoform is yet to be determined, structures of proteins composed of the four or five N-terminal EC domains of members of mouse Pcdhα, Pcdhβ, and Pcdhγ have been solved (Goodman et al. 2016a,b; Nicoludis et al. 2015, 2016; Rubinstein et al. 2015, 2017). Interestingly, all of the structures are highly similar. Each EC domain consists of approximately 100 amino acids that form two-layer β sheet structures. Homophilic interactions between Pcdh isoforms are Ca^{2+} dependent (Thu et al. 2014), as Ca^{2+}-binding linkers connect the six successive EC domains.

A series of computational modeling and mutagenesis experiments based on the EC1–EC3 crystal structures identified the *trans* homophilic interface of Pcdh isoforms between cells (Nicoludis et al. 2015, Rubinstein et al. 2015). In contrast to the homophilic interactions that occur in classical cadherins, which involve antiparallel interactions, Pcdh-mediated homophilic interactions occur in an antiparallel EC1–EC4 domain interaction (**Figure 3*c***). Thus, in the context of the *cis* dimer, the EC1, EC2, EC3, and EC4 domains of one protein interact with the EC4, EC3, EC2, and EC1 domains of its partner in *trans* (Rubinstein et al. 2015, 2017) (**Figure 3*c***). On the basis of these findings, it was proposed that cell-cell recognition is mediated by a mechanism that couples *cis* and *trans* interactions of Pcdh isoforms: Pcdhs form semipromiscuous EC5/EC6-dependent *cis* dimers at the cell surface that engage specifically in *trans* through an antiparallel interface via EC1–EC4 (Goodman et al. 2016a,b; Rubinstein et al. 2017) (**Figure 3*c***). This model was recently supported on the basis of crystal structures of Pcdh EC domains, which adopt dimeric conformations geometrically consistent with their functions in *trans* cell-cell recognition (Goodman et al. 2016a,b).

These results led to the proposal that Pcdh proteins form a zipper-like lattice structure involving asymmetrical *cis* and symmetrical *trans* interactions (**Figure 3*c***). In this model, each Pcdh *cis* dimer interacts with two dimers through independent *trans* binding to form a connected ribbon of molecules emerging from the two opposing cell surfaces (Goodman et al. 2016a; Rubinstein et al. 2015, 2017). According to this model, cells with matching Pcdh isoforms would form large Pcdh protein assemblies upon contact (**Figure 3*d***). In contrast, cells with different isoform compositions would incorporate mismatches, preventing further growth of the lattice (Goodman et al. 2016a,b; Rubinstein et al. 2015) (**Figure 3*c***). Interestingly, computational modeling of Pcdh assembly size (assuming that each cell contains a stable set of *cis* dimers formed from the random association of available monomers) also predicted that the size of Pcdh assemblies could depend on the number of mismatched Pcdh isoforms (Rubinstein et al. 2015). Thus, when all isoforms are matching, the assembly size is limited only by the number of copies of each Pcdh isoform (**Figure 3*c*,*d***). However, calculations predicted that the presence of even a single mismatched isoform would be sufficient to reduce the average size of a lattice assembly by at least two orders of magnitude (Rubinstein et al. 2015). The size of the proposed lattice may be of great importance for downstream signaling such that large assemblies could transduce high levels of intracellular signals but signals transduced by smaller lattices would be below the critical threshold. Interestingly, recent statistical calculations based on single-cell RNA-Seq data from OSNs and the functional interrogation of multicluster organization of *Pcdh* in the mouse olfactory system (see below) are consistent with the zipper-like lattice model (Goodman et al. 2017). Taken together, these findings establish a structurally based mechanism for how the clustered Pcdhs, with only a limited number of genes (~60), can provide the sufficiently high levels of diversity required for single-cell identity within the mammalian nervous system.

TRANSLATING THE PCDH CELL SURFACE RECOGNITION CODE

Initial efforts to understand the in vivo function of individual Pcdh gene clusters, accomplished by the phenotypic characterization of constitutive and conditional clustered Pcdh knockout mice, revealed the multifaceted roles of clustered Pcdhs in neural development, ranging from neuronal apoptosis to synaptic development to dendritic and axonal arborization (Chen et al. 2012, Garrett et al. 2012, Hasegawa et al. 2008, Katori et al. 2009, Ledderose et al. 2013, Lefebvre et al. 2008, Prasad et al. 2008, Suo et al. 2012, Wang et al. 2002b, Weiner et al. 2005; reviewed in Lefebvre 2017, Peek et al. 2017) (**Supplemental Text, Section 3**). Here, we focus on how the Pcdh cell surface code functions in neurite spacing (self-avoidance and tiling; see also **Supplemental Text, Section 1**) and how the clustered Pcdhs establish neuronal single-cell and cell type identities that mediate neural circuit assembly in the mammalian nervous system.

Supplemental Material >

The *Pcdhγ* Cluster Is Required for Dendritic Self-Avoidance of Retinal Starburst Amacrine Cells

A functional connection between the invertebrate Dscam1 proteins and the vertebrate clustered Pcdhs was first made by the discovery that the *Pcdhγ* gene cluster is required for dendritic self-avoidance in mouse retinal starburst amacrine cells (SACs) in a cell-autonomous manner (Lefebvre et al. 2012) (**Figure 4*a***). Abolishing Pcdhγ protein selectively in SACs led to the collapse of their radially symmetric dendritic arbors, resulting in extensive crossing and bundling of sister branches (Lefebvre et al. 2012) (**Figure 4*c***). These dendritic defects occurred early in cellular differentiation and persisted even when apoptosis was genetically blocked, eliminating any effect from the secondary loss of neurites from interacting partners of SACs (Lefebvre et al. 2012). Single-cell

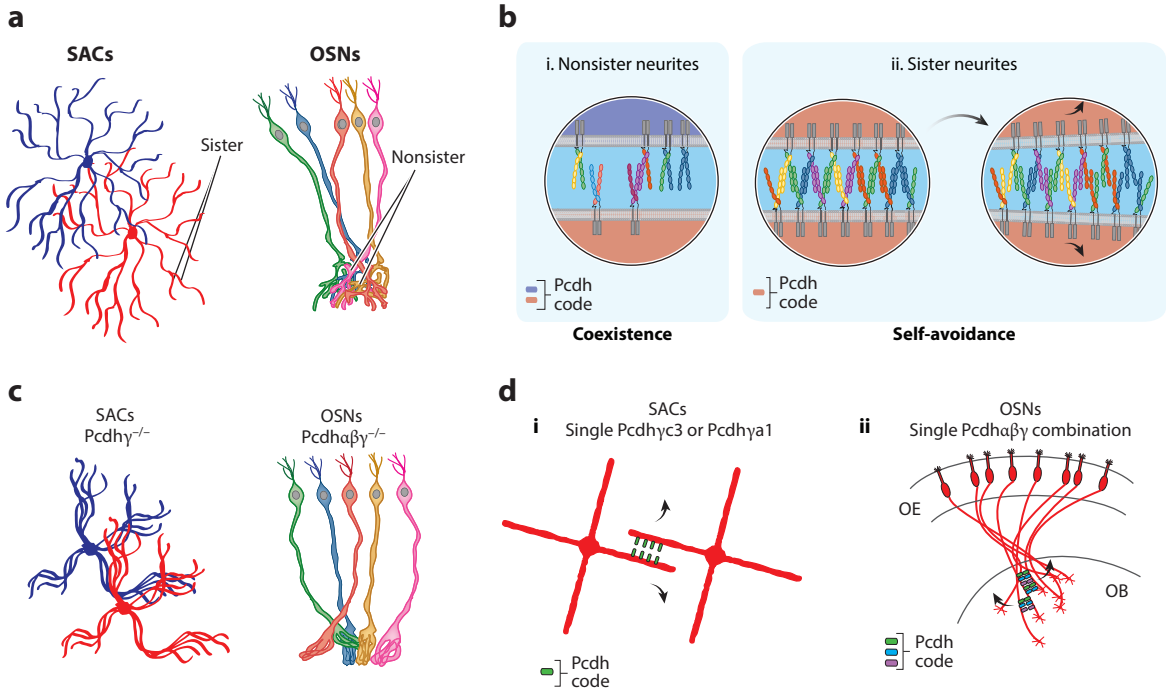


Figure 4

(*a*) A schematic illustration of neighboring starburst amacrine cells (SACs) and olfactory sensory neurons (OSNs). (*b*) A schematic illustration of the Pcdh isoform interaction model between adjacent neuronal membranes. (*i*) Each nonsister neurite has a unique Pcdh code. Mismatches between Pcdh protein isoforms from opposing interacting membranes obstruct the formation of an extensive, highly ordered Pcdh lattice, thus enabling neuronal processes to coexist within the target area (e.g., the glomerulus). (*ii*) In contrast, in the case of sister neurites, the formation of a Pcdh lattice leads to Pcdh-mediated repulsion and avoidance. The nature of this form of avoidance/repulsion is unknown, and this schematic merely represents the Pcdh interaction at the cell surface. (*c*) The effects of the loss of the Pcdh code on the distribution and configuration of neurites of SACs and OSNs, based on loss-of-function studies. Conditional deletion of the *Pcdhγ* gene cluster in SACs leads to the crossing and clumping of dendrites as a result of the loss of self-avoidance. Constitutive deletion of *Pcdhα*, *Pcdhβ*, and *Pcdhγ* (a tricluster deletion) leads to the collapse of axonal arbors in pioneer OSNs, reminiscent of loss of the self-avoidance phenotype. (*d*) (*i*) A single isoform, *Pcdhγc3* or *PcdhγA1*, prevents neurites from the same cell from clumping, but as a consequence of adjacent neurons expressing the same Pcdh code, these neurites fail to form synaptic interactions due to dendritic repulsion. A limited number of dendrites are shown for simplicity reasons. (*ii*) Expression of the same set of the three *Pcdhα*, *Pcdhβ*, and *Pcdhγ* isoforms leads to avoidance/repulsion between like OSN axons and failure to converge to form normal glomeruli. Abbreviations: OB, olfactory bulb; OE, olfactory epithelium. Subpanels *i* and *ii* of panel *d* adapted from Kostadinov & Sanes (2015) and Mountoufaris et al. (2017), respectively.

recordings from *Pcdhγ*-null SACs revealed the presence of autaptic currents upon stimulation (Kostadinov & Sanes 2015), providing evidence that one role of *Pcdhγ*-mediated self-avoidance in SACs is to prevent neurons from forming autapses (synapses to themselves). The dendritic self-avoidance phenotype was also observed in cerebellar Purkinje cells in *Pcdhγ*-deficient mice (Lefebvre et al. 2012), although the phenotype was less severe than that observed in SACs.

An additional observation regarding the *Pcdhγ* knockout SACs was that reintroduction of a single Pcdhγ protein isoform (Pcdhγa1 or Pcdhγc3) restored dendritic self-avoidance (Lefebvre et al. 2012). However, the expression of either of these Pcdh isoforms also led to a statistically

significant reduction of the relative dendritic overlap observed between adjacent SACs expressing a single Pcdhγ isoform (Lefebvre et al. 2012) (**Figure 4*d*,*i***). In addition, electrophysiological studies revealed a reduction of the connectivity strength, likely due to fewer synapses and/or weaker synaptic transmission between neighboring SACs (Kostadinov & Sanes 2015). By contrast, connectivity between the SACs and the direction-selective ganglion cells was not affected (Kostadinov & Sanes 2015). Given that the extent of the full Pcdh repertoire in these neurons is unknown, the mechanism by which overexpressing a single isoform leads to reduction of dendritic overlap is unclear. One can speculate that this gain-of-function phenotype in SACs lacking the expression of endogenous *Pcdhγ* is explained by the loss of single-cell identity, leading neurites from neighboring SACs to recognize each other as self and thus avoid each other (**Figure 4*b***).

Multicluster Pcdhs Function in Axonal Patterning and Coexistence

Early studies of the in vivo roles of clustered Pcdhs utilized genetic methods to inactivate or delete individual gene clusters (**Supplemental Text, Section 3**). Despite the observation that Pcdhγ protein isoforms could mediate dendritic self-avoidance, the role of Pcdhα or Pcdhβ proteins in self-recognition and self-avoidance had not been demonstrated. Moreover, abolishing *Pcdhα* or *Pcdhγ* gene functions resulted in a variety of different phenotypes, suggesting that individual Pcdh gene clusters may play either distinct or complementary roles in different neuronal cell types (**Supplemental Text, Section 3**).

Supplemental Material >

Significant progress was recently made in understanding the complexity of multicluster Pcdh function (Chen et al. 2017; Hasegawa et al. 2016, 2017; Katori et al. 2017; Mountoufaris et al. 2017). A series of single, double, or tricluster Pcdh deletion mutants were generated in mice, and the loss-of-function effects were examined. These independent studies provided a number of fundamentally important insights into the functional complexity of the clustered Pcdhs.

For example, OSNs express single olfactory receptors (ORs), and OSNs expressing the same OR (like OSNs) project their axons to stereospecific locations in the olfactory bulb, where the axons converge to form neuropil structures termed glomeruli (Mombaerts et al. 1996) (**Figure 4*a***). Glomeruli present an extreme case of neurite coexistence in which thousands of OSN axons intermingle (Bressel et al. 2016, Takeuchi & Sakano 2014). Single-cell RNA-Seq experiments revealed that Pcdhα, Pcdhβ, and Pcdhγ isoforms are stochastically expressed in OSNs (**Figure 2*b***) and localize throughout OSN axons and axonal termini (Hasegawa et al. 2008, Mountoufaris et al. 2017). By P0, distinct early glomeruli (referred as protoglomeruli) have begun to emerge, which sets the stage for the formation of the olfactory sensory map during the first postnatal week. Deletion of individual Pcdh gene clusters led to the appearance of somewhat less compact protoglomeruli (in the case of $Pcdh\alpha^{-/-}$) or had no major effect on OSN wiring (protoglomerulus formation) (in the case of $Pcdh\beta^{-/-}$ or $Pcdh\gamma^{-/-}$) when compared to heterozygote littermate pups (Mountoufaris et al. 2017; G. Mountoufaris, C.L. Nwakeze & T. Maniatis, unpublished observations), although quantitative differences in the number and size of the protoglomeruli were reported (Hasegawa et al. 2016). By contrast, deletion of two ($Pcdh\alpha\beta^{-/-}$, $Pcdh\beta\gamma^{-/-}$) or all three ($Pcdh\alpha\beta\gamma^{-/-}$) gene clusters resulted in a profound disruption of protoglomerular structures (Hasegawa et al. 2016, Mountoufaris et al. 2017). The total absence of all clustered Pcdhs does not affect OSN survival (Mountoufaris et al. 2017), and $Pcdh\alpha\beta\gamma^{-/-}$ OSN axons expressing the same OR still converge to approximately normal locations in the olfactory bulb (Mountoufaris et al. 2017). Strikingly, however, single-neuron labeling studies revealed that pan-Pcdh-deficient OSNs fail to arborize properly in the olfactory bulb; the arbors appeared highly distorted, enlarged, and less elaborated (Mountoufaris et al. 2017), reminiscent of the axonal self-avoidance phenotype

of *Drosophila Dscam1* mutant OSNs (Hummel et al. 2003, Zhu et al. 2006) (**Figure 4*c***). Taken together, these results demonstrate the importance of the multicluster Pcdh code in OSN wiring.

Additional insights into the role of multiclustered Pcdhs in OSN wiring were obtained with a gain-of-function experiment (Mountoufaris et al. 2017). In this approach, endogenous Pcdh diversity in OSNs was overridden by the expression of high levels of distinct combinations of Pcdhα, Pcdhβ, and Pcdhγ isoforms, thus endowing all OSNs with the same tricluster Pcdh codes (**Figure 4*d,ii***). Mutant mice were engineered in such a way that these codes could be either expressed broadly in all mature OSNs or restricted in expression to like OSNs. Remarkably, OSN axons bearing the same tricluster Pcdh codes largely failed to form glomeruli but rather projected diffusely to the approximate target sites in the olfactory bulb (Mountoufaris et al. 2017). This result is likely due to ectopic repulsion between the matching sets of Pcdh isoforms expressed in neighboring OSN axons (**Figure 4*b*** and **Figure 4*d,ii***). Surprisingly, although distinct glomerular structures are largely absent, these mutant mice were not anosmic (Mountoufaris et al. 2017). Rather, their ability to discriminate between different mixtures of odorants was adversely affected (Mountoufaris et al. 2017). Overexpression of truncated Pcdh mutant isoforms lacking their intracellular domains failed to impede glomeruli formation, highlighting the importance of intracellular signaling mediated by Pcdh proteins in generating a functional Pcdh code (Mountoufaris et al. 2017).

Ing-Esteves et al. (2018) recently demonstrated the functional contribution of the *Pcdhα* and *Pcdhγ* gene clusters to postnatal cell survival and neural dendritic patterning. To overcome the neonatal lethality caused by germline deletion of the *Pcdhγ* gene cluster, the authors generated mice bearing a conditional *Pcdhγ* allele and a constitutive *Pcdhα*-null allele. A synergistic loss of inner retinal layers was observed in these double mutant mice, consistent with a previous report (Hasegawa et al. 2016). Interestingly, the loss of both Pcdhα and Pcdhγ protein isoforms exacerbated the disruption of dendritic self-avoidance phenotypes previously reported for Pcdhγ-null SACs and Purkinje cells (Lefebvre et al. 2012). These results further highlight the functional significance of multicluster *Pcdh* expression during neural patterning of the mammalian central nervous system.

Pcdhαc2 Mediates Axonal Tiling of Serotonergic Neurons

Serotonin is a monoamine neurotransmitter that regulates a broad array of behaviors such as cardiovascular regulation, thermoregulation, pain sensitivity, respiration, sleep, appetite, cognition, and mood (Berger et al. 2009) (**Figure 5*a***). Abnormalities in serotonin levels and signaling are associated with multiple psychiatric disorders such as major depression, bipolar disorder, anxiety, aggression, schizophrenia, OCD, and ADHD (Olivier 2015). In humans, there are only 300,000 serotonergic neurons located in the raphe nuclei in the brain stem—a very small number compared to the billions of neurons that these serotonergic neurons influence. Serotonergic neurons, to accomplish their essential functions in neural modulation, project their axons throughout the entire CNS in a space-filling pattern, with their axon terminals diffusely distributed in various target areas in a tiled arrangement (Jacobs & Azmitia 1992, Kiyasova & Gaspar 2011). In most target regions, the serotonergic "synapses" are nonjunctional, lacking postsynaptic partners for target specificity. Instead, serotonin is released from these presynaptic varicosities by volume transmission, modulating neural activities by interacting with a variety of serotonin receptors (Daubert & Condron 2010). Since the functional concentration of serotonin is maintained only within a short distance from the site of release (Daubert & Condron 2010), a mechanism that controls even spacing between serotonergic axon terminals is required to prevent fluctuations in local serotonin levels.

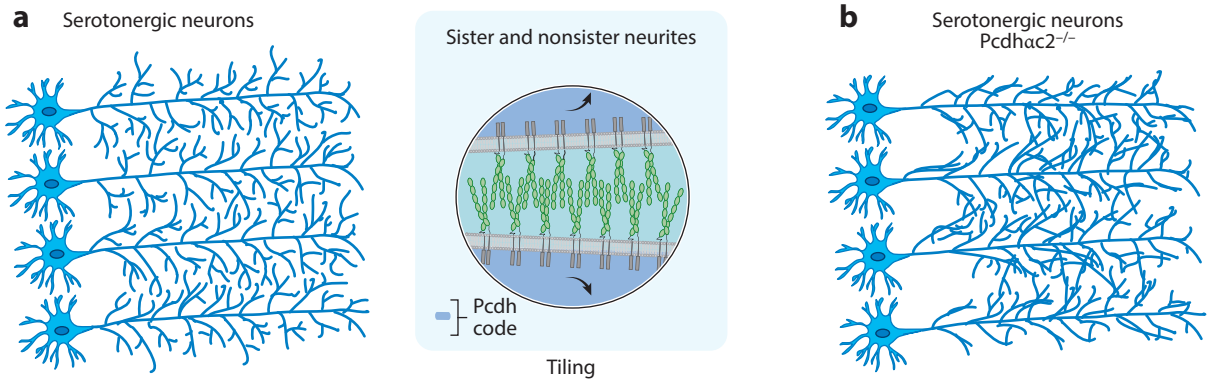


Figure 5

(*a*) A schematic illustration of neighboring serotonergic neurons (*left*). Each serotonergic neuron expresses the same Pcdh code (*Pcdhαc2*), which results in avoidance/repulsion between sister and nonsister neurites (*right*). The nature of this form of avoidance/repulsion is unknown, and this schematic merely represents the Pcdhαc2-mediated interactions at the cell surface. (*b*) Constitutive deletion of *Pcdhαc2* results in disruption of tiling between serotonergic axon terminals in target fields. Serotonergic neurons send out long-range axonal projections throughout the CNS and elaborate enormous axonal arbors; these schematic drawings are only to illustrate Pcdhac2-mediated axonal tiling (and the lack of it) between both sister and nonsister serotonergic axonal terminals in target fields.

Recent studies have shown that the *Pcdhα* gene cluster—specifically, a *Pcdhα* isoform, *Pcdhαc2*—plays a critical role in this process (Chen et al. 2017, Katori et al. 2017). While Pcdhα proteins are broadly expressed in the CNS, they are highly enriched in serotonergic neurons (Chen et al. 2017, Katori et al. 2009). In Pcdhα-deficient mice or severe hypomorphs, serotonergic axons failed to spread efficiently and formed clumps throughout the brain (Chen et al. 2017, Katori et al. 2009), and this phenotype is strictly cell autonomous to serotonergic neurons, as shown with conditional knockouts (Chen et al. 2017, Katori et al. 2017) (**Figure 5*b***). By using brainbow adeno-associated viral vectors (Cai et al. 2013) to differentially label individual serotonergic neurons and their axons, it was found that the mutant axons clump due to the disrupted tiling of the serotonergic axons originating from different neurons (Chen et al. 2017) (**Figure 5*b***). Surprisingly, genetic dissection of the entire Pcdh cluster revealed that both the *Pcdhβ* and *Pcdhγ* gene clusters, as well as all alternate *Pcdhα* isoforms, are not required for proper serotonergic wiring (Chen et al. 2017). Rather, the serotonergic axonal tiling defect was reproduced only when the two C-type *Pcdhα* genes (*Pcdhαc1* and *Pcdhαc2*) were deleted (Chen et al. 2017). Therefore, in contrast to the self-recognition of OSN axons that requires multicluster Pcdh diversity, a diversity-independent mechanism controls the tiling of serotonergic axons.

Fundamental insights into this mechanism were obtained by expression profiling of serotonergic neurons by TRAP-Seq and single-cell RNA-Seq experiments, which revealed that the only *Pcdhα* isoform expressed in serotonergic neurons is *Pcdhαc2* (Chen et al. 2017). In fact, out of the 58 Pcdh isoforms encoded in the mouse genome, only three C-type isoforms—Pcdhαc2, Pcdhγc3, and Pcdhγc4—were detected, revealing an intriguing regulatory mechanism that selectively activates the expression of C-type *Pcdhα* isoforms in serotonergic neurons. The specific requirement for *Pcdhαc2* alone for serotonergic wiring was subsequently confirmed by its selective deletion (C.L. Nwakeze, W.V. Chen & T. Maniatis, unpublished observations) and in a more recent study by Katori et al. (2017). Collectively, these results support a model in which the constitutive expression of *Pcdhαc2* in all serotonergic neurons provides a common cell surface Pcdh identity code

(**Figure 5*b***). The Pcdhαc2 protein isoforms in neighboring serotonergic axons interact with one another, which presumably leads to contact-dependent repulsion (**Figure 5*b***). This repulsion is necessary to ensure even spacing and appropriate coverage of serotonergic neuron target territories (Chen et al. 2017). Therefore, the role of *Pcdhαc2* is strikingly similar to that of *Dscam2* in the *Drosophila* nervous system (**Supplemental Text, Section 1**), as both function as cell type identity codes to mediate axonal tiling.

Supplemental Material >

The importance of serotonin as a neurotransmitter in the regulation of cognitive and affective functions is well established, as dysfunction of serotonin synthesis, transmission, and reuptake has been implicated in multiple psychiatric disorders, including depression (Olivier 2015). *Pcdh$\alpha^{-/-}$* mutant mice display several cognitive-affective defects, including reduced immobility in the tail suspension test and the force swim test and contextual fear conditioning (Chen et al. 2017, Katori et al. 2009). Similar to serotonergic wiring abnormalities, these depression-like behaviors were reproduced in serotonergic-specific conditional knockouts, which strongly suggest that these behavior defects likely resulted from serotonergic miswiring and consequently from the altered distribution of serotonin throughout the central nervous system (Chen et al. 2017).

CONCLUSIONS AND FUTURE PERSPECTIVES

The extraordinary genomic organization, stochastic single-cell expression, and multifaceted function of the mammalian Pcdh gene cluster in neural development have provided fundamental advances in understanding a wide range of important biological mechanisms. Areas of advance include genomic organization and evolution, chromatin structure and chromosomal nuclear organization, mechanisms of stochastic single-cell gene expression at the levels of enhancer-promoter interactions and RNA splicing, mechanisms for generating protein diversity by combinatorial *cis* dimerization and a novel structure-based mechanism of *trans* homophilic interactions, and finally the role of these highly specific protein-protein interactions at the cell surface in establishing the single-cell identities essential for neural circuit assembly.

Ongoing challenges include understanding the mechanism by which stochastic promoter choice occurs and how the promoter-proximal 5′ splice site is selectively recognized by the spliceosome. Does promoter choice occur at the level of selective DNA methylation or demethylation, by random CTCF/Cohesin binding followed by methylation of the unoccupied CTCF binding sites, or by some other mechanism? Particularly important problems to be addressed with single-cell technology and viral barcoding are understanding when stochastic promoter choice occurs during neuronal differentiation and whether promoter choice is maintained during cell division prior to terminal differentiation.

Recent studies of the expression of alternate and C-type Pcdh genes revealed that they are differentially regulated during development and in a neuronal cell type–specific manner. The most striking example of this, as mentioned above, is provided by serotonergic neurons, in which only a subset of C-type Pcdhs, and not the alternate exons, are expressed (Chen et al. 2017). The mechanisms by which this deterministic expression pattern appears in a specific cell type are not understood. Similarly, although the three C-type *Pcdhγ* genes are essential for neuronal survival (Chen et al. 2012), it is puzzling why only certain types of neurons are affected when these genes are deleted in all cells. Moreover, we do not know whether all three C-type isoforms function synergistically for this function, whether deletion of only one of the isoforms is required for this phenotype, or finally why more extensive apoptosis was observed in multicluster knockouts (Hasegawa et al. 2016, Ing-Esteves et al. 2018). A systematic nervous system–wide analysis of neuronal cell type–specific Pcdh gene expression should shed light on these problems and provide new insights into the function of Pcdh genes in neural development.

With the recent development of methods for single-cell RNA sequencing, detection of chromosome modifications at the single-cell level, and in situ sequencing, it is now possible to probe deeply into cell-specific patterns of Pcdh gene expression throughout the nervous system. Moreover, the generation of single, double, and tricluster Pcdh gene knockout mice (Hasegawa et al. 2016, 2017; Mountoufaris et al. 2017) sets the stage for systematic nervous system–wide studies of Pcdh function in different neuronal cell types. These efforts will be enabled by advances in single-neuron labeling and mutagenesis methods (Espinosa et al. 2014), superresolution imaging methods (Bates et al. 2013), and whole-brain clearing methods (Chung & Deisseroth 2013, Renier et al. 2014). Taken together, studies of the Pcdh gene cluster are likely to provide fundamental insights into complex mechanisms of neural circuit development, and these studies are likely to contribute significantly to the international BRAIN initiative (Bargmann & Newsome 2014) and the Human Cell Atlas Project (Regev et al. 2017).

Whereas expression of the *Pcdhα* gene cluster appears to be specific to the nervous system, the *Pcdhβ* and *Pcdhγ* gene clusters are expressed in other tissues, such as the kidney and liver, and are thus likely to have nonneuronal functions. For example, aberrant DNA methylation in the Pcdh gene cluster has been reported to be associated with cancers of various tissues and organs (El Hajj et al. 2017). However, the mechanisms by which Pcdhs are expressed in nonneuronal cell types and how such Pcdhs function remain to be explored.

A key question of great importance is the role of the Pcdh gene cluster in human neurological diseases. *Pcdhα* gene cluster mutations in mice result in abnormalities in cognitive and affective functions such as depression (Chen et al. 2017). In humans, both DNA sequence variants and changes in DNA methylation in the Pcdh gene cluster have been implicated in a variety of neurodevelopmental or psychiatric disorders (El Hajj et al. 2017, Hirabayashi & Yagi 2014). However, definitive evidence for a causal relationship has not been demonstrated. Among the strongest genetic evidence for a connection to neuropsychiatric disease is the association with autism spectrum disorders, identified through large-scale family studies (**https://www.sfari.org/resource/simons-simplex-collection/**). Further exploration of the relationship between DNA sequence variants, DNA methylation, and expression of the clustered Pcdhs is likely to be a priority of future studies aimed at understanding the role of the clustered Pcdhs in human diseases.

Perhaps foremost among unsolved problems of the clustered Pcdhs is an understanding the mechanisms by which specific homophilic interactions at the cell surface lead to neurite repulsion. Sparse clues are available, such as the cell surface cleavage of Pcdhs by γ-secretase and metalloprotease, a process that occurs during the development of the nervous system. Pcdh cleavage requires endocytosis and results in the release of the intracellular domain into the cytoplasm, a mechanism that is strikingly similar to that required for Notch intracellular signaling (Buchanan et al. 2010). In addition, endocytosis and endolysosomal trafficking of Pcdh complexes have been proposed to be a part of the mechanism by which Pcdhs convert from adhesive to avoidance molecules (Phillips et al. 2017). Moreover, a role for intracellular signaling was also suggested by the isolation of high-molecular-weight complexes containing Pcdh signaling molecules and kinases and the phosphorylation of Pcdhs by associated kinases (Han et al. 2010, Mah & Weiner 2017, Schalm et al. 2010, Suo et al. 2012). The relationship between these observations; the mechanisms of Pcdh cleavage, trafficking, and intracellular signaling; and neurite repulsion remains to be established.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Bargmann CI, Newsome WT. 2014. The Brain Research Through Advancing Innovative Neurotechnologies (BRAIN) initiative and neurology. *JAMA Neurol.* 71:675–76

Bates M, Jones SA, Zhuang X. 2013. Stochastic optical reconstruction microscopy (STORM): a method for superresolution fluorescence imaging. *Cold Spring Harb. Protoc.* 2013:498–520

Berger M, Gray JA, Roth BL. 2009. The expanded biology of serotonin. *Annu. Rev. Med.* 60:355–66

Bressel OC, Khan M, Mombaerts P. 2016. Linear correlation between the number of olfactory sensory neurons expressing a given mouse odorant receptor gene and the total volume of the corresponding glomeruli in the olfactory bulb. *J. Comp. Neurol.* 524:199–209

Buchanan SM, Schalm SS, Maniatis T. 2010. Proteolytic processing of protocadherin proteins requires endocytosis. *PNAS* 107:17774–79

Cai D, Cohen KB, Luo T, Lichtman JW, Sanes JR. 2013. Improved tools for the Brainbow toolbox. *Nat. Methods* 10:540–47

Chen K, Hu J, Moore DL, Liu R, Kessans SA, et al. 2015. Genome-wide binding and mechanistic analyses of Smchd1-mediated epigenetic regulation. *PNAS* 112:E3535–44

Chen WV, Alvarez FJ, Lefebvre JL, Friedman B, Nwakeze C, et al. 2012. Functional significance of isoform diversification in the protocadherin gamma gene cluster. *Neuron* 75:402–9

Chen WV, Maniatis T. 2013. Clustered protocadherins. *Development* 140:3297–302

Chen WV, Nwakeze CL, Denny CA, O'Keeffe S, Rieger MA, et al. 2017. Pcdhαc2 is required for axonal tiling and assembly of serotonergic circuitries in mice. *Science* 356:406–11

Chung K, Deisseroth K. 2013. CLARITY for mapping the nervous system. *Nat. Methods* 10:508–13

Daubert EA, Condron BG. 2010. Serotonin: a regulator of neuronal morphology and circuitry. *Trends Neurosci.* 33:424–34

El Hajj N, Dittrich M, Haaf T. 2017. Epigenetic dysregulation of protocadherins in human disease. *Semin. Cell Dev. Biol.* 69:172–82

Espinosa JS, Tea JS, Luo L. 2014. Mosaic analysis with double markers (MADM) in mice. *Cold Spring Harb. Protoc.* 2014:182–89

Esumi S, Kakazu N, Taguchi Y, Hirayama T, Sasaki A, et al. 2005. Monoallelic yet combinatorial expression of variable exons of the protocadherin-α gene cluster in single neurons. *Nat. Genet.* 37:171–76

Garrett AM, Schreiner D, Lobas MA, Weiner JA. 2012. γ-Protocadherins control cortical dendrite arborization by regulating the activity of a FAK/PKC/MARCKS signaling pathway. *Neuron* 74:269–76

Golan-Mashiach M, Grunspan M, Emmanuel R, Gibbs-Bar L, Dikstein R, Shapiro E. 2012. Identification of CTCF as a master regulator of the clustered protocadherin genes. *Nucleic Acids Res.* 40:3378–91

Goodman KM, Rubinstein R, Dan H, Bahna F, Mannepalli S, et al. 2017. Protocadherin *cis*-dimer architecture and recognition unit diversity. *PNAS* 114:E9829–37

Goodman KM, Rubinstein R, Thu CA, Bahna F, Mannepalli S, et al. 2016a. Structural basis of diverse homophilic recognition by clustered α- and β-protocadherins. *Neuron* 90:709–23

Goodman KM, Rubinstein R, Thu CA, Mannepalli S, Bahna F, et al. 2016b. γ-Protocadherin structural diversity and functional implications. *eLife* 5:e20930

Guo Y, Monahan K, Wu H, Gertz J, Varley KE, et al. 2012. CTCF/cohesin-mediated DNA looping is required for protocadherin α promoter choice. *PNAS* 109:21081–86

Guo Y, Xu Q, Canzio D, Shou J, Li J, et al. 2015. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* 162:900–10

Han MH, Lin C, Meng S, Wang X. 2010. Proteomics analysis reveals overlapping functions of clustered protocadherins. *Mol. Cell. Proteom.* 9:71–83

Hasegawa S, Hamada S, Kumode Y, Esumi S, Katori S, et al. 2008. The protocadherin-α family is involved in axonal coalescence of olfactory sensory neurons into glomeruli of the olfactory bulb in mouse. *Mol. Cell. Neurosci.* 38:66–79

Hasegawa S, Kobayashi H, Kumagai M, Nishimaru H, Tarusawa E, et al. 2017. Clustered protocadherins are required for building functional neural circuits. *Front. Mol. Neurosci.* 10:114

Hasegawa S, Kumagai M, Hagihara M, Nishimaru H, Hirano K, et al. 2016. Distinct and cooperative functions for the protocadherin-α, -β and -γ clusters in neuronal survival and axon targeting. *Front. Mol. Neurosci.* 9:155

Hattori D, Millard SS, Wojtowicz WM, Zipursky SL. 2008. Dscam-mediated cell recognition regulates neural circuit formation. *Annu. Rev. Cell Dev. Biol.* 24:597–620
Hirabayashi T, Yagi T. 2014. Protocadherins in neurological diseases. *Adv. Neurobiol.* 8:293–314
Hirano K, Kaneko R, Izawa T, Kawaguchi M, Kitsukawa T, Yagi T. 2012. Single-neuron diversity generated by Protocadherin-β cluster in mouse central and peripheral nervous systems. *Front. Mol. Neurosci.* 5:90
Hirayama T, Tarusawa E, Yoshimura Y, Galjart N, Yagi T. 2012. CTCF is required for neural development and stochastic expression of clustered Pcdh genes in neurons. *Cell Rep.* 2:345–57
Hirayama T, Yagi T. 2017. Regulation of clustered protocadherin genes in individual neurons. *Semin. Cell Dev. Biol.* 69:122–30
Hummel T, Vasconcelos ML, Clemens JC, Fishilevich Y, Vosshall LB, Zipursky SL. 2003. Axonal targeting of olfactory receptor neurons in *Drosophila* is controlled by Dscam. *Neuron* 37:221–31
Ing-Esteves S, Kostadinov D, Marocha J, Sing AD, Joseph KS, et al. 2018. Combinatorial effects of alpha- and gamma-protocadherins on neuronal survival and dendritic self-avoidance. *J. Neurosci.* 38:2713–29
Isbel L, Prokopuk L, Wu H, Daxinger L, Oey H, et al. 2016. Wiz binds active promoters and CTCF-binding sites and is required for normal behaviour in the mouse. *eLife* 5:e15082
Jacobs BL, Azmitia EC. 1992. Structure and function of the brain serotonin system. *Physiol. Rev.* 72:165–229
Jiang Y, Loh YE, Rajarajan P, Hirayama T, Liao W, et al. 2017. The methyltransferase SETDB1 regulates a large neuron-specific topological chromatin domain. *Nat. Genet.* 49:1239–50
Kai N, Mishina M, Yagi T. 1997. Molecular cloning of Fyn-associated molecules in the mouse central nervous system. *J. Neurosci. Res.* 48:407–24
Kaneko R, Kato H, Kawamura Y, Esumi S, Hirayama T, et al. 2006. Allelic gene regulation of Pcdh-α and Pcdh-γ clusters involving both monoallelic and biallelic expression in single Purkinje cells. *J. Biol. Chem.* 281:30551–60
Katori S, Hamada S, Noguchi Y, Fukuda E, Yamamoto T, et al. 2009. Protocadherin-α family is required for serotonergic projections to appropriately innervate target brain areas. *J. Neurosci.* 29:9137–47
Katori S, Noguchi-Katori Y, Okayama A, Kawamura Y, Luo W, et al. 2017. Protocadherin-αC2 is required for diffuse projections of serotonergic axons. *Sci. Rep.* 7:15908
Kawaguchi M, Toyama T, Kaneko R, Hirayama T, Kawamura Y, Yagi T. 2008. Relationship between DNA methylation states and transcription of individual isoforms encoded by the protocadherin-α gene cluster. *J. Biol. Chem.* 283:12064–75
Kehayova P, Monahan K, Chen W, Maniatis T. 2011. Regulatory elements required for the activation and repression of the protocadherin-α gene cluster. *PNAS* 108:17195–200
Kiyasova V, Gaspar P. 2011. Development of raphe serotonin neurons from specification to guidance. *Eur. J. Neurosci.* 34:1553–62
Kohmura N, Senzaki K, Hamada S, Kai N, Yasuda R, et al. 1998. Diversity revealed by a novel family of cadherins expressed in neurons at a synaptic complex. *Neuron* 20:1137–51
Kostadinov D, Sanes JR. 2015. Protocadherin-dependent dendritic self-avoidance regulates neural connectivity and circuit function. *eLife* 4:e08964
Ledderose J, Dieter S, Schwarz MK. 2013. Maturation of postnatally generated olfactory bulb granule cells depends on functional γ-protocadherin expression. *Sci. Rep.* 3:1514
Lefebvre JL. 2017. Neuronal territory formation by the atypical cadherins and clustered protocadherins. *Semin. Cell Dev. Biol.* 69:111–21
Lefebvre JL, Kostadinov D, Chen WV, Maniatis T, Sanes JR. 2012. Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature* 488:517–21
Lefebvre JL, Zhang Y, Meister M, Wang X, Sanes JR. 2008. γ-Protocadherins regulate neuronal survival but are dispensable for circuit formation in retina. *Development* 135:4141–51
Li SA, Cheng L, Yu Y, Chen Q. 2016. Structural basis of Dscam1 homodimerization: insights into context constraint for protein recognition. *Sci. Adv.* 2:e1501118
Mah KM, Weiner JA. 2017. Regulation of Wnt signaling by protocadherins. *Semin. Cell Dev. Biol.* 69:158–71
Maiuri P, Knezevich A, De Marco A, Mazza D, Kula A, et al. 2011. Fast transcription rates of RNA polymerase II in human cells. *EMBO Rep.* 12:1280–85
Mason AG, Slieker RC, Balog J, Lemmers R, Wong CJ, et al. 2017. SMCHD1 regulates a limited set of gene clusters on autosomal chromosomes. *Skelet. Muscle* 7:12

Mombaerts P, Wang F, Dulac C, Chao SK, Nemes A, et al. 1996. Visualizing an olfactory sensory map. *Cell* 87:675–86

Monahan K, Rudnick ND, Kehayova PD, Pauli F, Newberry KM, et al. 2012. Role of CCCTC binding factor (CTCF) and cohesin in the generation of single-cell diversity of protocadherin-α gene expression. *PNAS* 109:9125–30

Mountoufaris G, Chen WV, Hirabayashi Y, O'Keeffe S, Chevee M, et al. 2017. Multicluster Pcdh diversity is required for mouse olfactory neural circuit assembly. *Science* 356:411–14

Murata Y, Hamada S, Morishita H, Mutoh T, Yagi T. 2004. Interaction with protocadherin-γ regulates the cell surface expression of protocadherin-α. *J. Biol. Chem.* 279:49508–16

Nicoludis JM, Lau SY, Scharfe CP, Marks DS, Weihofen WA, Gaudet R. 2015. Structure and sequence analyses of clustered protocadherins reveal antiparallel interactions that mediate homophilic specificity. *Structure* 23:2087–98

Nicoludis JM, Vogt BE, Green AG, Scharfe CP, Marks DS, Gaudet R. 2016. Antiparallel protocadherin homodimers use distinct affinity- and specificity-mediating regions in cadherin repeats 1–4. *eLife* 5:e18449

Noonan JP, Grimwood J, Schmutz J, Dickson M, Myers RM. 2004. Gene conversion and the evolution of protocadherin gene cluster diversity. *Genome Res.* 14:354–66

Olivier B. 2015. Serotonin: a never-ending story. *Eur. J. Pharmacol.* 753:2–18

Peek SL, Mah KM, Weiner JA. 2017. Regulation of neural circuit formation by protocadherins. *Cell. Mol. Life Sci.* 74:4133–57

Phillips GR, LaMassa N, Nie YM. 2017. Clustered protocadherin trafficking. *Semin. Cell Dev. Biol.* 69:131–39

Prasad T, Wang X, Gray PA, Weiner JA. 2008. A differential developmental pattern of spinal interneuron apoptosis during synaptogenesis: insights from genetic analyses of the protocadherin-γ gene cluster. *Development* 135:4153–64

Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159:1665–80

Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, et al. 2017. The Human Cell Atlas. *eLife* 6:e27041

Renier N, Wu Z, Simon DJ, Yang J, Ariel P, Tessier-Lavigne M. 2014. iDISCO: a simple, rapid method to immunolabel large tissue samples for volume imaging. *Cell* 159:896–910

Ribich S, Tasic B, Maniatis T. 2006. Identification of long-range regulatory elements in the protocadherin-α gene cluster. *PNAS* 103:19719–24

Rubinstein R, Goodman KM, Maniatis T, Shapiro L, Honig B. 2017. Structural origins of clustered protocadherin-mediated neuronal barcoding. *Semin. Cell Dev. Biol.* 69:140–50

Rubinstein R, Thu CA, Goodman KM, Wolcott HN, Bahna F, et al. 2015. Molecular logic of neuronal self-recognition through protocadherin domain interactions. *Cell* 163:629–42

Sano K, Tanihara H, Heimark RL, Obata S, Davidson M, et al. 1993. Protocadherins: a large family of cadherin-related molecules in central nervous system. *EMBO J.* 12:2249–56

Schalm SS, Ballif BA, Buchanan SM, Phillips GR, Maniatis T. 2010. Phosphorylation of protocadherin proteins by the receptor tyrosine kinase Ret. *PNAS* 107:13894–99

Schmucker D, Clemens JC, Shu H, Worby CA, Xiao J, et al. 2000. *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* 101:671–84

Schreiner D, Weiner JA. 2010. Combinatorial homophilic interaction between γ-protocadherin multimers greatly expands the molecular diversity of cell adhesion. *PNAS* 107:14893–98

Suo L, Lu H, Ying G, Capecchi MR, Wu Q. 2012. Protocadherin clusters and cell adhesion kinase regulate dendrite complexity through Rho GTPase. *J. Mol. Cell Biol.* 4:362–76

Suzuki ST. 1996. Structural and functional diversity of cadherin superfamily: are new members of cadherin superfamily involved in signal transduction pathway? *J. Cell. Biochem.* 61:531–42

Takeuchi H, Sakano H. 2014. Neural map formation in the mouse olfactory system. *Cell. Mol. Life Sci.* 71:3049–57

Tarusawa E, Sanbo M, Okayama A, Miyashita T, Kitsukawa T, et al. 2016. Establishment of high reciprocal connectivity between clonal cortical neurons is regulated by the Dnmt3b DNA methyltransferase and clustered protocadherins. *BMC Biol.* 14:103

Tasic B, Nabholz CE, Baldwin KK, Kim Y, Rueckert EH, et al. 2002. Promoter choice determines splice site selection in protocadherin α and γ pre-mRNA splicing. *Mol. Cell* 10:21–33

Thu CA, Chen WV, Rubinstein R, Chevee M, Wolcott HN, et al. 2014. Single-cell identity generated by combinatorial homophilic interactions between α, β, and γ protocadherins. *Cell* 158:1045–59

Toyoda S, Kawaguchi M, Kobayashi T, Tarusawa E, Toyama T, et al. 2014. Developmental epigenetic modification regulates stochastic expression of clustered protocadherin genes, generating single neuron diversity. *Neuron* 82:94–108

Wang X, Su H, Bradley A. 2002a. Molecular mechanisms governing Pcdh-γ gene expression: evidence for a multiple promoter and *cis*-alternative splicing model. *Genes Dev*. 16:1890–905

Wang X, Weiner JA, Levi S, Craig AM, Bradley A, Sanes JR. 2002b. Gamma protocadherins are required for survival of spinal interneurons. *Neuron* 36:843–54

Weiner JA, Wang X, Tapia JC, Sanes JR. 2005. Gamma protocadherins are required for synaptic development in the spinal cord. *PNAS* 102:8–14

Wojtowicz WM, Flanagan JJ, Millard SS, Zipursky SL, Clemens JC. 2004. Alternative splicing of *Drosophila* Dscam generates axon guidance receptors that exhibit isoform-specific homophilic binding. *Cell* 118:619–33

Wojtowicz WM, Wu W, Andre I, Qian B, Baker D, Zipursky SL. 2007. A vast repertoire of Dscam binding specificities arises from modular interactions of variable Ig domains. *Cell* 130:1134–45

Wu Q, Maniatis T. 1999. A striking organization of a large family of human neural cadherin-like cell adhesion genes. *Cell* 97:779–90

Wu Q, Zhang T, Cheng JF, Kim Y, Grimwood J, et al. 2001. Comparative DNA sequence analysis of mouse and human protocadherin gene clusters. *Genome Res*. 11:389–404

Yagi T. 2008. Clustered protocadherin family. *Dev. Growth Differ.* 50(Suppl. 1):131–40

Yin M, Wang J, Wang M, Li X, Zhang M, et al. 2017. Molecular mechanism of directional CTCF recognition of a diverse range of genomic sites. *Cell Res*. 27:1365–77

Yokota S, Hirayama T, Hirano K, Kaneko R, Toyoda S, et al. 2011. Identification of the cluster control region for the protocadherin-β genes located beyond the protocadherin-γ cluster. *J. Biol. Chem.* 286:31885–95

Yu WP, Rajasegaran V, Yew K, Loh WL, Tay BH, et al. 2008. Elephant shark sequence reveals unique insights into the evolutionary history of vertebrate genes: a comparative analysis of the protocadherin cluster. *PNAS* 105:3819–24

Yu WP, Yew K, Rajasegaran V, Venkatesh B. 2007. Sequencing and comparative analysis of fugu protocadherin clusters reveal diversity of protocadherin genes among teleosts. *BMC Evol. Biol.* 7:49

Zhu H, Hummel T, Clemens JC, Berdnik D, Zipursky SL, Luo L. 2006. Dendritic patterning by Dscam and synaptic partner matching in the *Drosophila* antennal lobe. *Nat. Neurosci.* 9:349–55

Zipursky SL, Grueber WB. 2013. The molecular basis of self-avoidance. *Annu. Rev. Neurosci.* 36:547–68