Seminars in Cell & Developmental Biology 69 (2017) 172–182

Contents lists available at ScienceDirect

Seminars in Cell & Developmental Biology

journal homepage: www.elsevier.com/locate/semcdb

Review

Epigenetic dysregulation of protocadherins in human disease

Nady El Hajj*, Marcus Dittrich, Thomas Haaf

Institute of Human Genetics, Julius Maximilians University, Biozentrum, Am Hubland, Würzburg, Germany

ARTICLE INFO

Article history:
Received 14 March 2017
Received in revised form 3 July 2017
Accepted 6 July 2017
Available online 8 July 2017

Keywords:
Protocadherins
DNA methylation
Brain disorders
Epigenetics

ABSTRACT

Protocadherins (Pcdhs) are a group of cell-cell adhesion molecules that are highly expressed in the nervous system and have a major function in dendrite development and neural circuit formation. However, the role protocadherins play in human health and disease remains unclear. Several recent studies have associated epigenetic dysregulation of protocadherins with possible implications for disease pathogenesis. In this review, we briefly recap the various epigenetic mechanisms regulating protocadherin genes, particularly the clustered *Pcdh*s. We further outline research describing altered epigenetic regulation of protocadherins in neurological and psychiatric disorders, as well as in cancer and during aging. We additionally present preliminary data on DNA methylation dynamics of clustered protocadherins during fetal brain development, as well as the epigenetic differences distinguishing adult neuronal and glial cells. A deeper understanding of the role of protocadherins in disease is crucial for designing novel diagnostic tools and therapies targeting brain disorders.

Contents

Abbreviations: Pcdhs, Protocadherins; VE, variable exon; CSE, conserved sequence element; SSE, sequence specific element; CTCF, CCTC binding factor; NSRE, neuron-restrictive silencer element; DNMT3B, DNA methyltransferase 3B; GWAS, genome-wide association study; DS, Down syndrome; WBS, Williams-Beuren syndrome; Dup7, 7q11.23 duplication syndrome; CNV, copy number variation; MZ, monozygotic; ME, metastable epiallele; HPA, hypothalamic–pituitary–adrenal axis; PAE, prenatal alcohol exposure; FASD, fetal alcohol spectrum disorder; ASD, autism spectrum disorders; MeCP2, methyl-CpG-binding protein 2; DMRs, differential methylated regions; CGI, CpG island.

* Corresponding author at: Institute of Human Genetics, Julius-Maximilians-Universität Würzburg, Biozentrum, Am Hubland, 97074 Würzburg, Germany.
E-mail address: nady.elhajj@uni-wuerzburg.de (N. El Hajj).

http://dx.doi.org/10.1016/j.semcdb.2017.07.007

1. Introduction

The term "epigenetics" was coined by Conrad Waddington in 1942 and means "on top of" or "over" genetics. Every normal somatic cell of a human individual contains the same genome consisting of approximately 3 billion base pairs. However, a plethora of epigenetic modifications regulates the accessibility of genes in a given cell type and developmental stage to transcriptional machinery. Epigenetic information in the form of DNA methylation patterns and histone modifications orchestrates gene regulation that is highly coordinated both temporally and spatially [1,2]. It is estimated that during ontogeny each individual is endowed with more than 100 epigenomes [3]. Once established, epigenetic modifications can be stably transmitted to daughter cells during mitotic cell divisions and potentially during meiosis from one generation to the next [4].

The most prominent epigenetic modification of DNA itself is methylation, which occurs at the carbon 5′ position of cytosine, mainly in the context of CpG dinucleotides [5]. Non-coding regions of the genome are usually methylated to prevent retrotransposition causing genome instability, whereas most CpG islands and promoter regions are demethylated [6]. Promoter methylation during differentiation, development, or disease processes is associated with transcriptional silencing [7,8]. The epigenome is highly plastic during genome-wide reprogramming in the germline and early embryogenesis. Although during ontogeny this plasticity is steadily decreasing, the epigenome(s) of an individual can still be influenced by internal and external (*i.e.* environmental) cues [9–11]. One important hallmark of epigenetic modifications is their considerable variability among normal individuals, which is one or two orders of magnitude higher than genetic variation [12,13].

This review focusses on the epigenetics of protocadherin (*PCDH*) genes and the possible disease-relevance of epigenetic alterations. Pcdhs are a group of cell-cell adhesion proteins that belong to the cadherin family and are predominantly expressed in the nervous system [14,15]. They play a major role in neural circuit formation in the vertebrate nervous system during development, especially in regulating dendrite branching [14,16]. The family is divided into the clustered and non-clustered Pcdhs [16]. Kohmura et al. were the first to characterize a novel family of clustered protocadherins in the mouse designated as cadherin-related neuronal receptor (CNR) [15]. Wu and Maniatis analysed the human counterparts of the mouse protocadherins and discovered 52 novel human cadherin-like genes [17]. They identified a one-megabase region on human chromosome 5q31 endowed with three clusters containing the tandemly arrayed *PCDHA*, *PCDHB*, and *PCDHG* genes [17]. The clustered *PCDH*s are thought to generate unique single-cell identifiers (barcodes) for self-recognition in the vertebrate nervous system [18]. Individual neurons interact with each other *via* multi-synaptic connections in a complex neural network. During development, neurites (*i.e.* axons and dendrites) travel great distances to find and connect with their target neurons, while at the same time avoiding connections with sister neurites from their own cell body. Pcdhs are known to mediate homophilic matching between different cell types leading to the establishment of complex neural circuits in the brain. This has been particularly observed for the γ-Pcdh cluster between neurons [19,20], between neurons and astrocytes [21], and in cells transfected with individual *Pcdhg* isoforms [22]. Gene knockout studies have demonstrated that α-Pcdh- and γ-Pcdh have crucial roles in neuron survival [23], maturation of serotonergic projections [24], as well as in synaptic development and connectivity [25].

Reminiscent of the genomic architecture of the immunoglobulin and T cell receptor gene clusters [26], the paralogous genes in the *PCDHA*, *PCDHB*, and *PCDHG* clusters can generate substantial diversity [17]. However, unlike immunoglobulin and T cell receptor diversity which is the result of genetic recombination, diversity of clustered *PCDH*s relies on epigenetic regulation of promoter choice and alternative transcripts [27].

In the mouse, each of the 58 Pcdh proteins is composed of a variable intracellular and extracellular domain encoded by a large variable exon (VE) [28]. The diverse extracellular domains play a role in self-recognition *via* homophilic cell-cell interactions whereas the intracellular domain is involved in cell signaling [22,29–31]. The α-*Pcdh* and γ-*Pcdh* clusters share a constant intracellular domain encoded by three common exons downstream of the cluster [17,28]. Some of these variable cytoplasmic domains are required for trafficking and organelle tabulation, as well as for the regulation of the WNT canonical pathway [32,33]. Each of the 14 VEs in *Pcdha* and the 22 VEs in *Pcdh*g has their own upstream promoter. Promoter choice determines which exon is transcribed [26,34,35], leading to cell-specific expression [36]. All promoters include a 20 bp sequence specific element (SSE) that is upstream of a 22 bp conserved sequence element (CSE). Both CSE and SSE, are crucial for transcription and are controlled by CCTC binding factor (CTCF) that acts as a positive regulator [37]. Here, CTCF binding and promoter activity are controlled by promoter methylation [38]. In addition, there are long-distance enhancers that can increase the probability of transcription. Two cis-regulatory elements, HS7 and HS5-1, in the *Pcdha* cluster function as enhancers in the brain [34]. HS5-1 contains sequence elements for CTCF, as well as for the neuron-restrictive silencer element (NSRE) [39]. In addition, HS16-20 downstream of *Pcdh*g is involved in regulating the expression of *Pcdhb* and to a lesser extent a few *Pcdhg* genes [40]. In each neuron, the Pcdh clusters exhibit monoallelic and combinatorial gene regulation, generating individual patterns of cell identity on the neuronal surface [41]. Gene regulation mechanisms within the *Pcdh* cluster involve both random and constitutive expression [41]. Most genes are differentially expressed apart from the 2 C-type isoforms in the *Pcdha* cluster (*αC1* and *αC2*), as well as the 3 C-type isoforms in the *Pcdhg* cluster (*γC3*, *γC4*, and *γC5*) which are constitutively expressed across all neurons [36,41,42]. The location of a variable exon within the *Pcdha* cluster defines the manner by which it is expressed (*i.e.* stochastically or constitutively), as well as its CpG methylation level [43]. Furthermore, deletion and duplication mutants showed that the *Pcdha* gene cluster is regulated as a unit where gene duplications caused a decrease in gene expression whereas gene deletions were largely compensated for *in vivo* [44].

Genomic organization of the *PCDHA* and *PCDHG* clusters in humans and mice are very similar, whereas the human *PCDHB* cluster contains fewer genes than its mouse orthologue [28]. In both species, a high density of CpG sites is found upstream of the VE transcriptional start site [28]. VE expression is controlled by DNA methylation; the promoters of all non-transcribed VEs are heavily methylated [35]. In neuroblastoma cell lines expressing specific isoform combinations, transcriptional activity correlated with the methylation state of both the promoter and 5′ upstream region of the first exon [35,45]. Luciferase assays showed that treatment with the demethylating drug 5-azacytidine increased transcription of all isoforms, whereas promoter hypermethylation led to gene silencing [45]. The isoform-specific promoters are methylated by DNA methyltransferase 3B (DNMT3B) during early embryonic development, regulating the stochastic expression of *Pcdh* isoforms in single cells [27]. Recently, it has been shown that postnatal establishment of lineage-specific neural connections is predetermined by *Dnmt3b* and clustered *Pcdh*s during embryonal development [46]. Additional epigenetic modifiers are known to regulate expression of clustered *PCDH*s. For example, the repressor *Smchd1* (structural maintenance of chromosomes flexible hinge domain containing 1) functionally interacts with CTCF and is essential for methylation of the *Pcdha* and