informa
healthcare

Genetic Basis of Neuronal Individuality in the Mammalian Brain

Takeshi Yagi

KOKORO-Biology Group, Laboratories for Integrated Biology, Graduate School of Frontier Biosciences, Osaka University,
Osaka, Japan; and Japan Science and Technology Agency–Core Research for Evolutional Science and Technology (CREST),
Yamadaoka, Suita, Japan

Abstract: The mammalian brain is a complex multicellular system involving enormous numbers of neurons. The neuron is the basic functional unit of the brain, and neurons are organized by specialized intercellular connections into circuits with many other neurons. Physiological studies have revealed that individual neurons have remarkably selective response properties, and this individuality is a fundamental requirement for building complex and functionally diverse neural networks. Recent molecular biological studies have revealed genetic bases for neuronal individuality in the mammalian brain. For example, in the rodent olfactory epithelium, individual olfactory neurons express only one type of odorant receptor (OR) out of the over 1000 ORs encoded in the genome. The expressed OR determines the neuron's selective chemosensory response and specifies its axonal targeting to a particular olfactory glomerulus in the olfactory bulb. Neuronal diversity can also be generated in individual cells by the independent and stochastic expression of autosomal alleles, which leads to functional heterozygosity among neurons. Among the many genes that show autosomal stochastic monoallelic expression, approximately 50 members of the clustered protocadherins (Pcdhs) are stochastically expressed in individual neurons in distinct combinations. The clustered Pcdhs belong to a large subfamily of the cadherin superfamily of homophilic cell-adhesion proteins. Loss-of-function analyses show that the clustered Pcdhs have critical functions in the accuracy of axonal projections, synaptic formation, dendritic arborization, and neuronal survival. In addition, *cis*-tetramers, composed of heteromultimeric clustered Pcdh members, represent selective binding units for cell-cell interactions, and provide exponential numbers of possible cell-surface relationships between individual neurons. The extensive molecular diversity of neuronal cell-surface proteins affects neurons' individual properties and connectivities. The molecular features of the diverse clustered Pcdh molecules suggest that they provide a genetic basis for neuronal individuality and appropriate neuronal wiring in the brain.

Keywords: epigenetic, neural network, protocadherin, stochastic expression

INTRODUCTION

The mammalian brain contains an enormous number of neurons and complex topological neural networks, in which differences at the level of individual neurons play important roles. The neuron is the basic functional unit of the complex brain system. Neuronal individuality was demonstrated physiologically by the remarkably selective response properties of single neurons in specific functional situations and is related to the specific patterns of synaptic connections made by the huge numbers of neurons that form functional circuits. That is, the organization of neuronal networks, which is directed in part by the properties of individual neurons, determines the processing flow of information in neural circuits and has important implications for regulating brain functions. Recent physiological studies demonstrated that the bidirectional and highly

clustered connections that typify local neuronal networks exhibit specific features more frequently than would be expected if the circuits formed randomly (Markram, 1997; Kalisman et al., 2005; Song et al., 2005; Yoshimura et al., 2005; Koulakov et al., 2009; Yassin et al., 2010; Perin et al., 2011). Furthermore, individual neurons with similar response properties are highly selective in their synaptic connections and show preferential connectivity (Yoshimura & Callaway, 2005). These networks are characterized not only by their complex topology, but also by their non-Gaussian (long-tail) distributions of synaptic weight in the cortex and hippocampus (Song et al., 2005; Lefort et al., 2009; Ikegaya et al., 2013). In theoretical model, the long-tail distributions of connection weights can generate noise useful for associative memory (Hiratani et al., 2013). Recent research shows that cell lineage plays a key role in specifying neuronal connections

and functional properties in the cortex (Yu et al., 2009, 2012; Li et al., 2012; Ohtsuki et al., 2012). Therefore, the establishment of neuronal individuality, which helps determine neuronal connectivity and physiological function during development, is a fundamental mechanism underlying the building of complex neural networks in the brain.

At the molecular level, the mechanism of stochastic monoallelic expression generates cellular individuality and provides potential functional variation among the individual cells of a complex system. For example, the immune system is a complex system in which enormous numbers of diverse cells are developmentally generated by the random DNA rearrangement of immune receptors (Tonegawa, 1983; Lieber, 1992). This system can newly acquire specific immune responses to a nearly limitless number of antigens that attack an animal, and "memorize" the antigen. Thus, immune cell individuality is generated by the expression of a specific immune receptor, random somatic DNA rearrangements, the combination of heteromeric receptors, and selection by the specific response to an attacking antigen. This molecular mechanism that regulates the expression and function of immune receptors constitutes the predetermined immune memory system.

As another example, diverse olfactory receptors are expressed in the olfactory sensory system. In 1991, Buck and Axel discovered over 1000 odorant receptors (ORs) in olfactory sensory neurons in the olfactory epithelium in mice (Buck & Axel, 1991). This discovery led to further investigation of the complex olfactory system, and the individuality of olfactory sensory neurons. These investigations showed that an olfactory sensory neuron can express only one functional *OR* gene, and the specific OR protein expressed suppresses the expression of other *OR* genes and induces the neuron's axonal targeting of the appropriate glomerulus in the olfactory bulb by topological mapping. Thus, the individuality of olfactory sensory neurons is determined genetically by the OR expression and function (Sakano, 2010).

Thus, in both the immune and olfactory systems, cellular individuality is generated by the expression of diverse receptor-type molecules, and involves random allelic expression and specific functional properties. The elucidation of these mechanisms has been invaluable for increasing our understanding of these complex systems.

In humans, at the single-cell level, large numbers of genes with diverse functions are regulated by stochastic monoallelic expression (Gimelbrant et al., 2007), and random autosomal expression is a common mechanism in rodents and primate orthologs of these genes (Zwemer et al., 2012). Their genes with random monoallelic expression give a possible mechanism for generating cellular diversity. When there is their functional heterozygosity, the random allelic expression apparently leads to cellular functional diversity at the individual cell level. Studies of

random allelic expression patterns should shed light on the roles of cellular individuality. Interestingly, stochastic allelic expression is also observed in immortalized neural clones from human fetal brain, indicating that neurons exhibit clonal diversity in the developing brain (Jeffries et al., 2012).

In the brain, the complexity and diversity of neurons led to the suggestion that a somatic mutational mechanism was responsible for the evolution of diversified neuronal functions (Muotri & Gage, 2006; Rehen et al., 2005), and the endogenous retrotransposition of LINE-1 (L1) elements was proposed as a potential mechanism for generating neuronal genome diversity (Singer et al., 2010). Although it was estimated that in the amplification of the genome of a single neuron in the human brain, < 0.6 unique somatic insertions would occur, most neurons lack detectable somatic insertions, suggesting that L1 is not a major generator of neuronal diversity (Evrony et al., 2012).

Like the immune receptors and odorant receptors, the clustered protocadherins (Pcdhs), which are organized as gene clusters, undergo random allelic expression and elicit specific functional properties in individual neurons (Esumi et al., 2005; Kaneko et al., 2006; Hirano et al., 2012). The clustered Pcdhs are predominantly expressed in vertebrate brains and have functions in mice in neuronal survival, axonal targeting, dendrite arborization, synaptogenesis, neuronal migration, cortical activity, and learning and memory (Yagi, 2008; Zipursky & Sanes, 2010; Yamashita et al., 2012; Hirayama & Yagi, 2013; Weiner & Jontes, 2013; Ledderose et al., 2013). Similarly, clustered *Pcdh* genes are susceptibility genes for autism, bipolar disorder, and musical aptitude (Anitha et al., 2012; Redies et al., 2012; Ukkola-Vuoti et al., 2013). This review describes the molecular characteristics of the clustered Pcdh families and how they generate neuronal individuality, and it discusses how these findings increase our comprehension of neural circuit formation and function.

THE CLUSTERED PROTOCADHERINS

Cadherins are type I transmembrane proteins that were originally characterized as calcium-dependent cell-adhesion molecules (Takeichi, 2007). Cadherin family members are identified by the cadherin sequence repeats in their extracellular (EC) domain. Several subgroups of cadherins are classified by their sequence similarity: the classical (type I) and closely related type II cadherins, desmosomal cadherins, and protocadherins. The sequencing of vertebrate genomes has revealed more than 100 different cadherin-related genes (Morishita & Yagi, 2007). Protocadherins (Pcdhs) were first discovered by their sequence homology with the extracellular repeats of classical cadherins. Of the cadherin superfamily,

which has over 80 members, the *Pcdh*s constitute a major family, and are further divided into clustered *Pcdh*s and nonclustered *Pcdh*s. The clustered *Pcdh*s are divided into three smaller gene clusters: the *Pcdh*-α, -β, and -γ families (Wu & Maniatis, 1999). The *Pcdh*-α family was originally discovered in mouse brain and designated "cadherin-related neuronal receptors" (*CNR*s) (Kohmura et al., 1998). With over 50 members, the clustered Pcdhs are the largest subgroup within the cadherin superfamily. The nonclustered Pcdhs have 13 members: 4 in the *Pcdh-δ1*, 5 in the *Pcdh-δ2*, and 4 in other *Pcdh* subgroups. Interestingly all of the Pcdh-δ2 proteins have a similar structure to the clustered protocadherin members: five EC repeats and a Cys-(X)$_5$-Cys sequence in the EC1 domain (Morishita & Yagi, 2007). The structural characterization of the EC1 domain revealed that the protocadherin-specific disulfide-bounded Cys-(X)$_5$-Cys sequence is exposed to solvent as a loop (Morishita et al., 2006). Similarly, some Pcdh-δ1 proteins have Cys-(X)$_{8 \text{ or } 10}$-Cys sequence in the EC1 domain.

GENOMIC STRUCTURES OF CLUSTERED PROTOCADHERINS

Clustered *Pcdh* genes have been identified in vertebrates. They have an unusual genomic organization, similar to that of the immunoglobulin and T-cell receptor gene clusters. The N-terminal extracellular domains are different, whereas the cytoplasmic tail of each protein is identical among the Pcdh-α or -γ members (Kohmura et al., 1998; Wu & Maniatis, 1999). The Pcdh-α and -γ cytoplasmic tails can bind directly to non-receptor-type tyrosine kinases, FAK and PYK2, which play crucial roles in synaptic and dendrite maturation (Chen et al., 2009; Garret et al., 2012; Suo et al., 2012). Therefore, distinct extracellular interactions can be transferred to a common signaling pathway in the cytoplasm through the cytoplasmic tails of these molecules. The Pcdh-β members do not share a common constant cytoplasmic tail, but this region is highly conserved. In mouse chromosome 18, a total of 58 genes are arranged in the *Pcdh*-α, -β, and -γ clusters, which have 14, 22, and 22 members, respectively. In humans, a total of 53 *Pcdh*-α, -β, and -γ members are located at 5q31. There are many polymorphisms in the *Pcdh* gene cluster, including amino acid exchanges, extensive linkage disequilibrium, and deletions, in the human population (Noonan et al., 2003; Miki et al., 2005).

ROLES OF CLUSTERED PROTOCADHERINS IN NEURAL CIRCUIT FORMATION

The extracellular domains of the Pcdhs contain cadherin motifs. Many of the cadherin superfamily members have important roles in developmental processes, including synapse formation (Yagi & Takeichi, 2000). Mice lacking Pcdh-α are viable but exhibit abnormal learning and memory (Fukuda et al., 2008), and abnormal cortical activity (Yamashita et al., 2012). In addition, the precise targeting of olfactory and serotonergic axonal terminals is impaired in Pcdh-α knockout (KO) mice (Hasegawa et al., 2008, 2012; Katori et al., 2009). The loss of Pcdh-γ leads to neonatal death with neurological impairments, including the cell death of spinal cord interneurons and decreased numbers of synapses (Wang et al., 2002b). Conditional KO experiments for Pcdh-γ revealed that this subfamily has roles in the self-avoidance function in the dendrite formation of retinal starburst amacrine cells and cerebellar Purkinje neurons (Lefebvre et al., 2012). The dendritic phenotype (reduced arborization) was also observed in cortical neurons and granule cells in which Pcdh-γ was conditionally deleted (Garrett et al., 2012; Ledderose et al., 2013). The loss of Pcdh-γ in astrocytes and in neurons in the spinal cord and olfactory granule cells leads to a reduction in synapses (Garrett & Weiner, 2009; Ledderose et al., 2013). These results suggest that the clustered Pcdhs have important function for the proper building of neural networks in the brain. However, it does not have completely revealed their molecular mechanism for building specific connection of neural networks (including synapse formation) at an individual neuron level.

STOCHASTIC AND COMBINATORIAL GENE REGULATION IN INDIVIDUAL NEURONS

One of the most fascinating characteristics of the clustered *Pcdh* families is their differential and combinatorial expression in neurons in the brain. Over 50 clustered *Pcdh*s (i.e., all of the clustered Pcdhs except the 5 "C" members, *Pcdh*-αC1, -αC2, -γC3, -γC4, and -γC5) are expressed in a scattered pattern throughout the brain (Esumi et al., 2005; Kaneko et al., 2006; Noguchi et al., 2009; Yokota et al., 2011; Hirano et al., 2012). In addition, at the single Purkinje cell level, each *Pcdh*-α, -β, or -γ member is expressed monoallelically, although both the paternal and maternal cluster alleles are transcriptionally active in individual neurons (Esumi et al., 2005; Kaneko et al., 2006; Hirano et al., 2012). The promoter of each variable exon is independently chosen from both alleles. The exceptions to this rule are the constitutively expressed *Pcdh*-αC1, -αC2, -γC3, -γC4, and -γC5 variable exons, which are expressed from both alleles in individual neurons (Esumi et al., 2005; Kaneko et al., 2006).

Single-cell reverse transcriptase–polymerase chain reaction (RT-PCR) analysis of cerebellar Purkinje neurons indicated that individual neurons stochastically and mono-allelically express ~2 of the 12 *Pcdh*-α, ~4 of the 22 *Pcdh*-β,
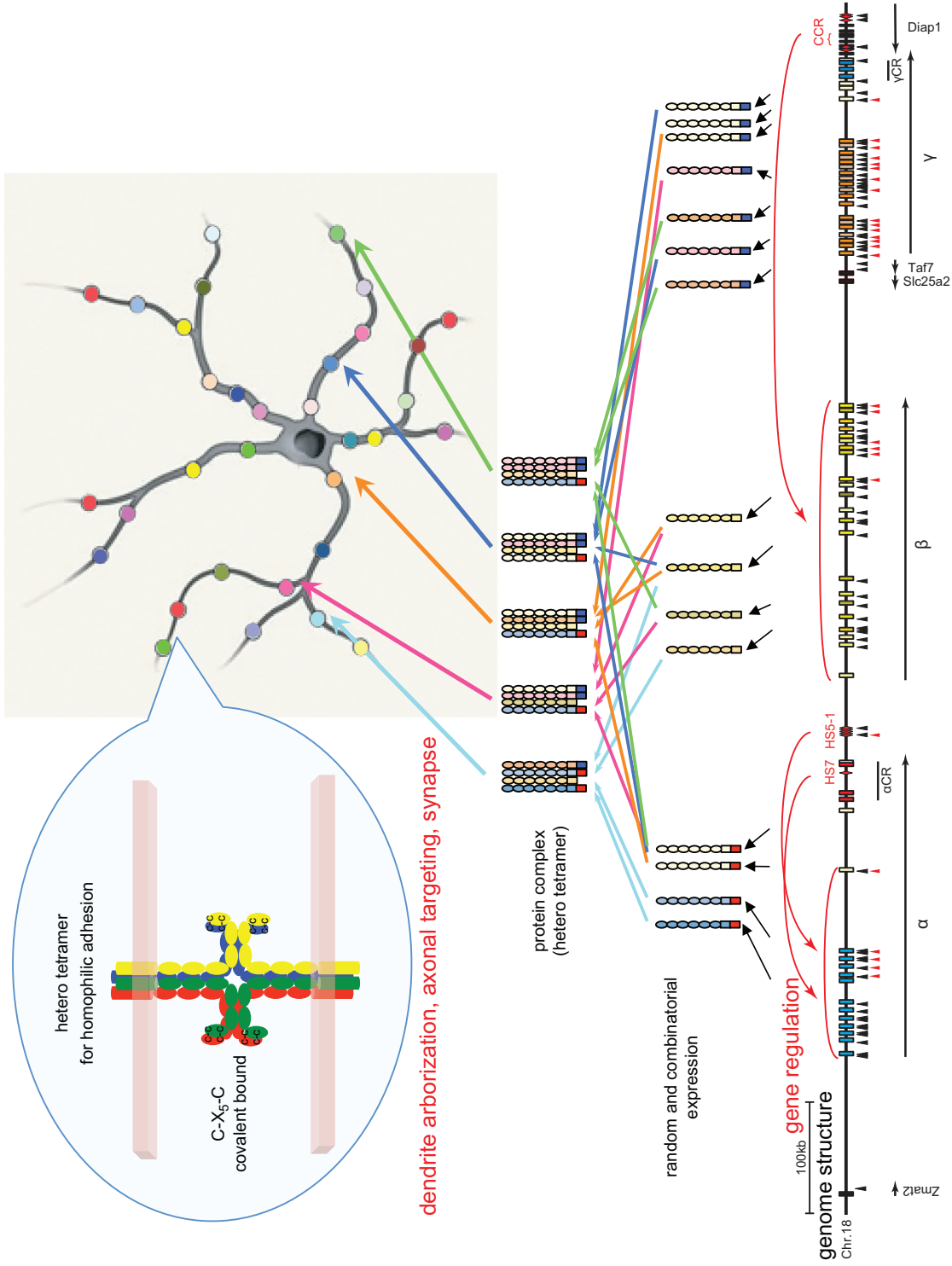
Figure 1. Neuronal individuality determined by the random and combinatorial expression of clustered Pcdh members and the random production of heteromultimeric *cis*-tetramers. From each clustered *Pcdh* allele, about 10 α, β, and γ members are randomly expressed, and 5 "C" members are constitutively expressed. The random expressions are regulated by the *cis*-elements HS5-1 for the *Pcdh-α* cluster and CCR for the *Pcdh-β* cluster; the CCR is located downstream from the *Pcdh-γ* cluster. The random expressions are also regulated by the *trans*-factors CTCF and cohesin-SA. Black and red triangles show CTCF-binding and cohesin-SA-binding sites, respectively. The expressed Pcdh member proteins form random heteromultimeric *cis*-tetramers. The C-X₅-C motif, which is important for forming the *cis*-tetramers, is conserved among all the clustered Pcdh and nonclustered Pcdh-δ2 members. The heteromultimeric *cis*-tetramers can bind homophilically between cells. Genetic analyses in mice suggested that these interactions influence dendritic arborization, axonal targeting, and synaptogenesis. This figure is modified from Yagi (2012).

and ~4 of the 19 *Pcdh-γ* genes (Figure 1). Therefore, 78 possible combinations from the 12 *Pcdh-α*, 26,796 from the 22 *Pcdh-β,* and 14,706 from the 19 *Pcdh-γ* genes are generated, and a total of $78 \times 26{,}796 \times 14{,}706 = 30{,}736{,}834{,}128$ (approximately 3×10^{10}) variations are possible in each individual neuron (Yagi, 2012). Notably, the cerebral cortex of the human brain contains approximately 10^{10} neurons. In addition, the 5 "*C*" members are constitutively and biallelically expressed in neurons, and increase the total number of clustered *Pcdh* members expressed in each neuron, but do not contribute to neuronal variation.

COMBINATORIAL COMPLEXITY OF THE CLUSTERED PROTOCADHERINS AT THE PROTEIN LEVEL

The clustered Pcdh proteins show a punctate localization in neurons in vitro and in vivo (Phillips et al., 2003; Murata et al., 2004; Femandez-Monreal et al., 2009). Pcdh-α, -β, and -γ protein members form a protein complex in the cytoplasmic membrane of cultured cells and in the brain (Murata et al., 2004; Han et al., 2010). These clustered Pcdh protein complexes have cell-adhesion activities, as reported by Schreiner and Weiner (2010), who showed that seven Pcdh-γ members exhibit isoform-specific homophilic binding. These authors suggested that heteromultimeric *cis*-tetramers function as a homophilic binding unit (Schreiner & Weiner, 2010).

In fact, the clustered Pcdhs form *cis*-tetramers before their translocation to the cell membrane. Therefore, the binding properties of the *cis*-tetramers of clustered Pcdhs are very different from those of classical cadherins. If two different clustered Pcdh isoforms are expressed in a cell, six different *cis*-tetramers are expressed on the cell membrane. If one isoform is common between two cells, only one of six tetramers will be common between the two cell surfaces, resulting in weak cell-cell aggregation (Schreiner & Weiner, 2010; Yagi, 2012). Thus, by forming functional units of hetero *cis*-tetramers, the Pcdhs exponential increase the number of variations in the protein complexes at the cell surface.

Individual neurons are estimated to express 15 clustered Pcdh members (i.e., 10 randomly expressed ones: 2 Pcdh-α, 4 Pcdh-β, and 4 Pcdh-γ; and 5 constitutively expressed ones: αC1, αC2, γC3, γC4, and γC5) (Figure 1). Random combinations could produce 12,720 15-member sets of Pcdhs and 15^{4} (50,625) possible tetramers (Yagi, 2012). Notably, the probability of matching tetramers occurring between a pair of neurons decreases exponentially as the number of different isoforms increases. If only 1 of the 15 clustered Pcdh members is different (i.e., 14 of the 15 are the same) between a pair of neurons, only 76% ($14^{4}/15^{4}$) of the tetramers would be matched. Even though the 5 "C" members are always expressed in

common, between a pair of neurons only 1.2% ($5^{4}/15^{4}$) of their tetramers will match. Therefore, the tetramers of the clustered Pcdh members can account for an exponential diversity in Pcdh tetramers between any given pair of neurons (Yagi, 2012). Interestingly, recent physiological studies revealed that the features of the neuronal connections in vivo include a non-Gaussian (long-tail) distribution of synaptic weight (Song et al., 2005; Lefort et al., 2009; Ikegaya et al., 2013).

REGULATION OF CLUSTERED PROTOCADHERIN EXPRESSION

The clustered *Pcdh* genes, except for the 5 "C" members, are randomly expressed from individual alleles in individual neurons. The sequence upstream of each variable exon contains its own promoter, which includes a conserved sequence element (CSE) (Wu & Maniatis, 1999). The CSE is required for the expression of the variable exons (Tasic et al., 2002; Wang et al., 2002a). The *cis*-regulatory elements are identified by their sequence conservation and hypersensitivity to DNase I (DNase I hypersensitive site; HS). Interestingly, the *cis*-regulatory elements independently control each gene cluster (Ribich et al., 2006; Yokota et al., 2011). Deletion of the HS5-1 site in mice leads to significantly reduced expressions of *Pcdh-α*6 to -α12 and -αC1 (Kehayova et al., 2011; Yokota et al., 2011). Deletion of the HS16–20 site (called the cluster control region; CCR) leads to a nearly complete loss of random allelic expression across the *Pcdh-β* cluster. Notably, deletion of the CCR has a much smaller effect on the closer *Pcdh-γ* cluster, and no effect on the *Pcdh-α* cluster in mice (Yokota et al., 2011), demonstrating that the CCR is an important *cis*-element for long-range independent regulation of the *Pcdh-β* cluster (Figure 1).

The random expressions of the *Pcdh-α* isoforms depend on the number of variable exons in the *Pcdh-α* cluster. In cells bearing an allele in which variable exons α2 to α11 (from the total α1 to α12) are deleted, the expressional frequencies of the remaining variable exons α1 and α12 are increased, compared with those of the wild-type allele. That is, each neuron expresses α*1*, α12, or both from the deletion allele, whereas α*1* and α12 are expressed only rarely (i.e., stochastically) from the wild-type allele. Interestingly, the total *Pcdh-α* expression level is maintained even in the deletion allele (Noguchi et al., 2009). Thus, the expressions of the variable exons are stochastic, like the results of throwing dice.

Results from several mouse lines with deletions or duplications within the *Pcdh-α* variable exon cluster support the hypothesis that the variable isoforms are stochastically expressed, that the frequency of their expression depends on the number of variable exons in the cluster, and that the total expression level is maintained

by long-range regulation by the outer *cis*-regulatory elements. Interestingly, in these deletion constructs, the 3'-most variable exon in each construct assumed a ubiquitous and biallelic expression pattern similar to that of αC2 in the *Pcdh-α* cluster in the wild-type allele.

The CCCTC-binding factor (CTCF), a zinc finger transcription factor, completely regulates the random allelic expression of the clustered *Pcdh*s at the individual neuron level, but does not down-regulate the constitutively expressed αC2, γC3, γC4, and γC5 (Hirayama et al., 2012). CTCF-binding sites are found in the randomly expressed *Pcdh* promoters and in HS5-1 and HS16–20 (Figure 1) (Kehayova et al., 2011; Golan-Mashiach et al., 2012; Hirayama et al., 2012). CTCF is known to mediate enhancer and promoter interactions by DNA looping (Merkenschlager & Odom, 2013). In long-range regulation, the DNA looping between a promoter and enhancer mediated by CTCF account for the "random choice" of clustered *Pcdh* genes in individual neurons. Notably, neurons lacking CTCF exhibit abnormal dendritic arborization and decreased spine density (and therefore decreased numbers of synapses), but show no effects on synapse maturation or the amplitude of excitatory synaptic currents. In addition, mice with a conditional CTCF KO in differentiated excitatory neurons exhibit a somatosensory map defect (a loss of barrel structure) in the brain. Thus, CTCF affects the final refinement of functional neural circuits during postnatal brain development (Hirayama et al., 2012).

The *Pcdh* expression in individual neurons is epigenetically controlled, independently and monoallelically (Tasic et al., 2002; Kawaguchi et al., 2008). Using two mouse cell lines that express different *Pcdh-α* isoform combinations, the DNA methylation of each *Pcdh-α* promoter was found to be negatively correlated with its expression level (Kawaguchi et al., 2008). In addition, the induction of DNA demethylation with 5-azacytidin increases the *Pcdh-α* transcription. Consistent with their constitutive expression in neurons, the promoters of *Pcdh*-αC1 and -αC2 are hypomethylated in the brain (Kawaguchi et al., 2008).

The hypermethylation of multiple clustered *Pcdh* genes is observed in mouse offspring receiving a low frequency of maternal care (licking and grooming behaviors), whereas many of the clustered *Pcdh* genes show significantly higher expression in offspring receiving a high frequency of maternal care (McGowan et al., 2011). Similar results were found in humans; hippocampal samples from suicide victims with a history of severe child abuse exhibit hypermethylation across the clustered *Pcdh* genes (Suderman et al., 2012). It is possible that the clustered *Pcdh* genes play an important role in the response to environmental stress. Interestingly, Tet1 protein, a DNA hydroxylase that converts 5-methyl cytosine to 5-hydroxymethylcytosine in DNA demethylation, binds

across the clustered *Pcdh* genes (Xu et al., 2011). These epigenetic studies and the molecular functions of the clustered Pcdhs in axonal targeting and dendritic arborization suggest that they are important mediators of neural circuit formation that respond to environmental stress during brain development. Furthermore, since epigenetic modifications in individual cells are maintained during cell proliferation, the neuronal individuality based on the random and combinatorial expression of the clustered *Pcdh* genes could be responsible for the clonal diversity of differentiated neurons related to cell lineage.

DETERMINANTS OF NEURONAL INDIVIDUALITY

Combinatorial explosion can generate a nearly limitless variety of information through the genetic code simply through variations in ATGC sequence combinations. In the immune system, combinations of random DNA rearrangements of variable regions, somatic mutations, and receptor heteromultimers generate the predetermined immune response to an unpredictable antigen. In the brain, neurons connect to form specific neural networks. Donald Hebb hypothesized the "cell assembly" hypothesis for the encoding of neural information by functional neuronal groups (Hebb, 1949). In Hebb's cell assembly hypothesis, a nearly limitless number of combinatorial neuronal groups is produced from a limited number of neurons by combinatorial explosion. Recent physiological studies involving large-scale recordings from individual neurons have experimentally defined putative cell assemblies (reviewed by Buzsaki, 2010). Neuronal individuality is likely to be generated by both a neuron's Pcdh expression pattern and its participation in circuits.

CONCLUSION

Genetic studies of the clustered Pcdhs have gradually unveiled their functions in axonal targeting, dendritic arborization, self-avoidance of dendrites, and synapse formation. In addition, the clustered Pcdhs contribute to neuronal individuality (approximate 10^{10} neuronal diversity) by random allelic expression. This individuality is generated by epigenetic regulation in the genome during development. Furthermore, heterotetramers of the clustered Pcdhs provide specific cell-adhesion activities and exponential neuronal diversity at the cell surface.

Recent physiological studies have revealed that local networks form complex topological neuronal ensembles in which there are differences at the individual neuron level (Markram et al., 1997; Kalisman et al., 2005; Song et al., 2005; Yoshimura et al., 2005; Koulakov et al., 2009; Yassin et al., 2010; Perin et al., 2011). In addition, the synaptic

weight of the local connectivity has a non-Gaussian (long-tail) distribution (Song et al., 2005; Lefort et al., 2009; Ikegaya et al., 2013). Specific local connectivities develop preferentially among clonally related excitatory cortical neurons (Yu et al., 2009, 2012; Li et al., 2012; Ohtsuki et al., 2012). Theoretical analyses of neural networks have suggested that complex topological networks exist in the brain (Sporns, 2011). Interestingly, Watts and Strogatz showed that "small-world" networks with high clustering coefficients and short characteristic path lengths, but without strong habbs, emerge between random and regular networks (Watts & Strogatz, 1998). It is estimated that the neural wiring in the brain is a Watts and Strogatz–type "small-world" network. The molecular features of the clustered Pcdhs, which permit random and selective connectivity, provide a fascinating mechanism underlying the building of complex networks in the brain.

ACKNOWLEGMENTS

REFERENCES

Anitha, A., Thanseem, I., Nakamura, K., Yamada, K., Iwayama, Y., Toyota, T., Iwata, Y., Suzuki, K., Sugiyama, T., Tsujii, M., Yoshikawa T., & Mori N. (2012). Protocadherin alpha (PCDHA) as a novel susceptibility gene for autism. *J Psychiatry Neurosci, 37*, 12.

Buck, L., & Axel, R. (1991). A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell, 65*, 175–187.

Buzsaki, G. (2010). Neural syntax: Cell assemblies, synapsembles, and readers. *Neuron, 68*, 362–385.

Chen, J., Lu, Y., Meng, S., Han, M. H., Lin, C., & Wang, X. (2009). alpha- and gamma-Protocadherins negatively regulate PYK2. *J Biol Chem, 284*, 2880–2890.

Esumi, S., Kakazu, N., Taguchi, Y., Hirayama, T., Sasaki, A., Hirabayashi, T., Koide, T., Kitsukawa, T., Hamada, S., & Yagi, T. (2005). Monoallelic yet combinatorial expression of variable exons of the protocadherin-alpha gene cluster in single neurons. *Nat Genet, 37*, 171–176.

Evrony, G. D., Cai, X., Lee, E., Hills, L. B., Elhosary, P. C., Lehmann, H. S., Parker, J. J., Atabay, K. D., Gilmore, E. C.,

Poduri, A., Park, P. J., & Walsh, C. A. (2012). Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell, 151*, 483–496.

Fernandez-Monreal, M., Kang, S., & Phillips, G. R. (2009). Gamma-protocadherin homophilic interaction and intracellular trafficking is controlled by the cytoplasmic domain in neurons. *Mol Cell Neurosci, 40*, 344–353.

Fukuda, E., Hamada, S., Hasegawa, S., Katori, S., Sanbo, M., Miyakawa, T., Yamamoto, T., Yamamoto, H., Hirabayashi, T., & Yagi, T. (2008). Down-regulation of protocadherin-alpha A isoforms in mice changes contextual fear conditioning and spatial working memory. *Eur J Neurosci, 28*, 1362–1376.

Garrett, A. M., Schreiner, D., Lobas, M. A., & Weiner, J.A. (2012). gamma-protocadherins control cortical dendrite arborization by regulating the activity of a FAK/PKC/MARCKS signaling pathway. *Neuron, 74*, 269–276.

Garrett, A. M., & Weiner, J. A. (2009). Control of CNS synapse development by {gamma}-protocadherin-mediated astrocyte-neuron contact. *J Neurosci, 29*, 11723–11731.

Gimelbrant, A., Hutchinson, J. N., Thompson, B. R., & Chess, A. (2007). Widespread monoallelic expression on human autosomes. *Science, 318*, 1136–1140.

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*, 3378–3391.

Han, M. H., Lin, C., Meng, S., & Wang, X. (2010). Proteomics analysis reveals overlapping functions of clustered protocadherins. *Mol Cell Proteomics, 9*, 71–83.

Hasegawa, S., Hamada, S., Kumode, Y., Esumi, S., Katori, S., Fukuda, E., Uchiyama, Y., Hirabayashi, T., Mombaerts, P., & Yagi, T. (2008). The protocadherin-alpha 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., Hirabayashi, T., Kondo, T., Inoue, K., Esumi, S., Okayama, A., Hamada, S., & Yagi, T. (2012). Constitutively expressed Protocadherin-alpha regulates the coalescence and elimination of homotypic olfactory axons through its cytoplasmic region. *Front Mol Neurosci, 5*, 97.

Hebb, D.O. (1949). *The organization of behavior.* New York: John Wiley & Sons.

Hirano, K., Kaneko, R., Izawa, T., Kawaguchi, M., Kitsukawa, T., & Yagi, T. (2012). Single-neuron diversity generated by Protocadherin-beta cluster in mouse central and peripheral nervous systems. *Front Mol Neurosci, 5*, 90.

Hiratani, N., Teramae, J. N., & Fukai, T. (2013). Associative memory model with long-tail-distributed Hebbian synaptic connections. *Front Comput Neurosci, 6*, 102.

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*, 345–357.

Hirayama, T., & Yagi, T. (2013). Clustered protocadherins and neuronal diversity. *Prog Mol Biol Transl Sci, 116*, 145–167.

Ikegaya, Y., Sasaki, T., Ishikawa, D., Honma, N., Tao, K., Takahashi, N., Minamisawa, G., Ujita, S., & Matsuki, N. (2013). Interpyramid Spike transmission stabilizes the sparseness of recurrent network activity. *Cereb Cortex, 23*, 293–304.

Jeffries, A. R., Perfect, L. W., Ledderose, J., Schalkwyk, L. C., Bray, N. J., Mill, J., & Price, J. (2012). Stochastic choice of allelic expression in human neural stem cells. *Stem Cells, 30*, 1938–1947.

Kalisman, N., Silberberg, G., & Markram, H. (2005). The neocortical microcircuit as a tabula rasa. *Proc Natl Acad Sci U S A, 102*, 880–885.

Kaneko, R., Kato, H., Kawamura, Y., Esumi, S., Hirayama, T., Hirabayashi, T., & Yagi, T. (2006). Allelic gene regulation of Pcdh-alpha and Pcdh-gamma clusters involving both monoallelic and biallelic expression in single Purkinje cells. *J Biol Chem, 281*, 30551–30560.

Katori, S., Hamada, S., Noguchi, Y., Fukuda, E., Yamamoto, T., Yamamoto, H., Hasegawa, S., & Yagi, T. (2009). Protocadherin-alpha family is required for serotonergic projections to appropriately innervate target brain areas. *J Neurosci, 29*, 9137–9147.

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-alpha gene cluster. *J Biol Chem, 283*, 12064–12075.

Kehayova, P., Monahan, K., Chen, W., & Maniatis, T. (2011). Regulatory elements required for the activation and repression of the protocadherin-alpha gene cluster. *Proc Natl Acad Sci U S A, 108*, 17195–17200.

Kohmura, N., Senzaki, K., Hamada, S., Kai, N., Yasuda, R., Watanabe, M., Ishii, H., Yasuda, M., Mishina, M., & Yagi, T. (1998). Diversity revealed by a novel family of cadherins expressed in neurons at a synaptic complex. *Neuron, 20*, 1137–1151.

Koulakov, A. A., Hromadka, T., & Zador, A. M. (2009). Correlated connectivity and the distribution of firing rates in the neocortex. *J Neurosci, 29*, 3685–3694.

Ledderose, J., Dieter, S., & Schwarz, M. K. (2013). Maturation of postnatally generated olfactory bulb granule cells depends on functional gamma-protocadherin expression. *Sci Rep, 3*, 1514.

Lefebvre, J. L., Kostadinov, D., Chen, W. V., Maniatis, T., & Sanes, J. R. (2012). Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature, 488*, 517–521.

Lefort, S., Tomm, C., Floyd Sarria, J. C., & Petersen, C. C. (2009). The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. *Neuron, 61*, 301–316.

Li, Y., Lu, H., Cheng, P.L., Ge, S., Xu, H., Shi, S.H., & Dan, Y. (2012). Clonally related visual cortical neurons show similar stimulus feature selectivity. *Nature, 486*, 118–121.

Lieber, M. R. (1992). The mechanism of V(D)J recombination: A balance of diversity, specificity, and stability. *Cell, 70*, 873–876.

Markram, H. (1997). A network of tufted layer 5 pyramidal neurons. *Cereb Cortex, 7*, 523–533.

McGowan, P. O., Suderman, M., Sasaki, A., Huang, T. C., Hallett, M., Meaney, M. J., & Szyf, M. (2011). Broad epigenetic signature of maternal care in the brain of adult rats. *PLoS ONE, 6*, e14739.

Merkenschlager, M., & Odom, D. T. (2013). CTCF and cohesin: Linking gene regulatory elements with their targets. *Cell, 152*, 1285–1297.

Miki, R., Hattori, K., Taguchi, Y., Tada, M. N., Isosaka, T., Hidaka, Y., Hirabayashi, T., Hashimoto, R., Fukuzako, H., & Yagi, T. (2005). Identification and characterization of coding single-nucleotide polymorphisms within human protocadherin-alpha and -beta gene clusters. *Gene, 349*, 1–14.

Morishita, H., Umitsu, M., Murata, Y., Shibata, N., Udaka, K., Higuchi, Y., Akutsu, H., Yamaguchi, T., Yagi, T., & Ikegami, T. (2006). Structure of the cadherin-related neuronal receptor/protocadherin-alpha first extracellular cadherin domain reveals diversity across cadherin families. *J Biol Chem, 281*, 33650–33663.

Morishita, H., & Yagi, T. (2007). Protocadherin family: Diversity, structure, and function. *Curr Opin Cell Biol, 19*, 584–592.

Muotri, A. R., & Gage, F. H. (2006). Generation of neuronal variability and complexity. *Nature, 441*, 1087–1093.

Murata, Y., Hamada, S., Morishita, H., Mutoh, T., & Yagi, T. (2004). Interaction with protocadherin-gamma regulates the cell surface expression of protocadherin-alpha. *J Biol Chem, 279*, 49508–49516.

Noguchi, Y., Hirabayashi, T., Katori, S., Kawamura, Y., Sanbo, M., Hirabayashi, M., Kiyonari, H., Nakao, K., Uchimura, A., & Yagi, T. (2009). Total expression and dual gene-regulatory mechanisms maintained in deletions and duplications of the Pcdha cluster. *J Biol Chem, 284*, 32002–32014.

Noonan, J. P., Li, J., Nguyen, L., Caoile, C., Dickson, M., Grimwood, J., Schmutz, J., Feldman, M. W., & Myers, R. M. (2003). Extensive linkage disequilibrium, a common 16.7-kilobase deletion, and evidence of balancing selection in the human protocadherin alpha cluster. *Am J Hum Genet, 72*, 621–635.

Ohtsuki, G., Nishiyama, M., Yoshida, T., Murakami, T., Histed, M., Lois, C., & Ohki, K. (2012). Similarity of visual selectivity among clonally related neurons in visual cortex. *Neuron, 75*, 65–72.

Perin, R., Berger, T. K., & Markram, H. (2011). A synaptic organizing principle for cortical neuronal groups. *Proc Natl Acad Sci U S A, 108*, 5419–5424.

Phillips, G. R., Tanaka, H., Frank, M., Elste, A., Fidler, L., Benson, D. L., & Colman, D. R. (2003). Gamma-protocadherins are targeted to subsets of synapses and intracellular organelles in neurons. *J Neurosci, 23*, 5096–5104.

Redies, C., Hertel, N., & Hubner, C. A. (2012). Cadherins and neuropsychiatric disorders. *Brain Res, 1470*, 130–144.

Rehen, S. K., Yung, Y. C., McCreight, M. P., Kaushal, D., Yang, A. H., Almeida, B. S., Kingsbury, M. A., Cabral, K. M., McConnell, M.J., Anliker, B., et al. (2005). Constitutional aneuploidy in the normal human brain. *J Neurosci, 25*, 2176–2180.

Ribich, S., Tasic, B., & Maniatis, T. (2006). Identification of long-range regulatory elements in the protocadherin-alpha gene cluster. *Proc Natl Acad Sci U S A, 103*, 19719–19724.

Sakano, H. (2010). Neural map formation in the mouse olfactory system. *Neuron, 67*, 530–542.

Schreiner, D., & Weiner, J. A. (2010). Combinatorial homophilic interaction between gamma-protocadherin multimers greatly expands the molecular diversity of cell adhesion. *Proc Natl Acad Sci U S A, 107*, 14893–14898.

Singer, T., McConnell, M. J., Marchetto, M. C., Coufal, N. G., & Gage, F.H. (2010). LINE-1 retrotransposons: Mediators of somatic variation in neuronal genomes? *Trends Neurosci, 33*, 345–354.

Song, S., Sjostrom, P. J., Reigl, M., Nelson, S., & Chklovskii, D. B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol, 3*, e68.

Sporns, O. (2011). *Networks of the brain.* Cambridge: The MIT Press.

Suderman, M., McGowan, P. O., Sasaki, A., Huang, T. C., Hallett, M. T., Meaney, M. J., Turecki, G., & Szyf, M. (2012). Conserved epigenetic sensitivity to early life experience in the rat and human hippocampus. *Proc Natl Acad Sci U S A, 109* (Suppl 2), 17266–17272.

Suo, L., Lu, H., Ying, G., Capecchi, M. R., & Wu, Q. (2012). Protocadherin clusters and cell adhesion kinase regulate dendrite complexity through Rho GTPase. *J Mol Cell Biol, 4*, 362–376.

Takeichi, M. (2007). The cadherin superfamily in neuronal connections and interactions. *Nat Rev Neurosci, 8*, 11–20.

Tasic, B., Nabholz, C. E., Baldwin, K. K., Kim, Y., Rueckert, E. H., Ribich, S. A., Cramer, P., Wu, Q., Axel, R., & Maniatis, T. (2002). Promoter choice determines splice site selection in protocadherin alpha and gamma pre-mRNA splicing. *Mol Cell, 10*, 21–33.

Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature, 302*, 575–581.

Ukkola-Vuoti, L., Kanduri, C., Oikkonen, J., Buck, G., Blancher, C., Raijas, P., Karma, K., Lahdesmaki, H., & Jarvela, I. (2013). Genome-wide copy number variation analysis in extended families and unrelated individuals characterized for musical aptitude and creativity in music. *PLoS One, 8*, e56356.

Wang, X., Su, H., & Bradley, A. (2002a). Molecular mechanisms governing Pcdh-gamma gene expression: Evidence for a multiple promoter and *cis*-alternative splicing model. *Genes Dev, 16*, 1890–1905.

Wang, X., Weiner, J. A., Levi, S., Craig, A. M., Bradley, A., & Sanes, J. R. (2002b). Gamma protocadherins are required for survival of spinal interneurons. *Neuron, 36*, 843–854.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*, 440–442.

Weiner, J. A., & Jontes, J. D. (2013). Protocadherins, not prototypical: A complex tale of their interactions, expression, and functions. *Front Mol Neurosci, 6*, 4.

Wu, Q., & Maniatis, T. (1999). A striking organization of a large family of human neural cadherin-like cell adhesion genes. *Cell, 97*, 779–790.

Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A. J., Zheng, L., Zhang, H., Huang, S., et al. (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell, 42*, 451–464.

Yagi, T. (2008). Clustered protocadherin family. *Dev Growth Differ, 50*, S131–S140.

Yagi, T. (2012). Molecular codes for neuronal individuality and cell assembly in the brain. *Front Mol Neurosci, 5*, 45.

Yagi, T., & Takeichi, M. (2000). Cadherin superfamily genes: Functions, genomic organization, and neurologic diversity. *Genes Dev, 14*, 1169–1180.

Yamashita, H., Chen, S., Komagata, S., Hishida, R., Iwasato, T., Itohara, S., Yagi, T., Endo, N., Shibata, M., & Shibuki, K. (2012). Restoration of contralateral representation in the mouse somatosensory cortex after crossing nerve transfer. *PLoS ONE, 7*, e35676.

Yassin, L., Benedetti, B. L., Jouhanneau, J. S., Wen, J. A., Poulet, J. F., & Barth, A. L. (2010). An embedded subnetwork of highly active neurons in the neocortex. *Neuron, 68*, 1043–1050.

Yokota, S., Hirayama, T., Hirano, K., Kaneko, R., Toyoda, S., Kawamura, Y., Hirabayashi, M., Hirabayashi, T., & Yagi, T. (2011). Identification of the cluster control region for the protocadherin-beta genes located beyond the protocadherin-gamma cluster. *J Biol Chem, 286*, 31885–31895.

Yoshimura, Y., & Callaway, E. M. (2005). Fine-scale specificity of cortical networks depends on inhibitory cell type and connectivity. *Nat Neurosci, 8*, 1552–1559.

Yoshimura, Y., Dantzker, J. L., & Callaway, E. M. (2005). Excitatory cortical neurons form fine-scale functional networks. *Nature, 433*, 868–873.

Yu, Y. C., Bultje, R. S., Wang, X., & Shi, S. H. (2009). Specific synapses develop preferentially among sister excitatory neurons in the neocortex. *Nature, 458*, 501–504.

Yu, Y. C., He, S., Chen, S., Fu, Y., Brown, K. N., Yao, X. H., Ma, J., Gao, K. P., Sosinsky, G. E., Huang, K., & Shi, S. H. (2012). Preferential electrical coupling regulates neocortical lineage-dependent microcircuit assembly. *Nature, 486*, 113–117.

Zipursky, S. L., & Sanes, J. R. (2010). Chemoaffinity revisited: Dscams, protocadherins, and neural circuit assembly. *Cell, 143*, 343–353.

Zwemer, L. M., Zak, A., Thompson, B. R., Kirby, A., Daly, M. J., Chess, A., & Gimelbrant, A. A. (2012). Autosomal monoallelic expression in the mouse. *Genome Biol, 13*, R10.