Monoallelic expression of the human *FOXP2* speech gene

Abidemi A. Adegbola[a,b], Gerald F. Cox[c,d], Elizabeth M. Bradshaw[e], David A. Hafler[f], Alexander Gimelbrant[g,h], and Andrew Chess[a,i,1]

[a]Department of Developmental and Regenerative Biology, Mount Sinai School of Medicine, New York, NY 10029; [b]Center for Autism, Philadelphia, PA 19131; [c]Division of Genetics and Department of Pediatrics, Boston Children's Hospital, Boston, MA 02115; [d]Department of Clinical Development, Genzyme, a Sanofi Company, Cambridge, MA 02142; [e]Center for Neurologic Diseases, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115; [f]Departments of Neurology and Immunobiology, Yale School of Medicine, New Haven, CT 06520; [g]Department of Cancer Biology, Dana–Farber Cancer Institute, Boston, MA 02215; [h]Department of Genetics, Harvard Medical School, Boston, MA 02115; and [i]Fishberg Department of Neuroscience, Department of Genetics and Genomic Sciences, Friedman Brain Institute, Mount Sinai School of Medicine, New York, NY 10029

Edited by Eric B. Keverne, University of Cambridge, Cambridge, United Kingdom, and accepted by the Editorial Board October 8, 2014 (received for review July 16, 2014)

The recent descriptions of widespread random monoallelic expression (RMAE) of genes distributed throughout the autosomal genome indicate that there are more genes subject to RMAE on autosomes than the number of genes on the X chromosome where X-inactivation dictates RMAE of X-linked genes. Several of the autosomal genes that undergo RMAE have independently been implicated in human Mendelian disorders. Thus, parsing the relationship between allele-specific expression of these genes and disease is of interest. Mutations in the human forkhead box P2 gene, *FOXP2*, cause developmental verbal dyspraxia with profound speech and language deficits. Here, we show that the human *FOXP2* gene undergoes RMAE. Studying an individual with developmental verbal dyspraxia, we identify a deletion 3 Mb away from the *FOXP2* gene, which impacts *FOXP2* gene expression in cis. Together these data suggest the intriguing possibility that RMAE impacts the haploinsufficiency phenotypes observed for *FOXP2* mutations.

random monoallelic expression | FOXP2 | speech | language | developmental verbal dyspraxia

Various mechanisms of monoallelic gene expression have been characterized in humans, each with impact on neurodevelopment and disease. Potential mechanisms yielding monoallelic expression include those with a genetic basis, such as aneuploidies, copy number variants (CNVs), and nonsense gene mutations. Genetic causes of monoallelic expression can also be due to perturbation of cis-regulatory DNA sequences that control the level, timing, and location of gene expression. Such cis-regulatory sequences can be found within and near the promoter, as well as at large distances from the start site of transcription. Polymorphisms in cis-regulatory elements (1, 2) can lead to differences in levels of expression between the two alleles that can be extreme (greater than 10-fold difference) or less pronounced.

In the last few decades, a number of epigenetic mechanisms that can cause monoallelic expression have been discovered. For example, imprinting is a parent-of-origin–dependent monoallelic expression whereby expression of a locus differs between the maternally and paternally inherited alleles, and generally manifests as transcriptional silencing of one of the alleles. Moreover, some genes are imprinted in a tissue- or isoform-specific manner (3). A variety of human developmental disorders are directly traceable to dysregulation of imprinting, secondary to genetic mutations impacting imprinting. For example, Prader–Willi and Angelman syndromes are severe neurobehavioral disorders whose molecular bases are related to defects, either by deletion, uniparental disomy, or somatic mutation, of the Prader–Willi/Angelman-imprinted locus on 15q11-13. Various deletions lead to distinct phenotypic presentations, depending both on the number of genes impacted and which parental allele is mutated (4, 5).

For many genes at the Prader–Willi/Angelman locus, one allele (maternal or paternal) is transcribed. For a given imprinted gene, because there is monoallelic expression in the WT state, a deletion of one allele leads to complete loss of expression if the normally expressed allele is the deleted allele. Angelman syndrome always involves the imprinted gene ubiquitin-protein ligase E3A (*UBE3A*), which is a maternally expressed gene. Mechanisms leading to Angelman syndrome include a range of maternally inherited *UBE3A* mutations (ranging from point mutations to full deletion of the gene), as well as uniparental disomy (UPD) of the paternal chromosome. Another interesting mechanism causing Angelman syndrome is an imprinting defect. Some of these imprinting defects (10–20%) are due to microdeletions (6–200 kb) that include the AS imprinting center, a genomic region necessary for appropriate imprinting. The other 80–90% of imprinting defects are thought to be due to epimutations (i.e., alterations of the epigenetic status, not the DNA sequence) that occur during oogenesis (in the mother) or occur during early embryogenesis in the affected individual (6). All of the above mechanisms cause Angelman syndrome by perturbing the expression or function of the maternal allele of the *UBE3A* gene.

A similar range of molecular mechanisms (deletion, UPD, and imprinting defect) exists with respect to causation of Prader–Willi syndrome (PWS), except that, in PWS, it is the paternal allele that is perturbed. Although in PWS a number of genes are directly impacted, in 99% of cases, a molecular signature of PWS can be recognized by analysis of promoter, exon 1, and intron 1 regions of the small nuclear ribonucleoprotein polypeptide N gene, *SNRPN* (7, 8). Under normal circumstances, these regions of the *SNRPN* gene are unmethylated on the paternal, expressed allele and are methylated on the maternal, repressed allele. Most PWS imprinting defects are epimutations leading to loss of expression from active alleles of a number of genes on the paternal chromosome. As with AS, in PWS, for each impacted gene, the loss of one allele is not compensated for by expression of the other allele. We point the reader to various excellent reviews of imprinting in refs. 3, 9, and 10.

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "Epigenetic Changes in the Developing Brain: Effects on Behavior," held March 28–29, 2014, at the National Academy of Sciences in Washington, DC. The complete program and video recordings of most presentations are available on the NAS website at www.nasonline.org/Epigenetic_changes.

Author contributions: A.A.A., G.F.C., and A.C. designed research; A.A.A. and A.G. performed research; E.M.B., D.A.H., and A.C. contributed new reagents/analytic tools; A.A.A., A.G., and A.C. analyzed data; and A.A.A., G.F.C., and A.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. E.B.K. is a guest editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. Email: andrew.chess@mssm.edu.

www.pnas.org/cgi/doi/10.1073/pnas.1411270111

Downloaded at Palestinian Territory, occupied on November 27, 2021

www.manaraa.com

X-linked random monoallelic expression is another epigenetic mechanism of monoallelic expression. The initial random chromosome-wide choice between the two X chromosomes is followed by a stable mitotic transmission of monoallelic expression. There can be primary skewing of X-inactivation, which may be due to sequence polymorphism or stochastic expression of a few random choices. Another mechanism, secondary skewing, has an initial random equal choice followed by selection for cells expressing one or the other X chromosome due to differential growth/survival of cells (11). The extent to which skewing of X-inactivation plays a role in brain phenotypes is difficult to assess, but undoubtedly the potential for such a role is there (12). Mitigation of X-linked dominant neurodevelopmental phenotypes in females has been described in association with skewing of X-inactivation toward increased fraction of cells expressing the nonmutant allele. Skewing of X-inactivation can, of course, also lead to more severe phenotypes in cases where skewing leads to a majority of the cells expressing the deleterious allele (12–14).

A third epigenetic mechanism is autosomal random monoallelic expression (RMAE), which has been recently described on a genome-scale level (15–18) and has not yet been as well characterized as the others. A defining feature of autosomal genes that share similarities with the genes subject to X-chromosome inactivation is that they are monoallelically expressed in a random manner. For some of these genes, half the cells express the maternal allele, and half the cells express the paternal allele. Other genes also falling into the randomly monoallelically expressed class have some cells with biallelic expression in addition to the cells with monoallelic expression. RMAE as defined here indicates that, in cells monoallelic for one allele, the other allele is basically undetectable (as has been observed in X-inactivation) and practically (because of limits of detection) can be said to be at least 20- to 50-fold lower than the expressed allele. Autosomal random monoallelic expression thus leads to three distinct expression states for each gene: expression of both alleles in some cells and expression of either the maternal allele or the paternal allele in other cells. Autosomal RMAE can impact biological function by affording cells unique specificity when the products of heterozygous loci might otherwise compete, and it also enhances the phenotypic heterogeneity that is possible in a population of cells (19).

Genes encoding immunoglobulins, T-cell receptors, and odorant receptors were the earliest described autosomal examples of random monoallelic expression. Perhaps the clearest example of the functional importance of this mechanism is the RAG-mediated DNA rearrangement mechanism of generating antigen-receptor diversity. The proper functioning of the immune system relies on a single antigen receptor being expressed in each T cell and each B cell; RMAE is part of the V(D)J recombination process that generates the extraordinarily large repertoire of receptors and ensures that individual lymphocytes do not have dual specificities (20–23). Another well-characterized example of RMAE is in the olfactory system (24). The expression of a single odorant receptor (25) dictates both the olfactory sensitivity of the neuron choosing it and also is involved in axon guidance (26, 27). If an individual neuron were to express both alleles of a functionally heterozygous odorant-receptor gene, confusion in neuronal wiring could result. In addition, the expression of two receptors would make it difficult for the organism to detect differences in the molecules sensed by the two alleles of a heterozygous odorant-receptor gene.

The discrete repertoire of autosomal random monoallelically expressed genes was dramatically expanded after a genome-wide survey by Gimelbrant et al. (15) (using SNP arrays to differentially detect transcripts from each allele of heterozygous loci in human cell lines) that found that a surprisingly large subset of autosomal human genes (~5–10%) undergo random monoallelic expression. Most of the experiments examined clonal lymphoblastoid cell lines, where, within a given clone, there was stability in the allele-specific gene expression. Because in vivo there is also clonal expansion, random monoallelic expression can lead to growth of macroscopic patches of tissue with subtly distinct properties. Indeed, such patches were observed in normal placenta, and other tissues remain to be analyzed.

A property that autosomal RMAE does not share with X-inactivation is that autosomal RMAE can lead to some cells expressing two alleles and others expressing one. Thus, even in the absence of heterozygosity, the ability of cells to express either one or two alleles can lead to differences in levels of expression that can also contribute to cellular diversity (15–18). When there is functional heterozygosity, the chance to generate diversity by having independent expression of the two alleles is readily apparent.

The disease potential of RMAE genes with respect to brain development is suggested by the fact that some genes that undergo RMAE are known to cause disease phenotypes in contexts defined by genetic gain of function. Dominant mutations and duplication of *APP* and *SNCA* are associated with early-onset Alzheimer disease and Parkinson disease, respectively (28–34). Genes that are subject to an autosomal-dominant phenotype in the context of haploinsufficiency would be additional strong candidates to be affected or modified in their phenotypic presentation by RMAE. A nonsense mutation or deletion that obliterates transcription from one allele would, in this context, lead to a complete loss of gene expression in the population of cells that would normally express that allele exclusively. Cells that express only the WT allele would be unaffected. Biallelically expressing cells would express one intact functional allele. Similar to secondary skewing in X-inactivation, the potential exists for RMAE disease phenotypes to be modified by differential growth or survival of cells, leading to a secondary skewing of the cellular population in the organism. Understanding RMAE will allow appreciation of a gene regulation mechanism that can impact both normal brain function and brain pathology.

Here, we analyze *FOXP2,* a gene that was close to, but not above, the rather strict threshold we used for calling RMAE in our prior genome-scale analysis of RMAE (15). Mutations in *FOXP2,* a forkhead domain-containing transcription factor, are known to cause an autosomal-dominant condition characterized by abnormal development of several brain areas critical for both developmental verbal movements and sequential articulation, with largely incomprehensible speech and marked disruption of multiple aspects of grammar and language (35–37). Truncating mutations and translocations interrupting the *FOXP2* gene sequence are sufficient to cause the phenotype, which has led to the conclusion that haploinsufficiency is the likely etiology; in other words, a single copy of a WT allele is incapable of providing sufficient protein for normal function (38–41). Thus, haploinsufficiency is different from both a dominant-negative effect wherein a nonfunctional mutant polypeptide interferes with the function of the normal allele, or gain-of-function mutations, wherein a changed gene product is endowed with a new and abnormal function. In this study, we analyze a deletion 3 Mb away from *FOXP2,* which, although it does not include the *FOXP2* gene itself, leads to loss of expression from the allele in cis with the deletion. We also observe data in controls consistent with *FOXP2* being subject to RMAE, thus leading to the idea that the haploinsufficient phenotype of *FOXP2* loss-of-function alleles, in this case and others, is due to some cells expressing no FOXP2 protein, rather than the classical view of haploinsufficiency, wherein every cell expresses half the normal amount.

Results

Phenotypic Description. A neurodevelopmental deficit in acquisition of spoken language across intact intelligence is termed specific language impairment (SLI). Importantly, specific language impairment is not part of a larger diagnosable neurological syndrome such as global developmental delay or autism. Specific language impairment has a prevalence of ~2–7% (42, 43) in children

Downloaded at Palestinian Territory, occupied on November 27, 2021

www.manaraa.com

entering school and is associated with later difficulties in learning to read. Closely related to SLI is Speech-language disorder-1 (SPCH1), an autosomal-dominant disorder, characterized by severe developmental verbal dyspraxia (DVD), that results in marked disruption of speech and expressive language, with virtually every aspect of grammar and language affected (35). The disorder is characterized by abnormal development of several brain areas critical for both developmental verbal movements and sequential articulation and is caused by rare heterozygous mutations in the *FOXP2* gene (36, 37). We performed a comprehensive case study of patient HS, who suffers from profound developmental verbal dyspraxia and dyslexia with preserved intelligence and behavior. Patient HS, a female, also had unilateral eye-tracking deficits properly characterized as Duane anomaly.

Blood samples from the proband and a matching female control were independently obtained and subjected to whole-blood RNA and DNA extraction in parallel whereas B lymphocytes were EBV-immortalized following standard procedures. Cells were pelleted for RNA extraction 4 wk after immortalization.

Proband-Specific and Allele-Specific *FOXP2* Analyses. Because mutations in the human *FOXP2* gene are known to cause developmental verbal dyspraxia, we analyzed the *FOXP2* gene in the proband. DNA sequencing of all 17 exons, as well as the intron–exon junctions of *FOXP2* in the proband, revealed no mutations in *FOXP2*. We therefore carried out whole-genome comparative genomic hybridization (CGH) microarray analysis using the Agilent 244K microarray chip to identify large-scale deletions/duplications genome-wide. CGH revealed a 2-Mb microdeletion on chromosome 7q31 (hg19 coordinates: chr7:109049659–111130658), 3 Mb centromeric to *FOXP2*. This deletion eliminated four exons of *IMMP2L* [IMP2 inner mitochondrial membrane peptidase-like (*Saccharomyces cerevisiae*)], as well as the entire *LRRN3* (leucine rich repeat neuronal 3) gene embedded within an intron of *IMMP2L*. Additionally, a noncoding RNA, *EIF3IP1* (eukaryotic translation initiation factor 3, subunit I pseudogene 1), was deleted (Fig. 1*A*). Parental analysis using multiplex ligation-dependent probe amplification (MLPA) showed that the microdeletion was not present in either parent and was therefore de novo. Note that array CGH also excluded deletions or duplications of the *FOXP2* gene itself, as well as other genomic deletions or duplications.

The Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER) database (44) revealed that none of the deleted genes have been reported to cause developmental verbal dyspraxia. We therefore considered the possibility that the deletion may be disrupting a *FOXP2* regulatory element, given the proximity of the microdeletion to the *FOXP2* locus on 7q31.

We sought to determine whether aberrant *FOXP2* expression was present in the proband. We analyzed allele-specific expression of *FOXP2* by Sanger sequencing of *FOXP2* cDNA derived from white-blood cells, which revealed that expression derived from one of the two alleles only (Fig. 1*B*). Allele-specific expression analysis of *FOXP2* transcription in B lymphoblastoid cells from a control individual (without the microdeletion near *FOXP2*) revealed equal biallelic *FOXP2* expression, thus ruling out imprinting (Fig. 1*B*), consistent with other reports (45).

To examine allele-specific expression of other genes in the vicinity of the microdeletion, we used dbSNP to find candidate SNPs in nearby, expressed genes, focusing on SNPs within the mature transcript (coding exons and UTRs). Sequence analysis of PCR products of eight nearby genes in the proband found heterozygous SNPs in *IFRD1* (IFN-related developmental regulator 1), *DOCK4* (dedicator of cytokinesis 4), and *NRCAM* (neuronal cell adhesion molecule). Five other genes in the vicinity either were not expressed or lacked heterozygous SNPs. We extracted RNA from the proband's whole blood and from isolated B lymphocytes and performed reverse transcriptase PCR (RT-PCR). Comparing cDNA-derived genotypes to the genotype obtained from genomic DNA revealed the presence of biallelic expression of all three of these genes (Fig. 2*A*).

We next examined the parent of origin of the deletion 3 Mb away from *FOXP2* as well as the parent of origin of expression of *FOXP2* in the proband. SNP genotyping of 30 SNPs within the microdeletion region on both parents and the proband revealed the paternal origin of the microdeletion (Table 1). Comparison of the proband's DNA and cDNA genotypes with parental genotypes established that the proband's *FOXP2* monoallelic expression derived from the maternally inherited allele (Fig. 2*B*).

Random Monoallelic *FOXP2* Expression in Control Tissue. To explore allele specificity of *FOXP2* expression in normal individuals, we

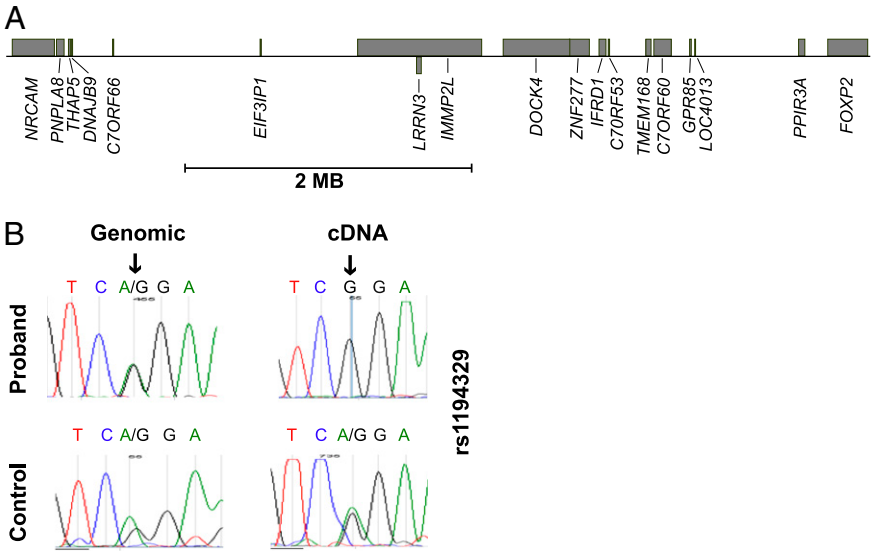

Fig. 1. Proband deletion region and allele-specific expression analysis. (*A*) Chromosome 7q31 deletion region in proband. Gray bars represent the genes; note that *LRRN3* is shown below the line merely to make it clear that it resides within the *IMMP2L* gene. (*B*) Comparison of genomic and cDNA showing biallelic *FOXP2* expression in control and monoallelic expression in the proband. Another SNP (rs1916980) also revealed monoallelic expression of *FOXP2*.

Downloaded at Palestinian Territory, occupied on November 27, 2021
www.manaraa.com

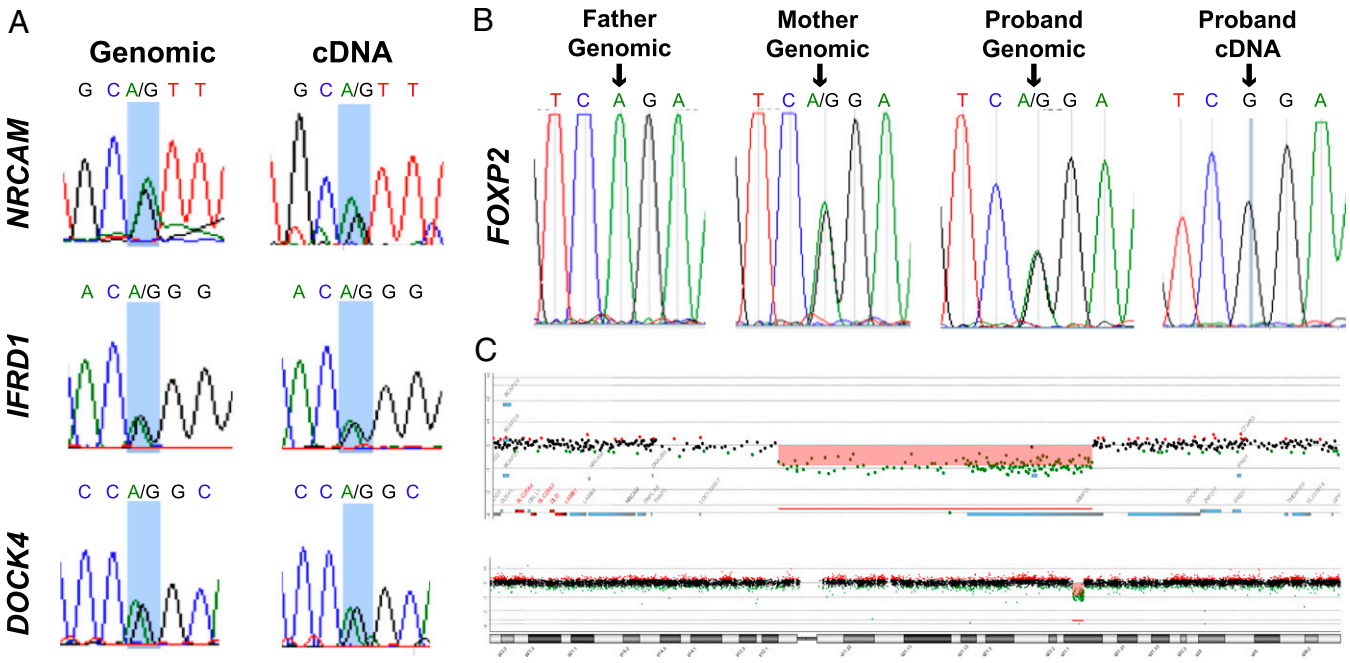

Fig. 2. Proband microarray and expression of genes flanking the deleted region. (*A*) cDNA genotypes of genes flanking the microdeletion in the proband show biallelic expression. (*B*) Parental (genomic) DNA genotypes and proband (genomic and cDNA) genotypes at rs1194329 confirm exclusively maternal expression of *FOXP2* in the proband. (*C*) Proband microarray with loss of probes at a 2-Mb locus on 7q31.

evaluated *FOXP2* expression in clonal B lymphoblastoid cell lines and clonal T-cell lines using the Affymetrix Human Mapping 500K array. We also carried out direct sequencing of *FOXP2* cDNA to validate the array findings. Briefly, the B-lymphoblastoid clones were those reported on in Gimelbrant et al. (15), and T-cell clones were derived from negatively isolated naive $CD4^+$ T cells activated with anti-CD3/28 beads (46). Both B- and T-cell clones were analyzed for allele-specific expression across the genome using the Affymetrix Human Mapping 500K array to obtain allele-specific expression information by comparing cDNA genotypes with regular DNA genotypes as we had previously (15) (Fig. 3*A*). We carried out locus validation by direct sequencing of cDNAs from these clonal B-cell lines. Nonclonal cells showed biallelic expression, ruling out imprinting. However, when subcloned cell lines were analyzed for allele-specific expression of *FOXP2*, we observed a pattern indicative of RMAE. One of two clones from the same individual revealed monoallelic expression whereas the other clone had biallelic expression, both on the array and confirmed by direct sequencing of cDNA (Fig. 3*B*). This pattern is consistent with the pattern observed in genes previously reported on in the genome-wide survey of random monoallelic expression (15).

Discussion

The human *FOXP2* phenotype is autosomal-dominant and thought to be a classic case of haploinsufficiency wherein two functional copies of *FOXP2* are necessary for acquisition of normal spoken language. What is the normal context within which to consider allele-specific perturbations of *FOXP2* expression? For most genes, it is generally assumed that there is biallelic expression, which would lead, in the case of a hemizygous deletion, to halving the gene dosage in all cells. Before our work, it was known that *FOXP2* translocations, deletions, and truncating mutations are disease-causing (37–41), and thus the presumptive etiology would be "homogeneous" haploinsufficiency wherein all cells express half the normal amount of FOXP2. However, here, we present data indicating that *FOXP2* is subject to RMAE. For each gene regulated by RMAE, an autosomal mechanism similar to X-inactivation, a random allelic choice initially made by each individual cell during development, is followed by a stable mitotic transmission of monoallelic expression to all daughter cells. In nonmutant individuals, some cells express the maternal allele, some cells express the paternal allele, and some cells express both alleles. In the case of a deletion eliminating a positively acting cis-regulatory element, there is the potential for a substantial fraction of the cells to express no *FOXP2* RNA because those were cells that should have expressed solely from the allele in cis with the deletion.

Thus, the autosomal-dominant phenotype of all *FOXP2* loss-of-function mutations may derive from the complete absence of functional transcript in the subset of cells that have inactivated the WT allele. It is worth noting that, as has been observed for X-linked mutations in females, it is possible that cells with no functional expression of *FOXP2* could be selected against at the cellular level during neural development. Indeed, the stochastic nature of allelic choice and clonal expansions during development could lead either to mitigation or enhancement of the phenotype. A caveat to the work we present is that we are analyzing allele-specific expression of *FOXP2* in immune system-derived cell lines. However, we note that RMAE is most readily observed in clonal cell lines, and thus direct interrogation of human brain neurons in general, and neurons responsible for control and fine coordination of the vocal apparatus in particular, is not feasible.

The causal association of *FOXP2* haploinsufficiency and developmental verbal dyspraxia is well-established, and the observed

Table 1. Informative SNPs in proband and parents within the microdeletion region

SNP	Maternal genotype	Paternal genotype	Proband genotype
rs425466	GA	A	G
rs1721899	G	A	G
rs3107945	AG	A	G

Observed genotypes are consistent with a deletion of the paternally inherited chromosome 7q31.

Downloaded at Palestinian Territory, occupied on November 27, 2021
www.manaraa.com

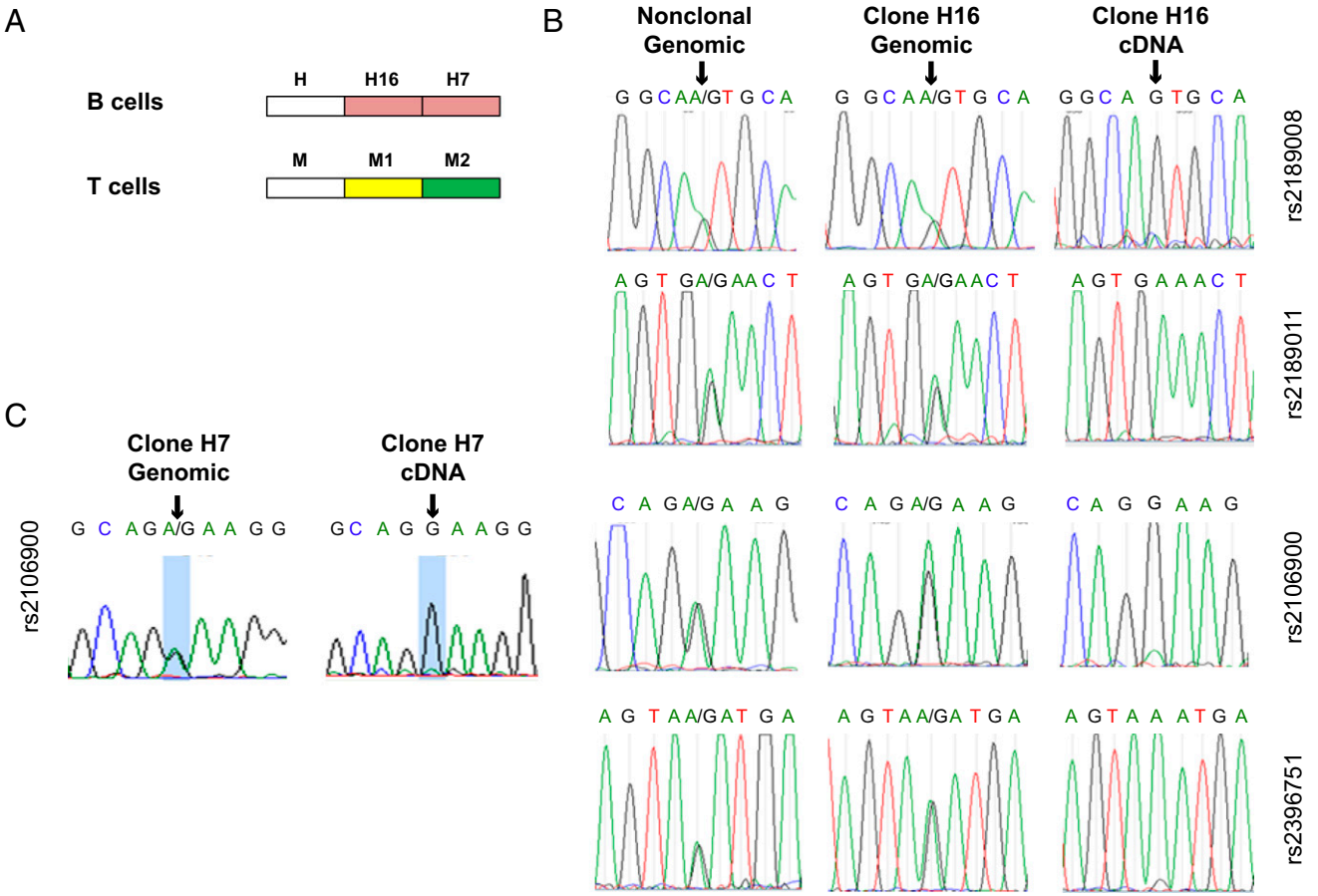

Fig. 3. RMAE assessed in normal cell lines. (*A*) A representation of the SNP array cDNA data for the *FOXP2* gene assessed for one SNP in B-cell subclones from one individual (15) and in T-cell subclones from a second individual. White, no signal; yellow, biallelic signal indicating expression of both alleles; pink, monoallelic maternal; green, monoallelic unknown. (Monoallelic for the paternal allele would have been blue, as in ref. 15.) H and M are nonclonal cell lines from which subclones (H16 and H7, B-cell clones, and M1 and M2, T-cell clones, respectively) were derived. (*B*) Electropherograms showing SNP genotypes in nonclonal B cells for individual H, as well as genomic DNA and cDNA genotypes for clone H16 at four different SNP loci showing heterozygosity in genomic and homozygosity in cDNA (arrowed). (*C*) A clone H7 SNP example is displayed.

association of the former with the latter in the proband is therefore biologically consistent and requires no further major elucidation. To investigate what aspects of the proband phenotype were ascribable to the deleted genes, we carried out a comparative assessment of overlapping *IMMP2L* and *LRRN3* microdeletion cases in the DECIPHER database and the database of genomic variants (DGV) (44, 47). Both genes are subject to polymorphic germ-line CNVs in humans that include exon deletions and duplications. We conclude that loss of *IMMP2L* and *LRRN3* is most likely responsible for the proband's dyslexia. The available evidence on the effect of isolated comorbid haploinsufficiency of both genes is found in the description of a two-generation pedigree with dyslexia associated with a genetic lesion that completely removed one copy of LRRN3 and rendered one copy of *IMMP2L* nonfunctional (48). Other clinical reports do not describe isolated simultaneous deletions of these two genes or fail to establish to a high degree of confidence the pathogenicity of the gene perturbation.

Hints of involvement of the relatively large (900-kb) *IMMP2L* gene in neurodevelopmental disorders are suggested by reports of association of autism spectrum disorder (ASD) with SNPs within the gene boundaries, which have not been validated by follow-up studies involving direct sequencing of the gene in large patient cohorts (49). Thus, the Online Database of Mendelian Inheritance in Man (OMIM) lists no phenotypes in association with allelic variants in *IMMP2L* or *LRRN3* (50), reflecting the absence of strong evidence on phenotypes attributable to these genes individually. Note that, in the two-generation pedigree mentioned in the previous paragraph (48), both genes are affected, thus mirroring closely the genetic loss in the proband analyzed here. Furthermore, humans with chromosome imbalances listed in the DECIPHER database involving *IMMP2L* and *LRRN3* predominantly had complex deletions involving seven or more other genes. Where language deficits were observed, the individuals presented with nonspecific language delay occurring within the context of global developmental delays, intellectual disability, or autism. The language delays would not be properly characterized as specific language impairment due to the extraneous presence of complicating factors. Indeed, deletion of *FOXP2* is the common factor in almost all reported developmental verbal dyspraxia 7q31 microdeletions reported to date, irrespective of whether these deletions also include *IMMP2L* or *LRRN3*, suggesting that this gene is a central cause of developmental verbal dyspraxia-related phenotypes in that region.

Of note, all prior reported instances of *FOXP2* haploinsufficiency (monoallelic expression) have been associated clinically with developmental verbal dyspraxia. The most striking example is a large three-generation family with developmental verbal dyspraxia where all 15 affected individuals were heterozygous for a *FOXP2* mutation, whereas all 12 unaffected individuals tested were homozygous for the wild type, showing perfect segregation (37), suggesting that the phenotypic expression of haploinsufficiency in this gene is highly penetrant.

Downloaded at Palestinian Territory, occupied on November 27, 2021
www.manaraa.com

A DECIPHER case with a single-gene deletion of 191 kb affecting exons 1–4 of *IMMP2L*, inherited from an unaffected parent, is listed in the database. The DECIPHER case shares with this proband the phenotype of Duane anomaly, a condition that normally occurs with a frequency of 1/1,000. Duane anomaly is not typically associated with *FOXP2* haploinsufficiency. Duane anomaly and specific language impairment are not commonly coincident, and the presence of Duane anomaly in the proband from this study, together with the 191-kb deletion from DECIPHER, suggests that *IMMP2L* may merit further investigation as a candidate gene for Duane anomaly.

The agent for cis-regulation of *FOXP2* is likely to be an enhancer or other cis-regulatory element within the microdeletion boundaries. The VISTA enhancer database (51) does not list any previously characterized enhancers within that region although at least two accelerated conserved noncoding sequences (ACNSs) are located within the microdeletion boundaries in the intragenic region between *IMMP2L* and *NRCAM*. The finding of two ACNSs of 992 across the genome described in Prabhakar et al. (52) is not statistically significant.

A number of studies have explored the mechanisms that regulate autosomal randomly monoallelically expressed genes. Although these investigations have been traditionally restricted to characterizing mechanisms that determine expression from individual loci (19), more recently, they have been broadened in an attempt to understand whether the various known gene regulatory mechanisms, including epigenetic marks and noncoding RNAs, might regulate autosomal randomly monoallelically expressed genes across the autosomes (17, 18, 53). The known mechanisms involved in establishing and maintaining RMAE are diverse, and it will be interesting to see the extent to which different mechanisms, as they are uncovered, will provide unifying concepts.

In conclusion, all different types of monoallelic expression have the potential to impact genotype–phenotype correlation. The expansion in the number of known randomly monoallelically expressed autosomal genes is remarkable and has the potential to impact our understanding of normal physiology, as well as pathological states. Instead of being mostly restricted to the immune system and chemosensory systems, it is now apparent that RMAE impacts a wide variety of different genes. We have focused here on neurons and brain pathophysiology, but RMAE has the potential to impact all organs. Moreover, RMAE can have an impact on phenotypic variability even without the presence of a clear mutation; there can be polymorphisms between two alleles of a given gene leading to expression of distinguishable proteins, and even if the two alleles are identical, RMAE can lead to different levels of expression in distinct cells as some cells express two alleles whereas others express only one allele (15). This type of cellular variability has the possibility of leading to variability in phenotype at the organismal level. Stochastic alternative splicing and stochastic alternative promoter use have been proposed as mechanisms for establishing unique neuronal identity (54–56), and RMAE has the possibility of subserving a similar function (15). Finally, as RMAE's mechanisms and impact on phenotype become better understood, it will be of interest to consider the potential effect of RMAE on the establishment and continued evolution of gene families, as we have discussed in detail previously (19).

Materials and Methods

Study Protocol. This study was conducted with the approval of the Children's Hospital Boston Institutional Review Board. All participants who were able, and parents or caregivers, gave written informed consent.

Proband Array CGH and *FOXP2* Sequencing. DNA extraction from whole blood obtained from the female proband and both unaffected parents was performed using the Qiagen DNA Blood Mini kit according to the manufacturer's guidelines (Qiagen). Parallel CGH microarray analysis on the proband and a female control individual was performed using the Agilent 244K microarray chip (Agilent). Analysis of parental samples for abnormal findings in the proband was carried out using targeted Multiplex Ligand Probe Amplification (MLPA) (MRC-Holland) of parent DNA according to the manufacturer's guidelines. All 17 exons and intron–exon junctions of *FOXP2* were sequenced from genomic DNA extracted from proband lymphocytes on the Applied Biosystems 3730XL DNA Analyzer.

Proband Allele-Specific Expression Analysis. RNA was extracted from the proband's whole blood using the PAXgene Blood RNA System Kit according to the manufacturer's guidelines, and from EBV-immortalized peripheral B lymphocytes, using the RNeasy Mini Kit (Qiagen). Total RNA was then diluted appropriately and treated with Turbo DNAfree (Ambion) according to the manufacturer's protocol for "strong DNA contamination." No significant DNA contamination remained, judging by lack of PCR amplification in the absence of reverse transcriptase (RT). The digestion of DNA allowed the evaluation of heterogeneous nuclear RNA (hnRNA) for intronic SNPs. RT-PCR was performed using the Qiagen one-step RT-PCR Kit. Primers were designed using the Primer3 program to flank SNPs in genes of interest (*FOXP2*, *DOCK4*, *IFRD1*, and *NRCAM*), and these SNPs were amplified from genomic DNA and from cDNA samples using intron-spanning primers when possible. Sequencing of amplicons was performed on an Applied Biosystems 3730XL DNA Analyzer, as before. cDNA-derived genotypes were compared with the genotype obtained from genomic DNA. Primers are available upon request.

Proband–Parental SNP Genotyping Analysis. To allow comparison of parental and proband genotypes in the microdeletion region, SNP genotyping was performed using the Sequenom iPLEX Gold Genotyping Assay in a multiplexed assay, using the MassARRAY System Designer software to design both PCR and MassEXTEND primers for multiplexed assays. For genotype calling, the extension products were spotted onto SpectroCHIP arrays and analyzed using a MALDI-TOF mass spectrometer.

Analysis of Random Monoallelic Expression. Analysis of random monoallelic expression was as previously described (15). Briefly, lymphoblast cell lines were obtained from the Coriell biorepository and subcloned as in ref. 15, and T-cell clones were derived from negatively isolated naive $CD4^+$ T cells activated with anti-CD3/28 beads as described previously (46). B- and T-cell clones were analyzed for allele-specific expression across the genome essentially as in ref. 15. Allele-specific expression was analyzed by comparing the "genotype" from cDNA with the regular DNA genotype. A finding of a homozygous cDNA genotype along with a heterozygous regular genomic DNA genotype indicates monoallelic expression. B-cell data were derived from the XbaI half of the Affymetrix Human Mapping 100k array. To validate the array findings, samples of cDNA from clones H7 and H16, as well as nonclonal cDNA from individual H, were subsequently subjected to PCR as previously described in *Proband Allele-Specific Expression Analysis*. Primers flanking five SNPs in *FOXP2* were designed using the Primer3 program, and used for PCR amplification from genomic DNA and from cDNA samples. Sanger sequencing was carried out as described in *Proband Allele-Specific Expression Analysis*. cDNA-derived genotypes were compared with the genotype obtained from genomic DNA.

ACKNOWLEDGMENTS. We thank present and past members of the A.C. laboratory for helpful discussions. This work was supported by grants from the National Institutes of Health (to A.C.), by a Brain and Behavior Foundation Young Investigator Award (to A.A.A.), and by a Pew Scholar award (to A.G.).

1. De Gobbi M, et al. (2006) A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science* 312(5777):1215–1217.
2. Pastinen T (2010) Genome-wide allele-specific analysis: Insights into regulatory variation. *Nat Rev Genet* 11(8):533–538.
3. Ferguson-Smith AC (2011) Genomic imprinting: The emergence of an epigenetic paradigm. *Nat Rev Genet* 12(8):565–575.
4. Nicholls RD, Knoll JH, Butler MG, Karam S, Lalande M (1989) Genetic imprinting suggested by maternal heterodisomy in nondeletion Prader-Willi syndrome. *Nature* 342(6247):281–285.
5. Buiting K, et al. (1995) Inherited microdeletions in the Angelman and Prader-Willi syndromes define an imprinting centre on human chromosome 15. *Nat Genet* 9(4): 395–400.

COLLOQUIUM PAPER
GENETICS
Downloaded at Palestinian Territory, occupied on November 27, 2021
www.manaraa.com

6. Buiting K (2010) Prader-Willi syndrome and Angelman syndrome. *Am J Med Genet C Semin Med Genet* 154C(3):365–376.
7. Glenn CC, et al. (1996) Gene structure, DNA methylation, and imprinted expression of the human SNRPN gene. *Am J Hum Genet* 58(2):335–346.
8. Kubota T, et al. (1997) Methylation-specific PCR simplifies imprinting analysis. *Nat Genet* 16(1):16–17.
9. Peters J (2014) The role of genomic imprinting in biology and disease: An expanding view. *Nat Rev Genet* 15(8):517–530.
10. Lawson HA, Cheverud JM, Wolf JB (2013) Genomic imprinting and parent-of-origin effects on complex traits. *Nat Rev Genet* 14(9):609–617.
11. Jeon Y, Sarma K, Lee JT (2012) New and Xisting regulatory mechanisms of X chromosome inactivation. *Curr Opin Genet Dev* 22(2):62–71.
12. Amir RE, et al. (2000) Influence of mutation type and X chromosome inactivation on Rett syndrome phenotypes. *Ann Neurol* 47(5):670–679.
13. Plenge RM, Stevenson RA, Lubs HA, Schwartz CE, Willard HF (2002) Skewed X-chromosome inactivation is a common feature of X-linked mental retardation disorders. *Am J Hum Genet* 71(1):168–173.
14. Gibson JH, Williamson SL, Arbuckle S, Christodoulou J (2005) X chromosome inactivation patterns in brain in Rett syndrome: Implications for the disease phenotype. *Brain Dev* 27(4):266–270.
15. Gimelbrant A, Hutchinson JN, Thompson BR, Chess A (2007) Widespread monoallelic expression on human autosomes. *Science* 318(5853):1136–1140.
16. Zwemer LM, et al. (2012) Autosomal monoallelic expression in the mouse. *Genome Biol* 13(2):R10.
17. Gendrel AV, et al. (2014) Developmental dynamics and disease potential of random monoallelic gene expression. *Dev Cell* 28(4):366–380.
18. Eckersley-Maslin MA, et al. (2014) Random monoallelic gene expression increases upon embryonic stem cell differentiation. *Dev Cell* 28(4):351–365.
19. Chess A (2012) Mechanisms and consequences of widespread random monoallelic expression. *Nat Rev Genet* 13(6):421–428.
20. Hozumi N, Tonegawa S (1976) Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc Natl Acad Sci USA* 73(10):3628–3632.
21. Held W, Roland J, Raulet DH (1995) Allelic exclusion of Ly49-family genes encoding class I MHC-specific receptors on NK cells. *Nature* 376(6538):355–358.
22. Bix M, Locksley RM (1998) Independent and epigenetic regulation of the interleukin-4 alleles in CD4+ T cells. *Science* 281(5381):1352–1354.
23. Holländer GA, et al. (1998) Monoallelic expression of the interleukin-2 locus. *Science* 279(5359):2118–2121.
24. Chess A, Simon I, Cedar H, Axel R (1994) Allelic inactivation regulates olfactory receptor gene expression. *Cell* 78(5):823–834.
25. Malnic B, Hirono J, Sato T, Buck LB (1999) Combinatorial receptor codes for odors. *Cell* 96(5):713–723.
26. Vassar R, et al. (1994) Topographic organization of sensory projections to the olfactory bulb. *Cell* 79(6):981–991.
27. Ressler KJ, Sullivan SL, Buck LB (1994) Information coding in the olfactory system: Evidence for a stereotyped and highly organized epitope map in the olfactory bulb. *Cell* 79(7):1245–1255.
28. Goate A, et al. (1991) Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 349(6311):704–706.
29. Tanzi RE, et al. (1992) Assessment of amyloid beta-protein precursor gene mutations in a large set of familial and sporadic Alzheimer disease cases. *Am J Hum Genet* 51(2):273–282.
30. Polymeropoulos MH, et al. (1997) Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* 276(5321):2045–2047.
31. Singleton AB, et al. (2003) alpha-Synuclein locus triplication causes Parkinson's disease. *Science* 302(5646):841.
32. Chartier-Harlin MC, et al. (2004) Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet* 364(9440):1167–1169.
33. Theuns J, et al. (2006) Promoter mutations that increase amyloid precursor-protein expression are associated with Alzheimer disease. *Am J Hum Genet* 78(6):936–946.
34. Ross OA, et al. (2008) Genomic investigation of alpha-synuclein multiplication and parkinsonism. *Ann Neurol* 63(6):743–750.
35. Fisher SE, Vargha-Khadem F, Watkins KE, Monaco AP, Pembrey ME (1998) Localisation of a gene implicated in a severe speech and language disorder. *Nat Genet* 18(2):168–170.
36. Vargha-Khadem F, et al. (1998) Neural basis of an inherited speech and language disorder. *Proc Natl Acad Sci USA* 95(21):12695–12700.
37. Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413(6855):519–523.
38. MacDermot KD, et al. (2005) Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am J Hum Genet* 76(6):1074–1080.
39. Rice GM, et al. (2012) Phenotype of FOXP2 haploinsufficiency in a mother and son. *Am J Med Genet A* 158A(1):174–181.
40. Zilina O, et al. (2012) Maternally and paternally inherited deletion of 7q31 involving the FOXP2 gene in two families. *Am J Med Genet A* 158A(1):254–256.
41. Turner SJ, et al. (2013) Small intragenic deletion in FOXP2 associated with childhood apraxia of speech and dysarthria. *Am J Med Genet A* 161A(9):2321–2326.
42. Bartlett CW, et al. (2002) A major susceptibility locus for specific language impairment is located on 13q21. *Am J Hum Genet* 71(1):45–55.
43. O'Brien EK, Zhang X, Nishimura C, Tomblin JB, Murray JC (2003) Association of specific language impairment (SLI) to the region of 7q31. *Am J Hum Genet* 72(6):1536–1543.
44. Firth HV, et al. (2009) DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet* 84(4):524–533.
45. Thomas AC, et al. (2012) The speech gene FOXP2 is not imprinted. *J Med Genet* 49(11):669–670.
46. Volpe E, et al. (2008) A critical function for transforming growth factor-beta, interleukin 23 and proinflammatory cytokines in driving and modulating human T(H)-17 responses. *Nat Immunol* 9(6):650–657.
47. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW (2014) The Database of Genomic Variants: A curated collection of structural variation in the human genome. *Nucleic Acids Res* 42(Database issue):D986–D992.
48. Girirajan S, et al. (2011) Relative burden of large CNVs on a range of neurodevelopmental phenotypes. *PLoS Genet* 7(11):e1002334.
49. Petek E, et al. (2007) Molecular and genomic studies of IMMP2L and mutation screening in autism and Tourette syndrome. *Mol Genet Genomics* 277(1):71–81.
50. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA (2005) Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res* 33(Database issue):D514–D517.
51. Visel A, Minovitsky S, Dubchak I, Pennacchio LA (2007) VISTA Enhancer Browser: A database of tissue-specific human enhancers. *Nucleic Acids Res* 35(Database issue):D88–D92.
52. Prabhakar S, Noonan JP, Pääbo S, Rubin EM (2006) Accelerated evolution of conserved noncoding sequences in humans. *Science* 314(5800):786.
53. Nag A, et al. (2013) Chromatin signature of widespread monoallelic expression. *eLife* 2:e01256.
54. Neves G, Chess A (2004) Dscam-mediated self- versus non-self-recognition by individual neurons. *Cold Spring Harb Symp Quant Biol* 69:485–488.
55. Neves G, Zucker J, Daly M, Chess A (2004) Stochastic yet biased expression of multiple Dscam splice variants by individual cells. *Nat Genet* 36(3):240–246.
56. Lefebvre JL, Kostadinov D, Chen WV, Maniatis T, Sanes JR (2012) Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature* 488(7412):517–521.

Downloaded at Palestinian Territory, occupied on November 27, 2021
www.manaraa.com