Developmental Cell
Article

Random Monoallelic Gene Expression Increases upon Embryonic Stem Cell Differentiation

Mélanie A. Eckersley-Maslin,[1,2] David Thybert,[3] Jan H. Bergmann,[1] John C. Marioni,[3] Paul Flicek,[3] and David L. Spector[1,2,*]
[1]Cold Spring Harbor Laboratory, One Bungtown Road, Cold Spring Harbor, NY 11724, USA
[2]Watson School of Biological Sciences, Cold Spring Harbor Laboratory, One Bungtown Road, Cold Spring Harbor, NY 11724, USA
[3]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK
*Correspondence: spector@cshl.edu
http://dx.doi.org/10.1016/j.devcel.2014.01.017

SUMMARY

Random autosomal monoallelic gene expression refers to the transcription of a gene from one of two homologous alleles. We assessed the dynamics of monoallelic expression during development through an allele-specific RNA-sequencing screen in clonal populations of hybrid mouse embryonic stem cells (ESCs) and neural progenitor cells (NPCs). We identified 67 and 376 inheritable autosomal random monoallelically expressed genes in ESCs and NPCs, respectively, a 5.6-fold increase upon differentiation. Although DNA methylation and nuclear positioning did not distinguish the active and inactive alleles, specific histone modifications were differentially enriched between the two alleles. Interestingly, expression levels of 8% of the monoallelically expressed genes remained similar between monoallelic and biallelic clones. These results support a model in which random monoallelic expression occurs stochastically during differentiation and, for some genes, is compensated for by the cell to maintain the required transcriptional output of these genes.

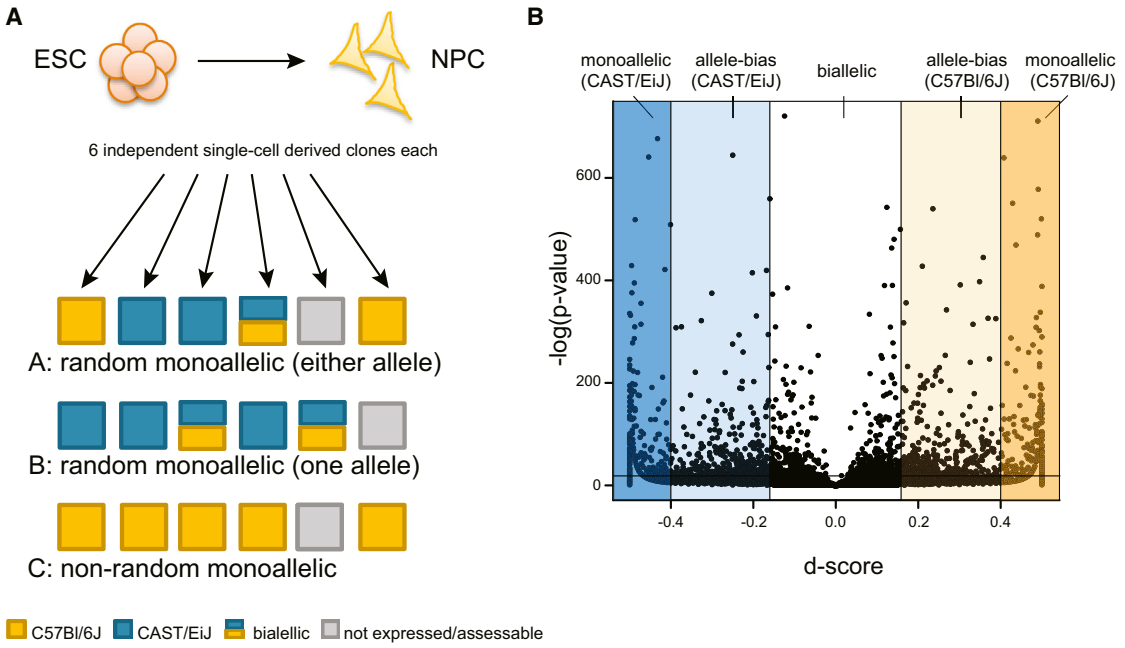
INTRODUCTION

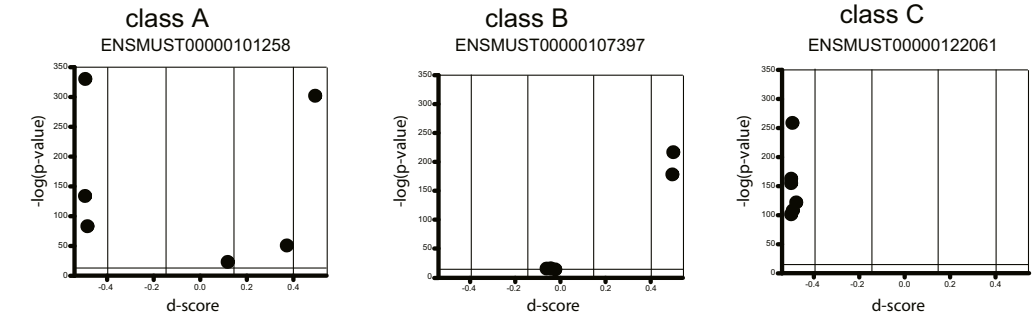
The majority of gene expression in diploid cells is carried out through expression of both alleles of each gene. However, several interesting cases of monoallelic expression, in which there is transcription from only one allele, have been documented. Well-characterized and extensively studied examples include X chromosome inactivation (reviewed in Guidi et al., 2004; Schulz and Heard, 2013) and genomic imprinting (reviewed in Bartolomei and Ferguson-Smith, 2011; McAnally and Yampolsky, 2010). Interestingly, random monoallelic expression can also occur on autosomes independently of parental origin and genotype (reviewed in Chess, 2012; Guo and Birchler, 1994). For example, the immune system utilizes monoallelic expression to ensure each B cell expresses a single uniquely rearranged immunoglobulin receptor (Pernis et al., 1965). Additionally, neurons express olfactory receptors (ORs) in a monogenic and monoallelic manner to provide cell identity and aid in neural connectivity (Chess et al., 1994). However, random autosomal monoallelic expression is not limited to specialized gene families, as it has been reported to occur at individual gene loci throughout the genome of a few cell types examined (Gimelbrant et al., 2007; Jeffries et al., 2012; Li et al., 2012; Zwemer et al., 2012). Yet, despite the identification of such genes, detailed molecular characterization and potential biological consequences of random monoallelic expression remain unknown.

The extent of random monoallelic expression varies from 2% in neural stem cells (Jeffries et al., 2012; Li et al., 2012) to 10% in lymphoblasts (Gimelbrant et al., 2007; Zwemer et al., 2012). Interestingly, only a small number of genes have been identified in common across these studies, suggesting that monoallelic expression may be established during development in a lineage- or cell-type-specific manner. However, random monoallelic expression has not been studied in the context of a developmental paradigm.

Exclusive expression from one allele renders the cell susceptible to loss-of-heterozygosity effects that could result in deleterious disease-related phenotypes. Monoallelic expression has been hypothesized to contribute to cellular diversity and identity, as is the case for ORs and immunoglobulins (reviewed in Chess, 2013), or may be a mechanism for regulating the transcriptional output of genes, although this has not been vigorously analyzed. Alternatively, rather than being an active process, the switch to monoallelic expression may instead reflect the stochastic nature of gene regulation occurring independently at the two alleles.

We performed an allele-specific RNA-sequencing screen for random autosomal monoallelic expression during differentiation of mouse embryonic stem cells (ESCs) to neural progenitor cells (NPCs). Interestingly, we observed a 5.6-fold increase in monoallelic expression during differentiation, from just 67 genes (<0.5%) in ESCs to 376 genes (3.0%) in NPCs, indicating that the establishment of monoallelic expression occurs during early development. Detailed genomic and molecular characterization of these genes revealed that DNA methylation was not sufficient for the mitotic inheritance of monoallelic expression, nor was there evidence for differential nuclear positioning of active versus





D

		Genes	Transcripts	% Genes
ESC	class A	1	1	~
	class B	1665	2499	12.15
	class B filtered	66	73	0.48
	class C	547	721	3.99
	Total Random Monoallelic (A+B filtered)	**67**	**74**	**0.49**

NPC	class A	86	135	0.69
	class B	1874	3004	14.95
	class B filtered	302	466	2.41
	class C	276	381	2.2
	Total Random Monoallelic (A+B filtered)	**376**	**602**	**3**

Figure 1. Identification of Monoallelically Expressed Genes in ESCs and NPCs

(A) Schematic of the allele-specific RNA-sequencing screen used to identify monoallelically expressed genes. For both ESCs and NPCs, six single-cell-derived clones were generated and transcripts were categorized as either C57Bl/6J biased (orange), CAST/EiJ biased (blue), biallelic (orange + blue), or not expressed/assessable (gray). Transcripts were further grouped into three classes based on their expression across clones, with class A representing high-confidence random monoallelically expressed genes; class B, following additional filtering, representing additional monoallelically expressed genes; and class C representing nonrandom monoallelically expressed genes.

(legend continued on next page)

inactive alleles. However, specific histone modifications were sufficient to distinguish the active and inactive alleles, and likely contribute toward maintaining monoallelic expression across cell divisions. Interestingly, in a subset of monoallelically expressed genes, transcriptional compensation through upregulation of the single active allele preserved the biallelic levels of the respective mRNA in the cell. These results support a model where stochastic gene regulation during differentiation results in monoallelic expression and, for some genes, the cell is able to compensate transcriptionally to maintain the required transcriptional output of these genes. Therefore, random monoallelic expression exemplifies the stochastic and plastic nature of gene expression in single cells.

RESULTS

Identification of Monoallelically Expressed Genes upon Differentiation of Mouse Embryonic Stem Cells to Neural Progenitor Cells

To identify random autosomal monoallelically expressed genes in mouse ESCs and NPCs, we used male cells derived from an F1 hybrid between C57Bl/6J and CAST/EiJ mice in which the high density of SNPs allowed us to quantify allele-specific expression for 82.8% of transcripts. We expanded six single-cell-derived clones from both ESCs and induced NPCs (Figure 1A; Figure S1 available online). Assuming inheritance of monoallelic expression across cell divisions, all cells within each single-cell-derived clone are expected to express the same alleles. However, different clones should show a random selection of alleles, allowing the identification of mitotically inheritable random monoallelically expressed genes.

For each clone, $\sim 5 \times 10^7$ reads were mapped to both the C57Bl/6J and CAST/EiJ transcriptomes using the Burrows-Wheeler Aligner (BWA) (Li and Durbin, 2010) (Figure S1E). To control for the possible loss of heterozygosity, mouse diversity SNP arrays were run on genomic DNA, and transcripts within aneuploid regions were excluded from further analysis (see Supplemental Experimental Procedures). For all assessable transcripts, the number of reads corresponding to each allele at each SNP position (minimum of five-read coverage) was used to determine whether there was evidence of allele-specific expression based on two metrics: a d score representing the ratio of allelic expression, and a p value (see Supplemental Experimental Procedures). Assessable transcripts were then classified as monoallelic ($|\text{d score}| \geq 0.4$, p value $< 10^{-8}$), allele biased ($0.18 \leq |\text{d score}| < 0.4$, p value $< 10^{-8}$), or biallelic ($|\text{d score}| < 0.18$) (Figure 1B). Based on the patterns of expression bias observed across clones, transcripts were subsequently grouped into one of three classes of monoallelically expressed genes (Figure 1C), as follows. Class A transcripts had at least one clone classified as monoallelic for the C57Bl/6J allele and at least one for the CAST/EiJ allele, and represent high-confidence random monoallelically expressed genes, as they clearly show a random choice of allele. Class B transcripts had at least one clone classified as monoallelic for either the C57Bl/6J or CAST/EiJ alleles, but not both. These transcripts were further filtered to select those in which there was one high-confidence biased clone (p value $< 10^{-10}$, $|\text{d score}| > 0.35$) and one high-confidence biallelic clone ($|\text{d score}| < 0.1$). With a larger number of clones, these filtered class B transcripts would likely be assigned to class A. Finally, class C transcripts represent nonrandom monoallelically expressed genes. In this class, all clones showed bias toward the same allele with no evidence that the second allele is transcribed. These transcripts include imprinted genes, in addition to genes in which one allele was inactive due to a *cis* mutation, and as such were not included in further analysis. Genes could be assigned to more than one class if at least one corresponding transcript was in each class.

In ESCs, of the 13,699 assessable genes, only 1 was classified as class A, with another 66 class B filtered, giving a total of 67 monoallelically expressed genes or 74 transcripts, representing only 0.49% of assessable genes (Figure 1D; Tables S1 and S2). Interestingly, this low number of genes increased 5.6-fold during differentiation to 376 genes (86 class A and 302 class B filtered) in the NPCs, corresponding to 602 transcripts or 3.0% of assessable genes (Figure 1D; Tables S1 and S2). This set included *Thrsp* and several members of the protocadherin family, which have been previously reported to be monoallelically expressed (Esumi et al., 2005; Kaneko et al., 2006; Wang et al., 2007). The increase in monoallelic expression during differentiation suggests that the establishment of monoallelic expression occurs upon cell-fate specification early in development.

Validation of Monoallelically Expressed Genes

Validation of the screen was first performed by Sanger sequencing of PCR products containing informative exonic SNPs for 20 different genes in both ESCs and NPCs. Clones were classified as monoallelic or biallelic (Figure 2A), and subsequently compared to the RNA-sequencing screen results. Seventy-six of 82 (93%) PCR products were in agreement with the RNA-sequencing screen (Figures S2A and S2B), demonstrating the robustness of our approach.

Next, monoallelic expression was confirmed at the level of transcription by chromatin immunoprecipitation (ChIP) for all forms of RNA polymerase II. Levels of pull-down were similar between monoallelic and biallelic clones within the body of four randomly selected genes (Figure 2B). Importantly, Sanger sequencing of SNPs within the amplicons used confirmed that RNA polymerase II was specifically associated with only the active allele in monoallelic clones compared to both alleles in biallelic clones (Figure 2C), confirming that monoallelic expression is due to the exclusive transcription of only one of the two alleles in the cell.

We further validated monoallelic expression at single-cell resolution by RNA fluorescence in situ hybridization (RNA-FISH). By using fluorescently labeled probes targeting both exonic and

(B) Graphical representation of the p value and d score thresholds used to categorize transcripts (●) in a given clone as either monoallelic ($|\text{d score}| > 0.4$), allele biased ($0.18 < |\text{d score}| < 0.4$), or biallelic ($|\text{d score}| < 0.18$).
(C) Examples of class A, B, and C transcripts showing the behavior of individual clones (●) with respect to the d score and p value.
(D) Summary table of the number of genes and transcripts from class A, B, B filtered, and C for both ESCs and NPCs and the percentage of all assessable genes.
See also Figure S1 and Tables S1 and S2.

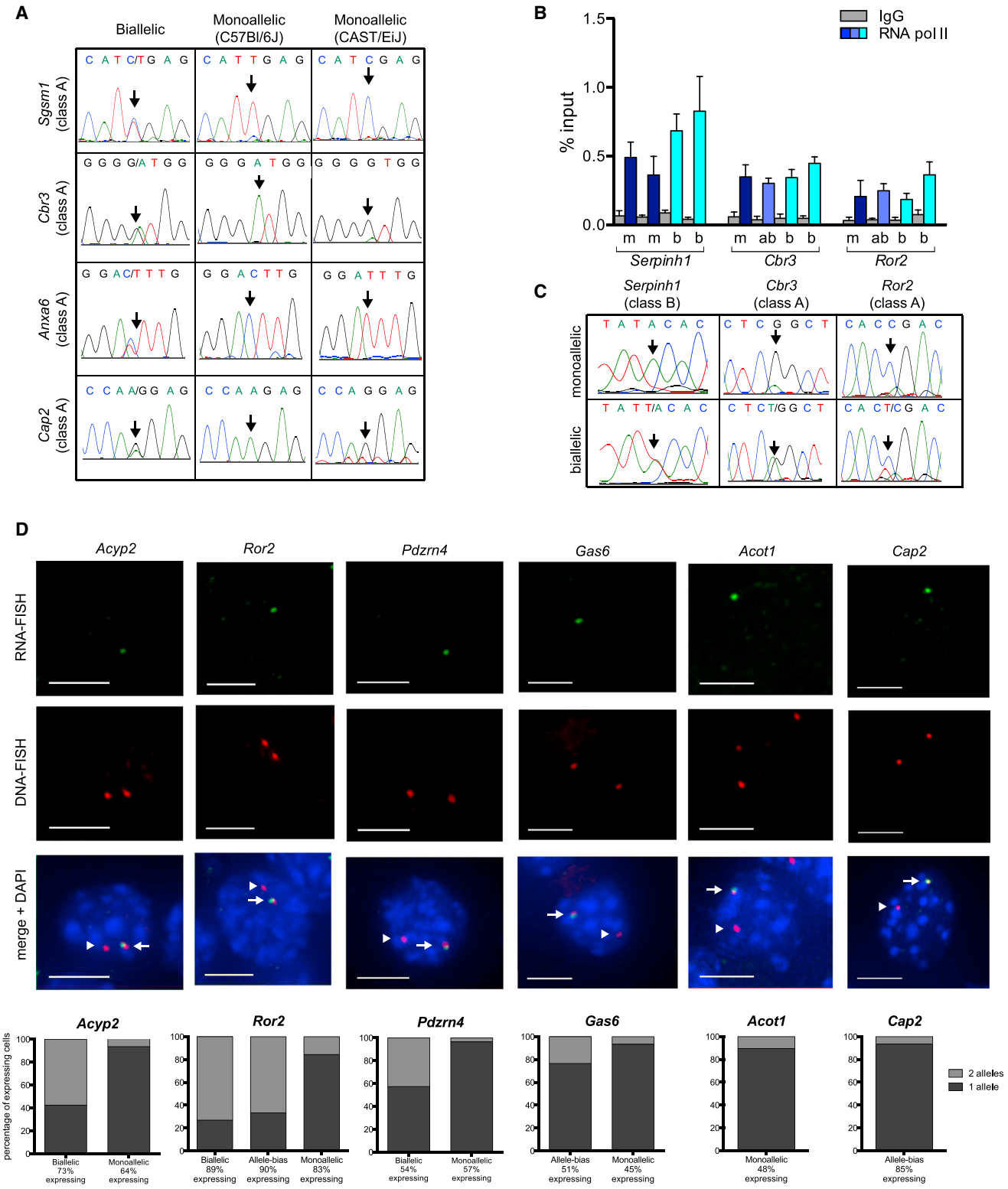


Figure 2. Validation of Monoallelic Gene Expression

(A) Representative traces from Sanger sequencing of PCR products containing informative exonic SNPs (arrows) from cDNA from biallelic (first column) or monoallelic (second and third columns) clones for four separate monoallelically expressed genes.

(B) Chromatin immunoprecipitation for RNA polymerase II large subunit (blue) or control IgG (gray) for three separate gene promoter regions between monoallelic (m; dark blue), allele-biased (ab; medium blue), or biallelic (b; light blue) clones. Error bars represent SEM of at least three biological replicates.

(legend continued on next page)

intronic sequences of the target gene, nascent RNA at the sites of transcription can be visualized as a fluorescent spot or spots within the nucleus (Figure 2D, first row). These RNA-FISH spots colocalize with the gene locus visualized by subsequent DNA-FISH in the same cells (Figure 2D, second row), confirming that they are indeed sites of transcription. We calculated the percentage of expressing cells exhibiting monoallelic or biallelic expression and successfully validated six out of six class A monoallelically expressed genes (Figure 2D). For example, a single *Acyp2* signal was detected by RNA-FISH in 93.5% of expressing cells in a monoallelic clone. In contrast, 57.5% of cells in a biallelic clone showed two active alleles of *Acyp2*. Likewise, expression from one allele was confirmed for 84.3%, 96.5%, 93.3%, and 89.6% of cells in monoallelic clones for *Ror2*, *Pdzrn4*, *Gas6*, and *Acot1*, respectively. In this way, the RNA-FISH analysis confirmed at single-cell resolution the results of RNA-sequencing analysis.

Importantly, RNA-FISH confirmed monoallelic expression for three class B genes in NPCs of a pure genetic background (Figure S2C). A single transcribing allele was observed in 54.8%, 67.7%, and 73.9% of expressing cells for *Atp1a2*, *Arap1*, and *Mavs*, respectively, confirming that monoallelic gene expression is independent of the genetic background and not due to differences between the two parental strains in the hybrid cell lines.

Dynamics of Monoallelic Expression during Differentiation

During differentiation there is a 5.6-fold increase in monoallelic expression, coinciding with the loss of pluripotency and gain of lineage commitment (Figure 1D). Interestingly, we observed very few (<2%) monoallelically expressed genes in common between ESCs and NPCs (Figure 3A). Instead, the majority of monoallelically expressed genes were biallelically expressed in the other cell type (Figures 3B, 3C, S3A, and S3B). Thus, monoallelic expression, although maintained across cell divisions, is not maintained during the transition from ESC to NPC.

Interestingly, for 98.9% of monoallelically expressed genes, at least one clone was either biallelic and/or did not express the respective gene (Table S2). Within a single clone, ~60% of the monoallelically expressed genes show biallelic expression (Figure 3D), suggesting that monoallelic expression may reflect variation in gene expression regulation between two homologous alleles. This contrasts with imprinted genes and X chromosome inactivation, where all cells exhibit strict monoallelic expression, and implies that, rather than being tightly regulated, random monoallelic expression is not an active decision required for cell survival or differentiation.

Importantly, the distribution of expression levels of the monoallelically expressed genes was not dramatically different from all assessable transcripts (Figures 3E, S3C, and S3D). The small yet statistically significant difference in the expression level for NPCs is unlikely to be of biological significance. Furthermore, reducing the stringent expression-level thresholds used for the screen did not result in a large increase in the number of monoallelically expressed genes (Table S3), confirming that monoallelically expressed genes have a similar expression profile to all expressed genes.

We next determined whether any genomic features of monoallelically expressed genes distinguished them from other expressed genes. Unlike for imprinting and ORs, the random monoallelically expressed genes were distributed throughout the genome and did not fall into any genomic clusters (Figure S4A). Monoallelically expressed genes showed similar GC density at their promoters to all assessable genes (Figure S4B), in contrast to the reduced GC density previously reported for ORs (Clowney et al., 2011). Analysis of 174 mammalian and 530 vertebrate transcription factor motifs revealed that although 10 motifs were differentially enriched at the promoters, they were not sufficient to distinguish monoallelically expressed genes from all assessable genes (Figure S4C). Furthermore, although there was a small decrease in evolutionary conservation of monoallelically expressed genes, this was not as dramatic as what is observed for ORs (Figure S4D). Finally, gene ontology analysis using DAVID (Huang et al., 2009) revealed a slight enrichment in glycoproteins involved in signaling (Figure S4E). Thus, random monoallelically expressed genes are not distinguished from other genes by these genomic features.

Finally, we compared the changes in expression levels of monoallelically expressed genes during differentiation (Figures 3F–3H). Expression of the majority of ESC monoallelically expressed genes either decreased (50%) or did not change (32.4%) during differentiation (Figure 3G), despite the majority being biallelically expressed in NPCs. Furthermore, only 13.1% of the NPC monoallelically expressed genes are expressed at lower levels in the NPCs compared to ESCs (Figure 3H), despite the fact that 55.2% switch from biallelic to monoallelic expression during differentiation (Figure 3B), again suggesting that monoallelic expression is not a mechanism for reducing transcript levels.

DNA Methylation Does Not Regulate Monoallelically Expressed Genes

One intriguing aspect of monoallelic expression is that the transcriptional imbalance between the active and inactive alleles is maintained across cell generations. DNA methylation is the most widely accepted mechanism through which the transcriptional state of a gene can be inherited and maintained in daughter cells (Smith and Meissner, 2013), and distinguishes active and inactive alleles of both imprinted (Kelsey and Feil, 2013) and X-linked genes (Schulz and Heard, 2013). To assess

(C) Representative traces of Sanger sequencing of ChIP products containing informative SNPs (arrows) revealing associated alleles for monoallelic and biallelic clones.

(D) RNA/DNA-FISH validation of six separate monoallelically expressed class A genes in NPCs. Representative 3D projections of RNA-FISH (first row, green) and DNA-FISH (second row, red) image stacks. The third row shows a merge of RNA- and DNA-FISH with DAPI to visualize total DNA (blue). Arrows denote actively transcribing alleles; arrowheads denote inactive alleles. The scale bars represent 5 μm. Bottom row shows quantification of independent clones for each of the six genes for NPC clones that were either biallelic, allele biased, or monoallelic for the respective gene. Percentage of expressing cells having either one (dark gray) or two (light gray) RNA-FISH signals representing monoallelic and biallelic cells, respectively. One hundred cells were analyzed per sample.
See also Figure S2.

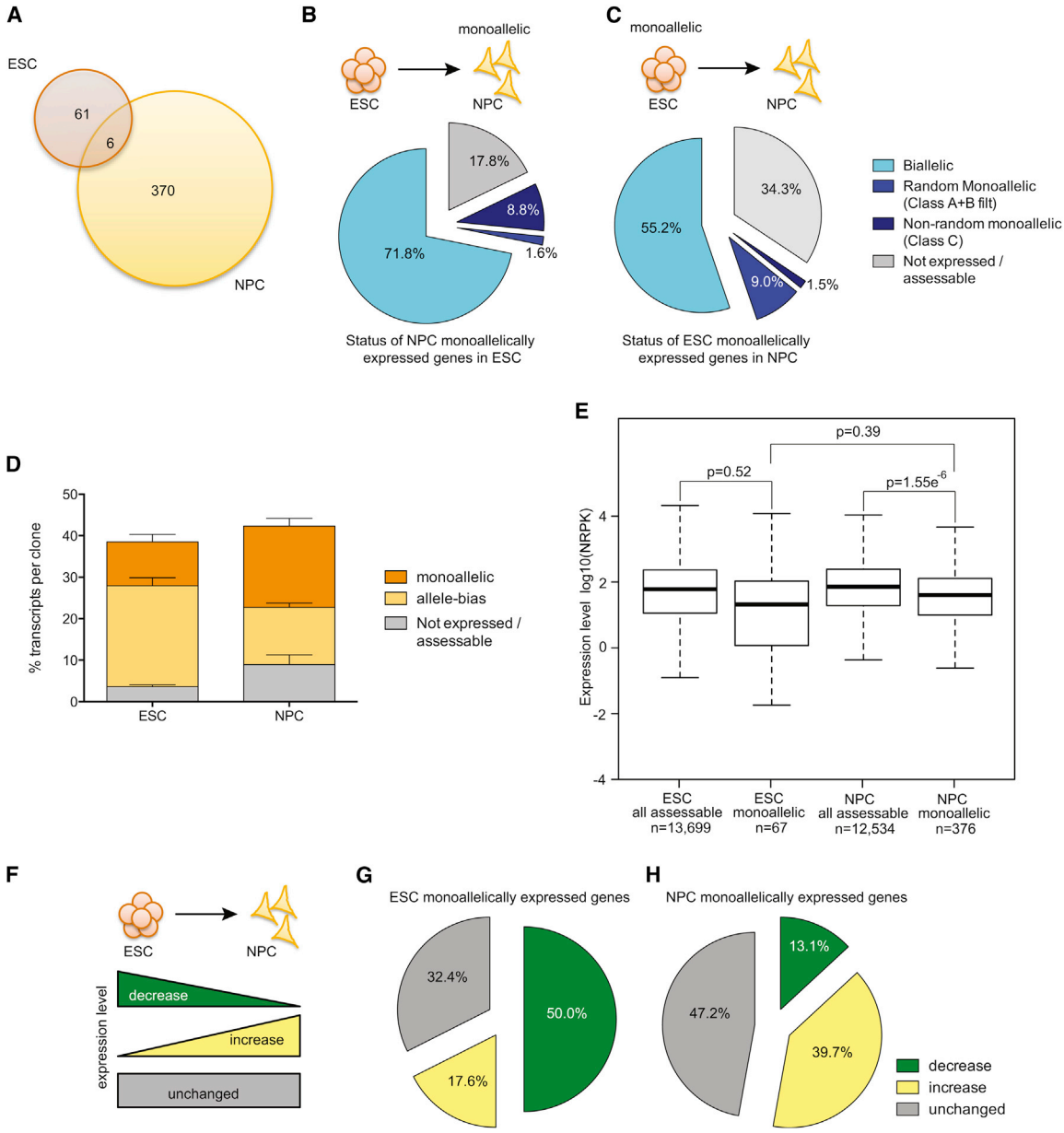


Figure 3. Dynamic Changes in Monoallelic Expression during Differentiation

(A) There is little overlap between ESC and NPC monoallelic transcripts.

(B) Status of the NPC monoallelic transcripts in ESCs.

(C) Status of the ESC monoallelic transcripts in NPCs.

(D) Quantification of ESC and NPC monoallelic transcripts as monoallelic (dark orange), allele biased (light orange), or not expressed/assessable (gray). Bars represent standard deviation among six clones.

(E) Box and whisker plot showing expression-level (normalized RPK) distribution of all assessable and monoallelically expressed genes in both ESCs and NPCs.

(F) Expression levels decrease, increase, or remain unchanged during ESC-to-NPC differentiation.

(G) Expression-level changes for ESC monoallelically expressed genes.

(H) Expression-level changes for NPC monoallelically expressed genes.

See also Figures S3 and S4.

a potential role for DNA methylation, we performed bisulfite analysis of both CpG high and CpG low promoters of ten monoallelically expressed genes, and compared methylation levels between biallelic, allele-biased, and monoallelic NPC clones (Figures 4A, 4B, and S5B). Globally, we did not observe a correlation between the extent of DNA methylation and the overall expression level (Figure S5A). If DNA methylation differentially marked the active and inactive alleles, monoallelic clones should contain a mix of methylated and unmethylated molecules. For seven out of the ten genes tested, we did not see any evidence

for allele-specific DNA methylation in the monoallelic clones (Figures 4A and S5B; data not shown). However, three genes (*Npl*, *Cbr3*, and *Fkbp7*) contained a monoallelic clone in which there was clear separation between methylated and unmethylated reads (Figures 4A and 4B; data not shown). In the case of *Cbr3*, the amplicon used contained an informative SNP that allowed us to assign the bisulfite-treated reads to either the C57Bl/6J or CAST/EiJ allele, confirming the unmethylated and methylated reads were in fact derived from the active and inactive alleles, respectively (Figure 4B). The bisulfite analysis predominantly reflected levels of 5-methylcytosine, as 5-methylcytosine (5meC) and 5-hydroxymethylcytosine (5hmeC) DNA immunoprecipitation revealed 5meC but little to no 5hmeC present at the promoters of six monoallelically expressed genes analyzed (Figures S5C–S5F).

To test whether this distinguishing differential DNA methylation was involved in maintaining monoallelic expression through the cell cycle, we treated the cells with 5-azacytidine, which inhibits DNA methyltransferases, leading to global DNA demethylation. Following 5 days of treatment, which was sufficient to allow several cell divisions to occur, the inactive allele lost all methylation marks (Figure 4B). However, when we analyzed expression from the two alleles by PCR amplification, including an informative SNP within the cDNA, we failed to see reactivation of the inactive allele for *Cbr3* (Figure 4C). This also held true for an additional six genes tested, including *Npl* and *Fkbp7* (data not shown). Thus, DNA methylation alone does not regulate the expression status of random monoallelically expressed genes.

Active and Inactive Alleles Are Differentially Marked by H3K4 and H3K9 Methylation

As we did not find a general role for DNA methylation, we next investigated whether histone modifications may distinguish the active and inactive alleles. We screened promoter regions of monoallelically expressed genes with a panel of nine well-characterized histone modifications implicated in gene transcription or gene silencing (Black et al., 2012) by ChIP (Table S4). Methylation of histone H3 at lysine 4 (H3K4) is associated with actively transcribed regions of the genome (Black et al., 2012). For all gene promoters tested, there was an increase in the levels of both associated H3K4me2 (Figure 5A) and H3K4me3 (Figure 5C) between monoallelic and biallelic clones, consistent with biallelic clones having twice the number of active alleles. Importantly, SNP analysis by Sanger sequencing revealed that only the active allele in monoallelic clones was associated with methylated H3K4 (Figures 5B and 5D), compared to both alleles in biallelic clones, for all genes tested.

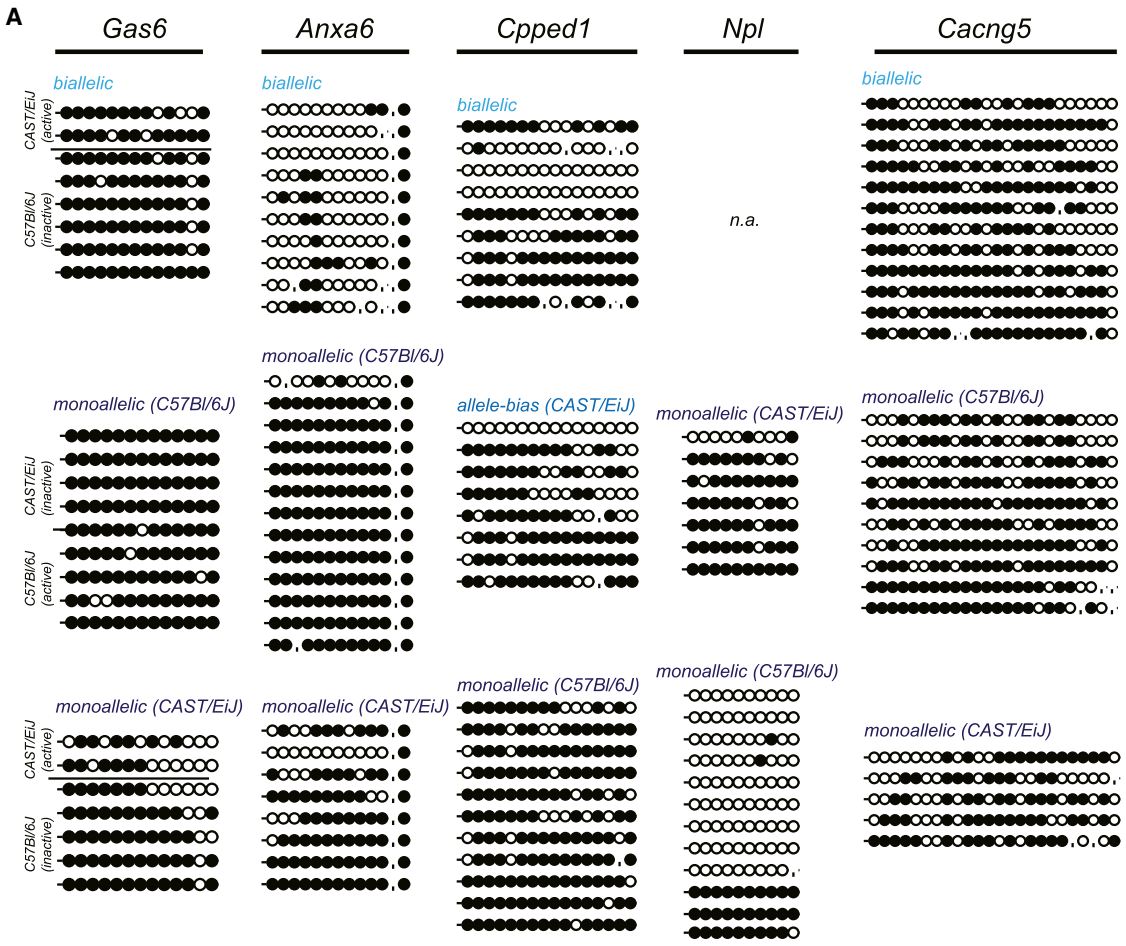
After identifying modifications specifically marking the active allele, we investigated whether there were any modifications associated with the inactive allele. Trimethylation of histone H3 lysine 9 (H3K9me3) and lysine 27 (H3K27me3) are two well-characterized marks of transcriptionally silent genes (Black et al., 2012). ChIP analysis revealed a decrease in the levels of H3K9me3 associated with the promoters of monoallelic versus biallelic clones (Figure 5E). Moreover, Sanger sequencing revealed that H3K9me3 was specifically associated with the inactive allele (Figure 5F). In all cases examined, we did not observe any specific association of H3K27me3 with the inactive allele (Figure S6A), and analysis of published H3K27me3 ChIP-sequencing data sets in NPCs (Mikkelsen et al., 2007) revealed that only 2% of monoallelically expressed genes have detectable H3K27me3 at their promoters (Figure S6B). We also did not observe any preferential association with the inactive allele for other marks of inactive chromatin (Table S4), including monomethylation of histone H4 lysine 20 (H4K20me1). Additionally, trimethylation of histone H4 lysine 20 (H4K20me3), which has been implicated in marking OR gene choice (Magklara et al., 2011), was not observed at these genes (Table S4).

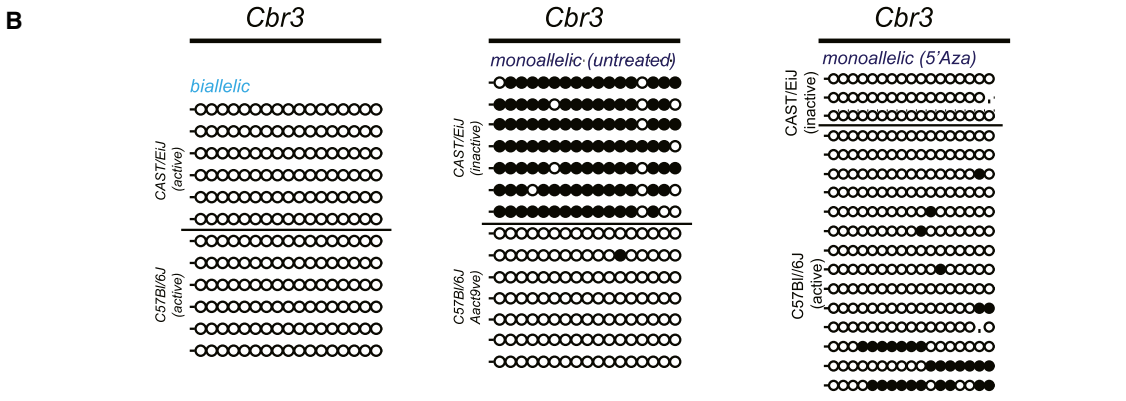
Nuclear Organization of Active versus Inactive Alleles

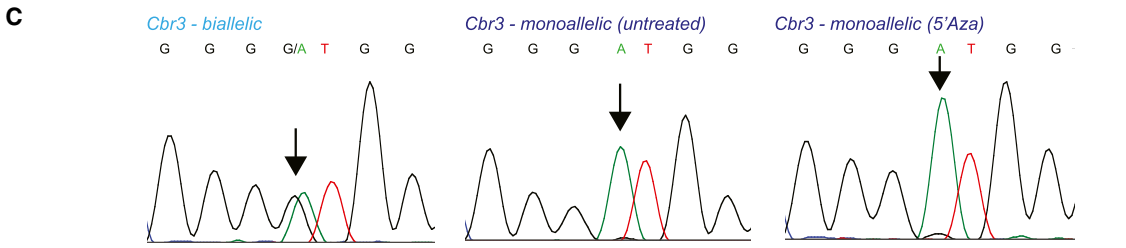
Because nuclear positioning of genes has been correlated with transcriptional activity (Hübner et al., 2013), we were interested in assessing whether differences in nuclear position may distinguish the active and inactive alleles. We performed RNA/DNA-FISH analysis in NPCs and analyzed the position of the active and inactive alleles in three dimensions. However, we did not find any evidence for preferential positioning of the inactive allele toward heterochromatic foci (Figure 6A), nor the nuclear periphery (Figure 6B), as, for most genes examined, both alleles had similar interaction frequencies with these domains despite their difference in transcriptional state. We also did not see evidence for allelic pairing of the active and inactive alleles (data not shown). Furthermore, global analysis of the monoallelically expressed genes did not show any preferential association within or at the borders of lamin-associated domains (LADs) (Figures 6C and 6D) (as defined in Peric-Hupkes et al., 2010). Therefore, the nuclear positioning of these genes does not play a determining role in distinguishing monoallelic expression.

A Subset of Monoallelically Expressed Genes Exhibits Transcriptional Compensation

Next, we examined the impact of monoallelic expression on the transcriptional output in the cell. We performed linear regression analysis to compare expression levels of individual monoallelically expressed genes across the independent NPC clones to determine whether there was a correlation between the extent of allelic imbalance, reflected in the d score, and total expression level. If the levels of expression remained constant across clones, the slope of the linear regression line, *a*, would be equal or close to 0 (Figure 7A, upper). Alternatively, if monoallelic clones had half the transcript level of biallelic clones, the slope *a* would be equal to the y intercept *b* (Figure 7A, lower). Thus, we classified genes as either following the dosage of active alleles $(0.75< -(a/b)<1.25)$ (Figures 7D and 7E) or showing evidence of transcriptional compensation $(|-(a/b)|<0.35)$ (Figures 7B and 7C). Using these criteria, we identified 30 monoallelically expressed genes (8%) with evidence for transcriptional compensation (Table S5) and 54 genes (15.4%) that followed the dosage of active alleles. The remaining genes either showed intermediate responses or were highly variable and so not able to be confidently classified based on data from six clones. We validated the linear regression analysis by quantitative RT-PCR, confirming transcriptional compensation for 7 out of 9 genes (78%) and dosage sensitivity for 7 out of 11 (64%) genes tested (Figures 7B–7E; data not shown). Interestingly, the genes that exhibited transcriptional compensation were enriched for DNA-binding proteins and transcription factor activity (Table S5),







(legend on next page)

although the confidence of enrichment was low (p value = 0.037). The transcriptional compensation is intriguing, as it suggests that for these genes the exact level of transcript is more critical than for others. Furthermore, it supports a model in which the biological consequence of monoallelic expression is not to reduce transcript levels in the cell but rather may be a reflection of the stochastic nature of gene regulation at independent alleles.

DISCUSSION

We performed an allele-specific RNA-sequencing screen and identified a 5.6-fold increase from just 67 to 376 genes in random autosomal monoallelic expression during differentiation of mouse ESCs to NPCs, indicating that monoallelic expression is acquired upon lineage commitment. Importantly, this study provides a detailed and extensive molecular characterization of random monoallelic expression, revealing that histone modifications, not DNA methylation or nuclear organization, distinguish active and inactive alleles. Interestingly, monoallelic expression is not required by the cell, because some clones exhibit biallelic expression, supporting a model in which stochastic gene regulation occurring independently at the two alleles results in monoallelic gene expression, and for some genes is compensated for transcriptionally to maintain the required level of expression of these genes.

We propose that random monoallelic expression exemplifies a stochastic aspect of gene regulation that takes place upon the initiation of specific differentiation programs resulting in global changes in chromatin and gene expression (Figure 7F). If the probability of gene activation or repression is less than 1, this would result in a mixed population of cells containing zero, one, or two active alleles, which, once established and not detrimental to the cell, could be subsequently maintained across cell generations and propagated clonally. Probabilistic models of stochastic gene regulation have been previously proposed for specific examples of monoallelic expression, including *Albumin* in hepatocytes (Michaelson, 1993), *Ly49* receptors in natural killer cells (Held and Kunz, 1998), and interleukins in T lymphocytes (Guo et al., 2005). In all cases, the two alleles are independently regulated with a low activation probability, possibly due to limiting accessibility of key activating factors. One outcome of this independent regulation is that it results in both monoallelic and biallelic cells in a mixed population. Indeed, at least one biallelic clone is observed for almost all monoallelically expressed genes, consistent with an independent stochastic regulation model. The outcome of monoallelic expression for some genes may be unfavorable if the cell requires a specific level of transcript that cannot be accommodated for by the single active allele, thus resulting in cell death. However, for those genes for which the exact level of transcript is not critical or for those that are able to compensate transcriptionally, monoallelic expression represents a viable outcome for the cell.

One key finding of our study is that there is significantly more monoallelic expression in NPCs compared to pluripotent ESCs, supporting the establishment of lineage- or cell-type-specific random monoallelic expression during early development (Figure 7F). Consistent with this, screens performed in neuronal cell types (Jeffries et al., 2012; Li et al., 2012; Wang et al., 2007) have a higher degree of overlap with our study than those in more distant cell types, such as lymphoblasts and fibroblasts (Gimelbrant et al., 2007; Zwemer et al., 2012). Several factors could contribute to the lack of extensive monoallelic expression in ESCs. Not only are ESCs unique in their pluripotent potential and dynamic open chromatin (reviewed in Fisher and Fisher, 2011; Mattout and Meshorer, 2010) but ESC populations are highly heterogeneous both in terms of transcriptional profiles and developmental potency (Huang, 2011; Martinez Arias and Brickman, 2011). Within a colony, ESCs cycle between different states of developmental potential, continuously adjusting their transcriptional program (Canham et al., 2010). Thus, although the initial frequency of monoallelic expression may be similar to that of differentiated cell types, these allelic imbalances may not be maintained as efficiently and thus not clonally propagated.

This study provides an extensive molecular characterization of the differences between active and inactive alleles of random autosomal monoallelically expressed genes. Intriguingly, DNA methylation, important for other examples of monoallelic expression including genomic imprinting (reviewed in Kelsey and Feil, 2013), was not sufficient to distinguish nor maintain monoallelic gene expression of the genes analyzed. Although allele-specific DNA methylation has been previously reported and used to identify random monoallelically expressed genes (Wang et al., 2007), a direct role for DNA methylation driving monoallelic expression has not been shown. Indeed, DNA methylation does not maintain active and inactive alleles of the monoallelically expressed *Cubilin* gene in kidney and intestinal cell lines (Aseem et al., 2013). Additionally, allele-specific DNA methylation does not drive monoallelic expression in human cells in the absence of DNA sequence variation effects (Gutierrez-Arcelus et al., 2013).

Importantly, we did, however, observe that the active and inactive alleles were associated with H3K4me2/3 and H3K9me3, respectively. Histone modifications have been shown to mark other examples of monoallelically expressed genes, including X-inactivated genes by promoter-restricted

Figure 4. DNA Methylation Does Not Regulate Monoallelic Expression

(A) Bisulfite traces of five separate class A gene promoters for NPC clones that were biallelic (top) or allele biased/monoallelic (middle and bottom). Filled and open circles represent methylated (●) and unmethylated CpGs (○), respectively. Each row within a group represents a single bisulfite-treated molecule. For regions containing allele information (*Gas6*), the alleles are separated by a line and labeled accordingly. n.a., not applicable.

(B) Bisulfite traces for *Cbr3* for a biallelic (left), untreated monoallelic (middle), and 5-azacytidine-treated monoallelic (right) clone. A line separates reads from CAST/EiJ (top) and C57Bl/6J alleles (bottom).

(C) Sanger sequencing results of cDNA from a biallelic (left), untreated monoallelic (center), and 5-azacytidine-treated monoallelic (right) clone of *Cbr3*. Both treated and untreated monoallelic samples show only one allele expressed (arrows).

See also Figure S5.

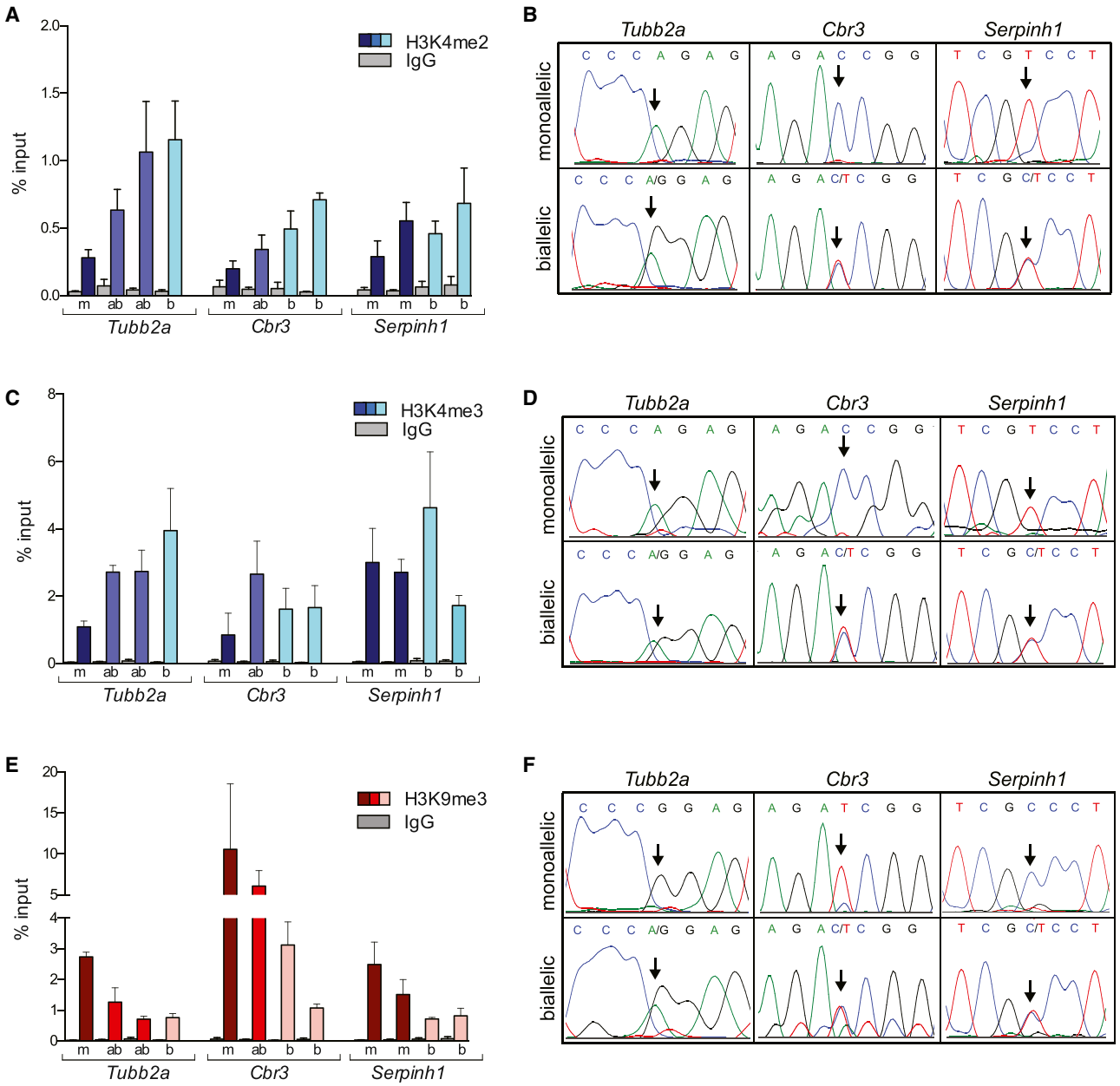


Figure 5. Alleles Are Marked by Differential Histone Modifications

(A–F) ChIP analysis for H3K4me2 (A and B), H3K4me3 (C and D), and H3K9me3 (E and F). Analysis of regions within 200 bp of the transcription start site for two class A (*Tubb2a* and *Cbr3*) and one class B (*Serpinh1*) genes. Pull-down quantification as the percentage of input for H3K4me2 (A), H3K4me3 (C), and H3K9me3 (E) for individual clones that are monoallelic (m; dark blue/red), allele biased (ab; medium blue/red), or biallelic (b; light blue/pink) for the respective clone. IgG (gray) shows nonspecific pull-down. Error bars represent SEM of three or four biological replicates. Sanger sequencing traces of ChIP-quantitative PCR products for H3K4me2 (B), H3K4me3 (D), and H3K9me3 (F) for a monoallelic and biallelic clone for each of the three genes tested showing an allele(s) associated with the respective histone modification. Arrows mark the positions of informative SNPs. See also Figure S6.

H3K4me2 (Rougeulle et al., 2003) and ORs by H3K9me3 and H4K20me3 (Magklara et al., 2011), consistent with our results. Interestingly, we did not see evidence for the Polycomb-associated H3K27me3 repressive mark, although H3K9me3 was present. It remains to be determined whether these histone modifications are actively involved in the inheritance of the transcriptional state or simply reflect the transcriptional status of the respective alleles.

The organization of genes within the nucleus has been linked to transcriptional output (reviewed in Hübner et al., 2013). Whereas nuclear positioning has been implicated in monoallelic expression of ORs (Clowney et al., 2012), immunoglobulins (Skok et al., 2001), and *Gfap* in astrocytes (Takizawa et al., 2008), we did not observe any differences in the position of active versus inactive alleles for the genes examined in this study.

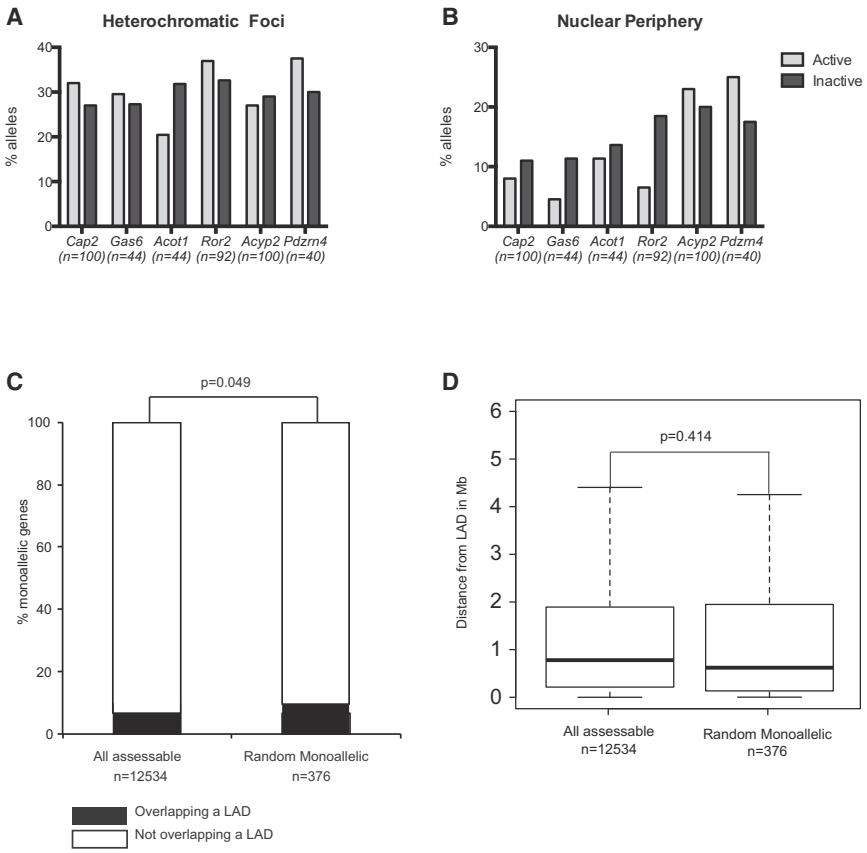


Figure 6. Active and Inactive Alleles Do Not Show Preferential Nuclear Positioning

(A and B) Bar graphs showing the proportion of active (light gray) and inactive (dark gray) alleles associated with either heterochromatic foci (A) or the nuclear periphery (B) for six separate class A monoallelically expressed genes in NPCs. Measurements were performed in three dimensions.

(C) Proportion of monoallelically expressed genes that are located within (black) or outside (white) LADs.

(D) Box and whisker plot showing minimal distance of genes from the nearest LAD in Mb.

There may be yet additional undetected characteristics that distinguish the active and inactive alleles of monoallelically expressed genes that may also play a role in maintaining the difference in transcriptional state across cell divisions. Asynchronous DNA replication timing, in which the active allele replicates earlier in S phase than the inactive allele (Hiratani and Gilbert, 2009), has been observed at monoallelically expressed genes (Donley et al., 2013; Dutta et al., 2009). However, as the monoallelically expressed genes are interspersed among biallelic genes within the same DNA replication timing domains (Alabert and Groth, 2012), it is unlikely that asynchronous DNA replication timing contributes to the monoallelic state.

Surprisingly, transcript levels for some monoallelically expressed genes did not follow the active allele dosage. Transcriptional compensation has been reported for heterozygous knockout mice that show comparable mRNA and/or protein levels to their wild-type counterparts, including *Mks1* (Wheway et al., 2013) and *Bag3* (Homma et al., 2006), both identified as monoallelically expressed in NPCs. However, there are examples of genes in which the heterozygous mice have reduced transcript levels, including the monoallelically expressed genes *Cth* (Kaasik et al., 2007) and *Cstb* (Ishii et al., 2010), suggesting that transcriptional upregulation is not only gene specific but also cell-type specific. Transcriptional compensation has also been observed in nonmammalian systems, including *Drosophila* (McAnally and Yampolsky, 2010) and maize (Guo and Birchler, 1994), in which mRNA levels do not strictly follow the dosage of the gene. For those genes that exhibit compensation, it will be of interest to determine the mechanisms by which transcriptional compensation maintains the total level of mRNA in the cell, potentially through levels and accessibility of specific transcription factors, feedback loops sensing the levels of mRNA and/or protein in the cell (Guidi et al., 2004), or autoregulation (Trieu et al., 2003). This ability to tune the transcriptional output of an allele in response to either genetic or epigenetic inactivation of the second allele has important biological consequences, especially in the interpretation of copy-number variants, as these may not necessarily result in a change of transcript and protein product. In this way, random autosomal monoallelic gene expression illustrates the remarkable plasticity and stochasticity of gene regulation in mammalian cells.

EXPERIMENTAL PROCEDURES

Cell Culture

ESCs were cultured using standard procedures in medium containing 1,000 U/ml leukemia inhibitory factor (Millipore) with irradiated mouse embryonic fibroblast (MEF) feeders (GlobalStem) on gelatin-coated plates. ESCs were removed from feeder cells by soaking twice on gelatin-coated plates for 1 hr each prior to sample collection or cell differentiation. Differentiation was performed using a protocol adapted from Conti et al. (2005) by culturing ESCs in the absence of MEF feeders in 50:50 DMEM/F12:Neurobasal medium (GIBCO) supplemented with 1× N2 (GIBCO), 1× B27 (GIBCO), 40 mg/l insulin (Sigma), 25 μg/ml BSA fraction V (GIBCO) at 0.5 × 10^6 to 2.0 × 10^6 cells per 10 cm plate for 6 days. Cells were then resuspended in N2 expansion medium (DMEM/F12, 50 μg/ml BSA fraction V, 10 ng/ml epidermal growth factor [PeproTech], 10 ng/ml fibroblast growth factor [PeproTech], 1 μg/ml laminin [Invitrogen], 1× N2) and plated onto uncoated T75 flasks to allow for neurosphere outgrowth. Following 4 days, neurospheres were collected by mild centrifugation and plated onto gelatin-coated plates in N2 expansion medium. Following two or three passages, cells represented a homogeneous population of NPCs. Single ESCs and NPCs were seeded through limiting dilutions in 96 wells and expanded to obtain clonal populations.

Allele-Specific RNA-Sequencing Screen

RNA from six independent single-cell clones was isolated using TRIzol reagent (Ambion) and polyA$^+$ RNA was isolated (Oligotex kit; QIAGEN). Stranded libraries were prepared using a protocol adapted from Parkhomchuk et al. (2009) for paired-end sequencing on the Illumina GA IIx platform. Reads were mapped with BWA to both C57Bl/6J and CAST/EiJ transcriptomes. Affymetrix Mouse Diversity SNP arrays (The Jackson Laboratory) were performed

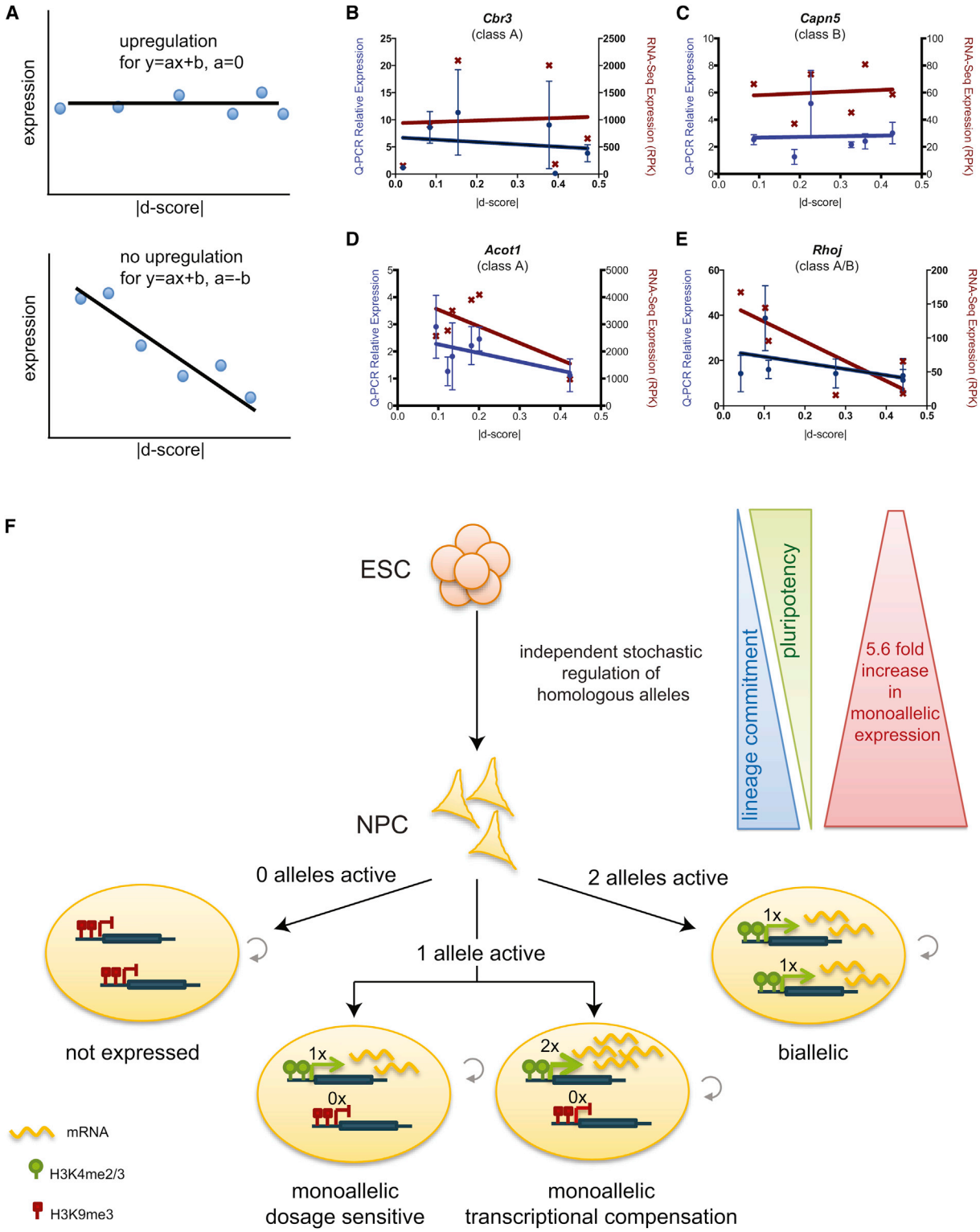


Figure 7. A Subset of Monoallelically Expressed Genes Upregulates the Single Active Allele

(A) A hypothetical gene that undergoes compensation (top) or no compensation (bottom). For the trend line y = ax + b, the noncompensated genes would be expected to have a = −b or a/−b = 1. Compensated genes would have a = 0 or a/−b < 1.

(B–E) Examples of genes classified as compensated (a/−b < 0.35) (B and C) or noncompensated (a/−b > 0.75) (D and E). Red crosses represent individual clone normalized RPK from RNA sequencing; blue circles represent mean normalized expression of three biological replicates by quantitative PCR. Errors bars represent standard deviation from the mean. Lines represent linear regression lines of best fit. See also Table S5.

(legend continued on next page)

for DNA from NPC clones to control for locus heterozygosity. For detailed statistical analysis, see the Supplemental Experimental Procedures. For each assessable transcript (≥5 reads per SNP) in each clone, a significance test and d score representing the weighted difference from the 50:50 expected ratio between the two alleles were calculated. Transcripts were then classified as monoallelic (|d score| ≥ 0.40, p value ≤ 10^{-8}), allele biased (0.18 ≤ |d score| < 0.40, p value ≤ 10^{-8}), biallelic (|d score| < 0.18 and/or p value > 10^{-8}), not expressed (expression is lower than 5.8 normalized reads per kb [NRPK]), nonassessable, or other, and assigned to three classes of monoallelically expressed transcripts based on patterns across clones, or a fourth class containing all other transcripts: class A transcripts have at least one C57Bl/6J and one CAST/EiJ biased clone; class B transcripts have at least one biased clone and one biallelic clone and were further filtered to obtain additional monoallelic transcripts; and class C transcripts showed bias in all clones but only toward one allele.

Validation of Monoallelically Expressed Genes

RNA was isolated using TRIzol reagent (Ambion) and converted to cDNA (Applied Biosciences; RT reagents). PCR amplification of exonic SNPs was performed (Phusion High-Fidelity Polymerase; New England BioLabs), and products were gel purified and subjected to Sanger sequencing. Sequencing traces were analyzed using 4Peaks 1.7.2 (Mekentosj) and scored independently of screen results. Nonclonal NPC cDNA or genomic DNA was used to confirm the presence of the SNP. See Table S3 for primer sequences and SNP information.

RNA and DNA Fluorescence In Situ Hybridization

Probes for both RNA- and DNA-FISH were generated from bacterial artificial chromosome or fosmid DNA (*Acot1* WI1-2795P12; *Acyp2* RP23-405O2; *Arap1* WI1-0101I18; *Atp1a2* WI1-1389O12; *Cap2* RP23-105K14; *Gas6* WI1-0153N19; *Mavs* WI1-1832A20; *Pdzrn4* RP23-322L18; *Ror2* RP23-280O5) by nick translation (Abbott Molecular) with red or green fluorescently conjugated dUTP nucleotides (Enzo Life Sciences) for 10 hr at 15°C. Probe size was verified by agarose gel electrophoresis to be 50–400 nt. Probes were mixed with competitor DNA, lyophilized, and resuspended in 50% deionized formamide, 2× SSC, 10% dextran sulfate.

RNA/DNA-FISH was performed sequentially, with separate images taken for both RNA-FISH and DNA-FISH. Cells were grown on coverslips, fixed in freshly prepared 4% formaldehyde for 15 min, and permeabilized in 0.1% Triton X-100 for 5 min on ice in the presence of 5 mM vanadyl ribonucleoside complex (New England BioLabs). RNA-FISH was performed by hybridizing prepared denatured probes on coverslips overnight at 40°C. DNA-FISH required prior treatment with 0.1 mg/ml RNase A (Invitrogen) for 1 hr at 37°C, followed by heat denaturation in 70% formamide, 2× SSC for 5 min at 80°C, and hybridization with denatured probe overnight at 37°C. Following hybridization, cells were washed in 2× SSC/50% formamide, 2× SSC, and then 1× SSC at 37°C, counterstained with DAPI and mounted in antifade containing 10% glycerol and 1 mg/ml *p*-phenylenediamine (Sigma).

Samples were imaged using an Applied Precision DeltaVision Core wide-field fluorescence microscope system (GE Healthcare) equipped with a PlanApo 60× 1.40 numerical aperture objective lens (Olympus America). Image stacks were taken at 0.2 nm intervals throughout the entire cell and deconvolved using Applied Precision softWoRx software version 4.2.1 with default parameters. Separate projections of RNA-FISH/DAPI and DNA-FISH/DAPI images were overlaid in Photoshop using heterochromatin foci as a guide. Image analysis was performed manually using Applied Precision softWoRx software.

Quantitative RT-PCR

cDNA was prepared as above and used for quantitative RT-PCR using SYBR Green reagents (Applied Biosciences). Quantitative PCR was performed using the following forward and reverse primer sequences: *18S* 5′-GGGCCC GAAGCGTTTACTTT-3′, 5′-CGCCGGTCCAAGAATTTCAC-3′; *Acot1* 5′-CAT CACCTTTGGAGGGGAGC-3′, 5′-TGTACCTTTCCCCAACCTCC-3′; *Capn5* 5′-ACACGTCAGAGGAATGGCAG-3′, 5′-GGATGCTCAGGTAGGACGTG-3′; *Cbr3* 5′-GTCCCTCTGACATGTCGTCC-3′, 5′-CGTTAAGTCCCCCGTACTCC-3′; *CycloB* 5′-GACAGACAGCCGGGACAAGC-3′, 5′-GGGGATTGACAGGA CCCACA-3′; *Gapdh1* 5′-GGTGGTGAAGCAGGCATCTG-3′, 5′-CGGCATC GAAGGTGGAAGAG-3′; *Rhoj* 5′-GGCCACTCTCTTACCCCAAC-3′, 5′-GAGG CATGCAGTCCTTCAGT-3′. Three biological replicates for each sample were used for each experiment, and values were normalized to the geometric mean of at least three separate housekeeping genes.

Bisulfite Analysis

Five hundred nanograms of purified DNA (TRIzol reagent; Ambion) was converted with bisulfite (EZ DNA Methylation-Gold kit; Zymo) according to the manufacturer's instructions. Primers amplifying A-T SNPs within promoter CpG islands were designed using MethPrimer (Li and Dahiya, 2002). PCR was performed using OneTaq Hot Start DNA Polymerase (New England BioLabs). See Table S5 for primers. Products were gel purified and cloned (Topo-TA; Invitrogen). Clones were sent for Sanger sequencing and analyzed using BiQ Analyzer (Bock et al., 2005).

Chromatin Immunoprecipitation

Chromatin was prepared from cells, and immunoprecipitations were performed as described in the Supplemental Experimental Procedures. Quantitative RT-PCR was performed on immunoprecipitated DNA using primers amplifying promoter or genic regions of monoallelically expressed genes (see Table S4) and normalized to input DNA. Products containing informative SNPs were subsequently analyzed by Sanger sequencing. At least three biological replicates were analyzed per sample.

ACCESSION NUMBERS

RNA-sequencing and microarray data have been deposited in the ArrayExpress database under accession numbers E-MTAB-1822 and E-MTAB-1823, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.devcel.2014.01.017.

AUTHOR CONTRIBUTIONS

M.A.E.-M. and D.L.S. conceived the study, designed the experiments, and wrote the manuscript. M.A.E.-M. and J.H.B. performed the experiments. D.T., J.C.M., and P.F. designed and performed all bioinformatic analysis.

ACKNOWLEDGMENTS

We thank C. Vakoc (Cold Spring Harbor Laboratory; CSHL) for kindly providing the C57Bl/6J × CAST/EiJ ESC cell line; A. Mills (CSHL) for the AB2.2 ESC line; C. Davis, J. Drenkow, and T. Gingeras (CSHL) for assistance in RNA-sequencing library preparation; E. Hodges and G. Hannon (CSHL) for assistance in DNA methylation studies; S. Hearn for assistance in microscopy; M.J. Delás for assistance with DNA-FISH; and members of the Spector laboratory for discussion and comments. M.A.E.-M. is supported by a Genentech Foundation Fellowship and George A. and Marjorie H. Anderson Fellowship.

(F) Monoallelic gene expression increases during differentiation of ESCs to NPCs, coinciding with the loss of pluripotency and gain of lineage commitment. Stochastic regulation of homologous alleles results in random monoallelic gene expression. Upon differentiation of ESCs to NPCs, there is independent low-probability regulation of the two alleles, resulting in a mixed population of NPCs containing zero, one, or two active alleles. The active and inactive alleles are distinguished through H3K4me2/3 (green circles) and H3K9me3 (red squares), respectively, which may contribute to the clonal inheritance of allelic imbalance. Monoallelic expression can result in either dosage sensitivity, where the cell has half the levels of mRNA (yellow line) as biallelic cells, or transcriptional compensation, in which the cell upregulates the single active allele.

D.T., J.C.M., and P.F. are supported by the European Molecular Biology Laboratory. D.T. and P.F. are supported by the Wellcome Trust (WT095908). J.H.B. is supported by a Deutscher Akademischer Austauschdienst Postdoctoral Fellowship. Research in the Spector laboratory is supported by National Institute of General Medical Sciences grant 42694 and National Cancer Institute grant 2P30CA45508.

Received: August 9, 2013
Revised: December 21, 2013
Accepted: January 21, 2014
Published: February 24, 2014

REFERENCES

Alabert, C., and Groth, A. (2012). Chromatin replication and epigenome maintenance. Nat. Rev. Mol. Cell Biol. *13*, 153–167.

Aseem, O., Barth, J.L., Klatt, S.C., Smith, B.T., and Argraves, W.S. (2013). Cubilin expression is monoallelic and epigenetically augmented via PPARs. BMC Genomics *14*, 405–423.

Bartolomei, M.S., and Ferguson-Smith, A.C. (2011). Mammalian genomic imprinting. Cold Spring Harb. Perspect. Biol. *3*, a002592.

Black, J.C., Van Rechem, C., and Whetstine, J.R. (2012). Histone lysine methylation dynamics: establishment, regulation, and biological impact. Mol. Cell *48*, 491–507.

Bock, C., Reither, S., Mikeska, T., Paulsen, M., Walter, J., and Lengauer, T. (2005). BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. Bioinformatics *21*, 4067–4068.

Canham, M.A., Sharov, A.A., Ko, M.S.H., and Brickman, J.M. (2010). Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. PLoS Biol. *8*, e1000379.

Chess, A. (2012). Mechanisms and consequences of widespread random monoallelic expression. Nat. Rev. Genet. *13*, 421–428.

Chess, A. (2013). Random and non-random monoallelic expression. Neuropsychopharmacology *38*, 55–61.

Chess, A., Simon, I., Cedar, H., and Axel, R. (1994). Allelic inactivation regulates olfactory receptor gene expression. Cell *78*, 823–834.

Clowney, E.J., Magklara, A., Colquitt, B.M., Pathak, N., Lane, R.P., and Lomvardas, S. (2011). High-throughput mapping of the promoters of the mouse olfactory receptor genes reveals a new type of mammalian promoter and provides insight into olfactory receptor gene regulation. Genome Res. *21*, 1249–1259.

Clowney, E.J., LeGros, M.A., Mosley, C.P., Clowney, F.G., Markenskoff-Papadimitriou, E.C., Myllys, M., Barnea, G., Larabell, C.A., and Lomvardas, S. (2012). Nuclear aggregation of olfactory receptor genes governs their monogenic expression. Cell *151*, 724–737.

Conti, L., Pollard, S.M., Gorba, T., Reitano, E., Toselli, M., Biella, G., Sun, Y., Sanzone, S., Ying, Q.-L., Cattaneo, E., and Smith, A. (2005). Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. PLoS Biol. *3*, e283.

Donley, N., Stoffregen, E.P., Smith, L., Montagna, C., and Thayer, M.J. (2013). Asynchronous replication, mono-allelic expression, and long range *cis*-effects of ASAR6. PLoS Genet. *9*, e1003423.

Dutta, D., Ensminger, A.W., Zucker, J.P., and Chess, A. (2009). Asynchronous replication and autosome-pair non-equivalence in human embryonic stem cells. PLoS ONE *4*, e4970.

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

Fisher, C.L., and Fisher, A.G. (2011). Chromatin states in pluripotent, differentiated, and reprogrammed cells. Curr. Opin. Genet. Dev. *21*, 140–146.

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

Guidi, C.J., Veal, T.M., Jones, S.N., and Imbalzano, A.N. (2004). Transcriptional compensation for loss of an allele of the Ini1 tumor suppressor. J. Biol. Chem. *279*, 4180–4185.

Guo, M., and Birchler, J.A. (1994). Trans-acting dosage effects on the expression of model gene systems in maize aneuploids. Science *266*, 1999–2002.

Guo, L., Hu-Li, J., and Paul, W.E. (2005). Probabilistic regulation in TH2 cells accounts for monoallelic expression of IL-4 and IL-13. Immunity *23*, 89–99.

Gutierrez-Arcelus, M., Lappalainen, T., Montgomery, S.B., Buil, A., Ongen, H., Yurovsky, A., Bryois, J., Giger, T., Romano, L., Planchon, A., et al. (2013). Passive and active DNA methylation and the interplay with genetic variation in gene regulation. Elife *2*, e00523.

Held, W., and Kunz, B. (1998). An allele-specific, stochastic gene expression process controls the expression of multiple Ly49 family genes and generates a diverse, MHC-specific NK cell receptor repertoire. Eur. J. Immunol. *28*, 2407–2416.

Hiratani, I., and Gilbert, D.M. (2009). Replication timing as an epigenetic mark. Epigenetics *4*, 93–97.

Homma, S., Iwasaki, M., Shelton, G.D., Engvall, E., Reed, J.C., and Takayama, S. (2006). BAG3 deficiency results in fulminant myopathy and early lethality. Am. J. Pathol. *169*, 761–773.

Huang, S. (2011). Systems biology of stem cells: three useful perspectives to help overcome the paradigm of linear pathways. Philos. Trans. R. Soc. Lond. B Biol. Sci. *366*, 2247–2259.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Hübner, M.R., Eckersley-Maslin, M.A., and Spector, D.L. (2013). Chromatin organization and transcriptional regulation. Curr. Opin. Genet. Dev. *23*, 89–95.

Ishii, I., Akahoshi, N., Yamada, H., Nakano, S., Izumi, T., and Suematsu, M. (2010). Cystathionine γ-lyase-deficient mice require dietary cysteine to protect against acute lethal myopathy and oxidative injury. J. Biol. Chem. *285*, 26358–26368.

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

Kaasik, A., Kuum, M., Aonurm, A., Kalda, A., Vaarmann, A., and Zharkovsky, A. (2007). Seizures, ataxia, and neuronal loss in cystatin B heterozygous mice. Epilepsia *48*, 752–757.

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

Kelsey, G., and Feil, R. (2013). New insights into establishment and maintenance of DNA methylation imprints in mammals. Philos. Trans. R. Soc. Lond. B Biol. Sci. *368*, 20110336.

Li, L.-C., and Dahiya, R. (2002). MethPrimer: designing primers for methylation PCRs. Bioinformatics *18*, 1427–1431.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589–595.

Li, S.M., Valo, Z., Wang, J., Gao, H., Bowers, C.W., and Singer-Sam, J. (2012). Transcriptome-wide survey of mouse CNS-derived cells reveals monoallelic expression within novel gene families. PLoS ONE *7*, e31751.

Magklara, A., Yen, A., Colquitt, B.M., Clowney, E.J., Allen, W., Markenscoff-Papadimitriou, E., Evans, Z.A., Kheradpour, P., Mountoufaris, G., Carey, C., et al. (2011). An epigenetic signature for monoallelic olfactory receptor expression. Cell *145*, 555–570.

Martinez Arias, A., and Brickman, J.M. (2011). Gene expression heterogeneities in embryonic stem cell populations: origin and function. Curr. Opin. Cell Biol. *23*, 650–656.

Mattout, A., and Meshorer, E. (2010). Chromatin plasticity and genome organization in pluripotent embryonic stem cells. Curr. Opin. Cell Biol. *22*, 334–341.

McAnally, A.A., and Yampolsky, L.Y. (2010). Widespread transcriptional autosomal dosage compensation in *Drosophila* correlates with gene expression level. Genome Biol. Evol. *2*, 44–52.

Michaelson, J. (1993). Cellular selection in the genesis of multicellular organization. Lab. Invest. *69*, 136–151.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–560.

Parkhomchuk, D., Borodina, T., Amstislavskiy, V., Banaru, M., Hallen, L., Krobitsch, S., Lehrach, H., and Soldatov, A. (2009). Transcriptome analysis by strand-specific sequencing of complementary DNA. Nucleic Acids Res. *37*, e123.

Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S.W.M., Solovei, I., Brugman, W., Gräf, S., Flicek, P., Kerkhoven, R.M., van Lohuizen, M., et al. (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. Mol. Cell *38*, 603–613.

Pernis, B., Chiappino, G., Kelus, A.S., and Gell, P.G. (1965). Cellular localization of immunoglobulins with different allotypic specificities in rabbit lymphoid tissues. J. Exp. Med. *122*, 853–876.

Rougeulle, C., Navarro, P., and Avner, P. (2003). Promoter-restricted H3 Lys 4 di-methylation is an epigenetic mark for monoallelic expression. Hum. Mol. Genet. *12*, 3343–3348.

Schulz, E.G., and Heard, E. (2013). Role and control of X chromosome dosage in mammalian development. Curr. Opin. Genet. Dev. *23*, 109–115.

Skok, J.A., Brown, K.E., Azuara, V., Caparros, M.L., Baxter, J., Takacs, K., Dillon, N., Gray, D., Perry, R.P., Merkenschlager, M., and Fisher, A.G. (2001). Nonequivalent nuclear location of immunoglobulin alleles in B lymphocytes. Nat. Immunol. *2*, 848–854.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nat. Rev. Genet. *14*, 204–220.

Takizawa, T., Gudla, P.R., Guo, L., Lockett, S., and Misteli, T. (2008). Allele-specific nuclear positioning of the monoallelically expressed astrocyte marker GFAP. Genes Dev. *22*, 489–498.

Trieu, M., Ma, A., Eng, S.R., Fedtsova, N., and Turner, E.E. (2003). Direct autoregulation and gene dosage compensation by POU-domain transcription factor Brn3a. Development *130*, 111–121.

Wang, J., Valo, Z., Smith, D., and Singer-Sam, J. (2007). Monoallelic expression of multiple genes in the CNS. PLoS ONE *2*, e1293.

Wheway, G., Abdelhamed, Z., Natarajan, S., Toomes, C., Inglehearn, C., and Johnson, C.A. (2013). Aberrant Wnt signalling and cellular over-proliferation in a novel mouse model of Meckel-Gruber syndrome. Dev. Biol. *377*, 55–66.

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