Trends in Genetics

CellPress
REVIEWS

Special Issue: Gene Expression in Time and Space
Review

# Parallels between Mammalian Mechanisms of Monoallelic Gene Expression

Ahmed Amine Khamlichi[1,*] and Robert Feil[2,*]

**Different types of monoallelic gene expression are present in mammals, some of which are highly flexible, whereas others are more rigid. These include allelic exclusion at antigen receptor loci, the expression of olfactory receptor genes, genomic imprinting, X-chromosome inactivation, and random monoallelic expression (MAE). Although these processes play diverse biological roles, and arose through different selective pressures, the underlying epigenetic mechanisms show striking resemblances. Regulatory transcriptional events are important in all systems, particularly in the specification of MAE. Combined with comparative studies between species, this suggests that the different MAE systems found in mammals may have evolved from analogous ancestral processes.**

## Highlights

MAE systems have diverse biological roles, but show similarities in the underlying epigenetic mechanisms.

Transcriptional events and non-coding RNAs are associated with the initiation of MAE.

Differential modification and architectures of chromatin are also important in MAE.

The different MAE systems found in mammals may have evolved from analogous ancestral processes.

## Introduction

In mammals, many genes show **monoallelic expression (MAE**, see Glossary), often in a lineage- or tissue-specific manner (Table 1 and Box 1). The first known example concerned **antigen receptor** *(AgR)* loci, which undergo a genetic recombination process called V(D)J recombination (Box 2). Each developing B or T lymphocyte generates one functionally expressed gene, on only one of the two parental chromosomes [1,2]. Analogously, in olfactory neurons, **olfactory receptor** *(OR)* genes are regulated such that only one of many genes is randomly 'chosen' to become activated, and on only one allele [3]. The MAE of *OR* genes, however, is orchestrated entirely by epigenetic processes [4]. MAE in neurons at protocadherin gene clusters shows similarities to *OR* expression [5]. Many individual genes also display **random** MAE (RMAE), which further contributes to the diversification of cellular identity [6], and provides functional specificity. X-chromosome inactivation serves a somewhat different biological purpose. This epigenetic process stably inactivates genes on one of the two X chromosomes in female embryos, and thus compensates gene dosage between females (which have two X chromosomes) and males (one X) [7,8]. In another epigenetic dosage mechanism, called genomic imprinting, autosomal genes become expressed in a monoallelic manner. The allelic expression of imprinted genes is different from other MAE systems in that it is non-random and is determined by the parental origin of the allele [9,10].

Although the MAE systems are functionally different (Table 1), they display similar regulatory mechanisms. Chromatin modifications differentially mark active and repressed alleles at one or multiple steps of the process. There is growing evidence for the involvement of transcriptional events as well. Differentially expressed non-coding RNAs (ncRNAs) play diverse roles in MAE, often through chromatin regulatory effects on close-by genes. Another level of convergence concerns the higher-order architecture of chromatin domains, which also contributes to MAE. We will not discuss this topic in depth because excellent reviews covering this topic are available [4,11–14]. We introduce MAE systems in mammals and highlight similarities between

[1]Institute of Pharmacology and Structural Biology (IPBS), Centre National de la Recherche Scientifique (CNRS) and Paul Sabatier University (UPS), 205 route de Narbonne, 31077 Toulouse, France
[2]Institute of Molecular Genetics of Montpellier (IGMM), CNRS and the University of Montpellier, 1919 route de Mende, 34293 Montpellier, France

*Correspondence:
robert.feil@igmm.cnrs.fr (R. Feil) and
ahmed.khamlichi@ipbs.fr
(A.A. Khamlichi) .

the mechanisms by which MAE is established and maintained. Instead of providing comprehensive descriptions, we present examples of convergence between different types of MAE. The ostensibly diverse mechanisms may have evolved from comparable ancestral processes, with gradual acquisition of more stable and diversified MAE, for instance through adaptation of DNA methylation or by inclusion of genetic recombination as part of the process.

## Monoallelic Systems Evolved through Different Biological Constraints

X-inactivation compensates for gene dosage between males and females, and controls genes across the X chromosome, except at the **pseudo-autosomal regions** (PARs) that are, present on both the X and the Y chromosome. In mice and humans, random X-inactivation is initiated in the early embryo through upregulation of the long ncRNA (lncRNA) *Xist* (X-inactive specific transcript) from one of the two alleles in each cell [7]. This triggers diverse repressive events on the chosen X chromosome, including exclusion of RNA polymerase II, installation of repressive histone modifications, and acquisition of DNA methylation at promoters [7,8].

Autosomal genes controlled by genomic imprinting play essential roles in development, metabolism, and behavior. As opposed to other MAE systems, their monoallelic expression is deterministic [9]. In the embryo, some imprinted genes are expressed from the maternal, oocyte-inherited, allele only, whereas others are expressed only from the paternal, sperm-inherited, allele, with no distinction between males and females [9]. Imprinted genes are clustered and fundamental to their allelic expression are DNA methylation marks at **imprinting control regions** (ICRs) that originate from either the sperm or the oocyte [15,16]. It is this differential methylation at ICRs that brings about imprinted gene expression in the embryo at many domains by mediating allelic expression of regulatory lncRNAs [15].

Recent mouse studies have revealed a mechanistically different type of genomic imprinting that is controlled by repressive histone H3 lysine 27 trimethylation (H3K27me3) acquired in the oocyte [17]. At multiple genes, maternal H3K27me3 imprints bring about a transient, partial repression of the maternal allele in the preimplantation embryo [18]. Interestingly, oocyte-acquired H3K27me3 also represses the maternally inherited allele of *Xist* in post-fertilization female mouse embryos [19], and this triggers imprinted X-inactivation, which we discuss below.

Adaptive immune responses in vertebrates rely on the ability of B and T lymphocytes to express a large and diverse repertoire of AgRs that can recognize countless antigens. In jawed vertebrates, AgR diversity is generated by V(D)J recombination that is catalyzed by the recombinase complex RAG1/RAG2 (hereafter called 'RAG'), which recognizes conserved 'recombination signal sequences' (RSSs) that flank the recombining segments of *AgR* loci [2]. Although **B cell receptor** (*BCR*) and **T cell receptor** (*TCR*) loci share similar RSSs and use the same recombination machinery, full V(D)J recombination at *BCR* loci occurs in B cells only, and that of *TCR* loci in T cells specifically, leading to monospecific B and T cells respectively [1].

In the same way as the humoral immune system recognizes antigens that vary in size and structure, the olfactory system can discern subtle differences among simple or complex mixtures of odorants. This discriminatory power is mediated by a vast repertoire of cell-surface G protein-coupled chemosensory receptors. The ORs comprise the largest family [4]. Each olfactory sensory neuron achieves selectivity by restricting expression to a single allele of one *OR* gene, and by keeping silent the second allele as well as all other *OR* genes, whether present on the same or a different chromosome. This ensures that each neuron will respond only to specific odorants (and will converge to specific glomeruli in the olfactory bulb).

## Glossary

**Allelic exclusion:** ensures MAE at olfactory and antigen receptor loci and the functional identity of sensory neurons or lymphocytes.
**Antigen receptor (AgR):** molecules that can bind antigens and are expressed on the surface of B lymphocytes (immunoglobulin or B cell receptor, BCR) or T lymphocytes (T cell receptor, TCR).
**B cell receptors (BCRs):** molecules that can bind antigens and are expressed on the surface of B lymphocytes (also called surface immunoglobulins).
**CpG islands:** region up to several kilobases in length that are rich in CpG dinucleotides. Most CpG islands are unmethylated and comprise promoters.
**Differentially methylated region (DMR):** a region at which DNA methylation marks one of the parental alleles only.
**Enhancer of Zeste homolog-2 (EZH2):** a methyltransferase of histone H3 lysine 27 (H3K27) that is part of the PRC2 complex.
**Imprinting control region (ICR):** a CpG island (-like) sequence with a germline-derived epigenetic mark that controls the parental allele-specific expression of nearby genes.
**Monoallelic expression (MAE):** expression uniquely, or predominantly, from one of the two gene copies (alleles). MAE can be random or deterministic.
**Olfactory receptors (ORs):** G protein-coupled chemosensory receptors expressed on the surface of olfactory neurons.
**Polycomb repressive complexes (PRC1 and PRC2):** complexes composed of Polycomb group proteins that mediate repressive histone modifications and gene silencing.
**Pseudo-autosomal region (PAR):** region(s) that are homologous between the X and Y chromosomes, and undergo recombination between X and Y during meiosis.
**Random monoallelic expression (RMAE):** a term often used to indicate transient random MAE at individual genes in stem and differentiated cells.
**Recombination centers (RCs):** encompass approximately the 'joining' (J) region of antigen receptor

The rapidly expanding repertoire of genes that display RMAE include those encoding interleukins, natural killer cell receptors, pheromone receptors, protocadherins, and many single-copy genes [5,6,20]. Genes with RMAE often encode cell-surface proteins. It has therefore been proposed that MAE may have evolved to provide a unique identity to individual cells, and specificity to cellular functions [6]. Concordantly, random MAE gene alleles display high genetic diversity within human populations [21].

## Mechanistic Models of MAE

With the exception of autosomal imprinted genes and imprinted X-inactivation, where allelic expression is predetermined by germline acquired epigenetic marks, and the special case of *AgR* loci (see below), the initial choice of the allele to be expressed or repressed is stochastic. Although **stochastic choice** remains poorly understood, several scenarios have been evoked [7,20,22–25]. These include limiting concentrations of transcription factors (TFs) that are recruited to promoters or other key regulatory regions, infrequent assembly of the transcriptional machinery, and differential kinetics of juxtaposition of distant regulatory elements between the alleles. Notwithstanding, a monoallelic choice involves a stochastic initiation step in which there is a high chance to activate one allele only before the possible activation of the other(s). This activation process may initially be inefficient and slow, and takes place during a limited window of time. The monoallelic gene activation is followed by a maintenance step, and this often ensures that the 'lagging' allele(s) will remain silent.

At *AgR* loci, asynchronous replication set up during early development was proposed to be an epigenetic mark that potentiates monoallelic recombination [26]. At the *Igκ* locus, the allele that

loci. They feature high levels of transcription, active chromatin, and RAG binding.
**Stochastic choice:** the probability to activate/repress an allele is comparable between the two or more alleles present. This makes allelic choice random.
**T cell receptors (TCRs):** molecules that can bind antigenic peptides and are expressed on the surface of T lymphocytes.
**X-inactivation center (*Xic*):** the locus which controls X-inactivation in mammals and expresses the lncRNA *Xist* (X-inactive specific transcript).

Table 1. MAE in the Mouse

| Genetic loci | Chromosome (Chr.); gene numbers | Nature of monoallelic expression | Mechanism(s) involved | Biological function(s) | Refs |
|---|---|---|---|---|---|
| Immunoglobulin gene (Ig) loci: *IgH* *IgL: Igκ*, *Igλ* | *IgH*: Chr. 12<br>*Igκ*: Chr. 6<br>*Igλ*: Chr. 22 | Random, in B cells; $>10^{10}$ possible different immunoglobulins | Genetic recombination, chromatin alterations, ncRNA | Immune system: expression of one heavy chain and one light chain in each B cell | [2,25] |
| T cell receptor (TCR) loci: *TCRβ*, *TCRα* | *TCRβ*: Chr. 6<br>*TCRα*: Chr. 14 | Random, in T cells; $>10^{8}$ possible different TCRs | Genetic recombination, chromatin alterations, ncRNA transcription | Immune system: expression of one β and one α polypeptide in each αβ T cell | [2,65] |
| Olfactory receptor *(OR)* genes | >1400 genes<br>Many gene clusters | Random | Epigenetic, predominantly chromatin modifications | Olfaction: expression of one receptor per olfactory neuron | [4,24] |
| Protocadherin genes | 58 genes<br>Three autosomal gene clusters | Random, tissue-specific | Epigenetic, promoter choice and alternative splicing | Neurons: cell-surface diversity, signaling, neuronal survival | [5,124] |
| RMAE of unique genes | >1000 genes<br>Autosomal and X chromosome | Random, often tissue-specific | Epigenetic, histone methylation | Not clear: may provide diversity in cell identity | [6,20] |
| Genomic imprinting at autosomal genes | ~120 protein- coding genes<br>Hundreds of regulatory ncRNAs | Deterministic, often tissue-specific | Epigenetic, DNA methylation, chromatin modifications, lncRNAs | Not clear: genes involved in development, metabolism, and behavior | [9,10,52] |
| X-chromosome inactivation in females | X-linked genes (~2000) | Random, deterministic (imprinted); early embryo and trophoblast | Epigenetic, lncRNAs, chromatin modifications, DNA methylation | Compensates for differential dosage of gene expression between females and males | [7,11,12] |

### Box 1. Approaches and Precautions in Assessing MAE

Different approaches are used to detect MAE. To study random X-inactivation, MAE of *OR* genes, and other random systems, RNA FISH can be used, but this is less suitable for poorly expressed genes. Alternatively, single-cell RNA sequencing can be applied to hybrid animals with genetically different parental genomes, with again a handicap for poorly expressed genes. For *AgR* loci, crosses between mice with different genetic backgrounds, cell sorting, and quantitative RT-PCR can reveal allelic differences and help to determine MAE at the cell surface of individual immune cells. Because at imprinted genes the same parental allele is expressed in all cells, PCR can be performed on cDNA of tissues of hybrid animals, followed by allelic discrimination using single-nucleotide polymorphisms (SNPs). Imprinted gene loci have also been pinpointed through the presence of DMRs. Current RNA-Seq protocols yield many reads across polymorphic nucleotides, and even minute allelic expression differences become statistically significant. However, a smaller than twofold allelic difference may not be biologically relevant. The stringency of bioinformatics and applied thresholds also influences how many genes emerge [119,120], and divergent analytical approaches have given rise to highly discordant estimates of imprinted gene numbers [120]. Furthermore, apparent MAE may result from genetic differences between the parental alleles [121,122]. Gene regulatory sequences in humans frequently contain SNPs that are linked to different levels of DNA methylation which also influence gene expression [122]. In mice, genetic background effects can be ruled out by performing reciprocal crosses between different inbred strains. Classical PCR-based approaches can be used to confirm RNA-Seq data, and to prove that allelic differences are indeed transcriptional in nature, nascent RNA fluorescent *in situ* hybridization (FISH) can be applied [123]. In addition to limited tissue availability, cell contamination may also affect studies on embryonic lineages. For placental tissue, for instance, contaminating maternal cells can often give rise to apparent maternal allele-specific expression of genes of interest [121].

replicated earlier in the early embryo was generally the first to recombine at the right B cell developmental stage [26]. This deterministic model does not exclude an element of stochasticity because the paternal and the maternal alleles have an equal chance to undergo early replication, and in the whole B cell populations both alleles would be equally marked and represented [23]. For *AgR* loci, nevertheless, the important question of stochastic choice versus deterministic selection of the allele remains controversial [27,28]. In mouse embryonic stem and differentiated cells, MAE does not generally correlate with asynchronous replication timing [29]. In various instances, asynchronous replication can also not explain initiation of MAE, notably in systems where multiple gene clusters are scattered across different chromosomes, such as *OR* loci, or *AgR* loci in cartilaginous fish [20,30,31], discussed below.

Compared to the initiation process, the mechanisms that maintain MAE are relatively well understood. For some systems, this involves feedback inhibition of the other allele(s) that may be enforced by additional layers of regulation such as up- or downregulation of key effectors, chromatin modifications, and nuclear localization changes. Enforced expression of an alternative splicing product of the E2A transcription factor, for instance, antagonized the feedback inhibition of *TCRβ* rearrangements (Box 2). This suggests that **allelic exclusion** is influenced by the relative abundance of key TFs [32].

*OR* genes illustrate the hypothesized role of differential nuclear localization in the regulation of MAE. Repressed *OR* genes aggregate centrally in the nucleoplasm, forming distinct foci that feature interchromosomal interactions [33], whereas the single transcribed allele is localized outside these foci [34]. Enhanced expression from the single active *OR* gene elicits a negative feedback mechanism by inducing the unfolded protein response. This endoplasmic reticulum pathway [35,36] may act as a first checkpoint, before cell-surface expression of the chosen OR and its feedback signaling. The latter is important for the maturation of olfactory sensory neurons and their functional identity, and stabilizes both *OR* gene expression and silencing [37,38].

## Box 2. Brief Outline of Allelic Exclusion in Mice

*AgR* loci comprise variable (V), diversity (D), and joining (J) gene segments (Figure I). Before recombination, *AgR* loci are in a 'germline configuration'. The J regions – called **recombination centers** (RCs) – show non-coding transcription, active chromatin, and RAG binding controlled by close-by enhancers. RCs provide ideal environments for the initial cleavage of RSSs and for recruiting a physically distant V segment by large-scale chromatin looping/contraction. Theoretically, each locus can randomly assemble different V, (D), or J segments on both alleles through V(D)J recombination, and potentially encode multiple AgRs with different specificities. The vast majority of lymphocytes, however, only produce a single monospecific AgR.

V(D)J recombination (the figure shows the *IgH* locus) is associated with various transcriptional events and chromatin modifications, and is regulated by 'accessibility control elements', including enhancers, insulators, and promoters, in a cell type- and developmental stage-specific manner. Additional levels of regulation include 3D architecture, nuclear location, *trans*-acting factors, and large-scale chromosome dynamics [2,14]. *AgR* loci recombine in an ordered manner. In developing B and αβT cells, rearrangements of immunoglobulin heavy chain *(IgH)* and *TCRβ* loci precede those of immunoglobulin light chain loci *(Igκ* and *Igλ)* and *TCRα* respectively.

At the *IgH* and *TCRβ* loci, D to J recombination occurs on both parental alleles, and a V segment is then appended to a DJ segment. Although they have the potential to undergo bi-allelic recombination, one allele heads first for V–DJ recombination. Junctional diversification mechanisms randomly add or delete nucleotides at D–J and V–DJ junctions such that only ~1/3 of rearrangements are productive (i.e., in the correct reading frame for the production of Igμ heavy-chain or β polypeptide). Igμ and β chains assemble with other proteins to form the pre-BCR and pre-TCR respectively. A signal emitted from the surface by the pre-BCR or the pre-TCR instructs the cell to inhibit V–DJ recombination on the second *IgH* or *TCRβ* allele, respectively.

Pre-BCR and pre-TCR also signal initiation of recombination at the *Igκ* and *TCRα* loci, respectively. A productive rearrangement on the first *Igκ* allele or *TCRα* allele eventually leads to κ light-chain assembly with Igμ, and of an α chain with a β chain, forming a BCR or a TCR, respectively. Surface BCR and TCR molecules signal a halt of recombination at the second *Igκ* and *TCRα* alleles.

If the first V–(D)J rearrangement is not productive (i.e., no feedback inhibition), the second allele can undergo V–(D)J recombination. If the second rearrangement is itself non-productive, the cell undergoes apoptosis. Thus, allelic exclusion of *AgR* loci is established during V(D)J recombination [1,115].

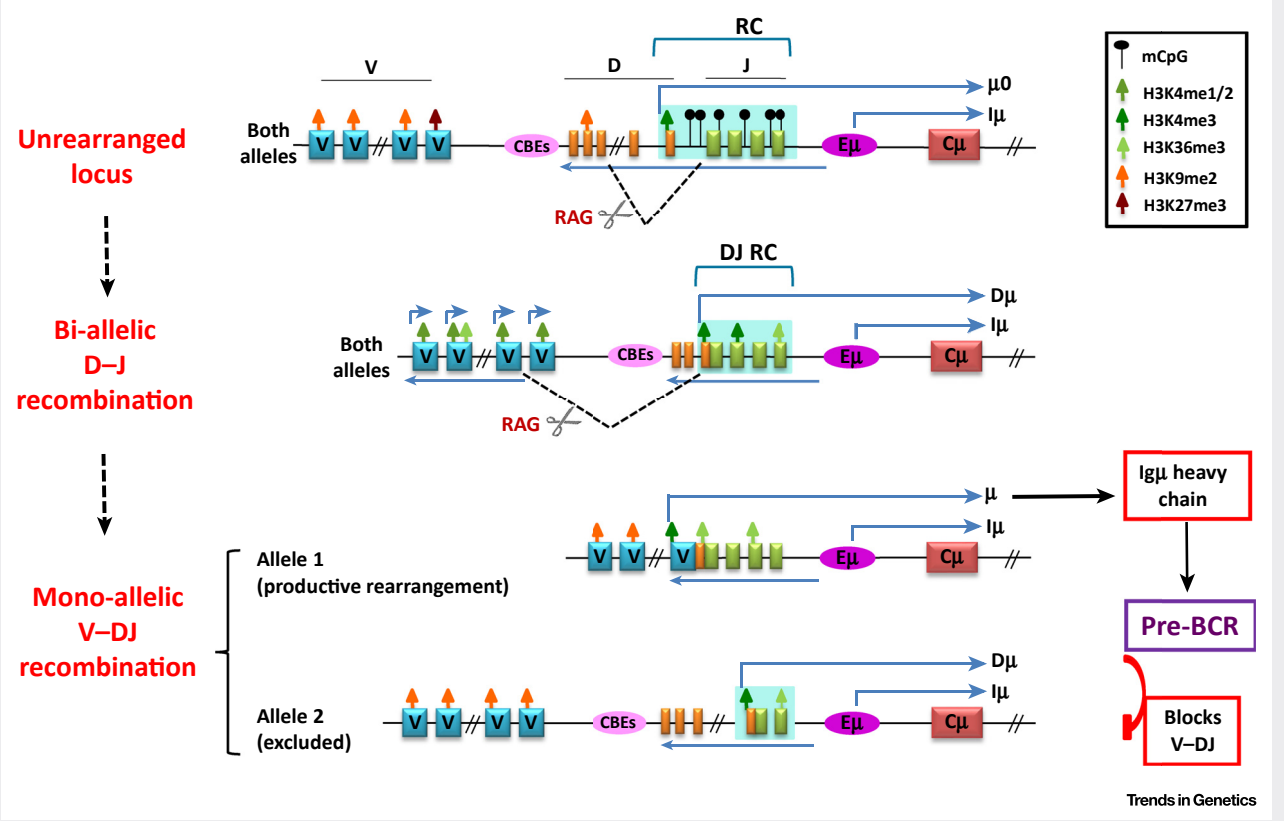


Figure I. V(D)J Recombination and Allelic Exclusion at *AgR* Loci.

## Transcriptional Events in the Acquisition of MAE

Regulatory sequences that control MAE are present on both alleles, but exert their effects on one allele only. One consequence is differential transcription, which has epigenetic effects as such, or produces regulatory RNAs that control chromatin and gene expression in *cis*. The best-understood example is genomic imprinting. Establishment of DNA methylation imprints at 'maternal ICRs' occurs during oogenesis (Figure 1A) and is guided by transcription [16,39,40]. Oocyte-specific promoters drive transcription through these **CpG islands**. For several maternal ICRs, this read-through was proven to be essential for methylation acquisition [40–42]. The process involves prior removal of H3K4 methylation, and recruitment of the DNMT3A–DNMT3L methyltransferase complex that is brought to the DNA by transcription-associated H3K36me3 [16,40,43,44]. Maternal ICRs do not become methylated in male germ cells in which 'paternal ICRs' become methylated, and in which methylation acquisition is not linked to transcription and is more widespread along the genome than in oocytes [39].

Most autosomal imprinted gene domains express at least one lncRNA during embryogenesis. At several, this mediates chromatin repression at close-by gene. The *Kcnq1* imprinted domain on mouse chromosome 7 has an intragenic ICR that comprises a promoter which transcribes an lncRNA (called *Kcnq1ot1*) from the unmethylated paternal allele only (Figure 2B). In the placenta, this lncRNA mediates repression at several nearby genes, through **Enhancer of Zeste homolog-2** (EZH2)/**Polycomb repressive complex 2** (PRC2)-controlled H3K27me3, RING1B/**Polycomb repressive complex 1 (**PRC1)-controlled H2AK119ub, and H3K9me2 controlled by KMT1C [45–49]. A repeat motif in its 5′ part is essential to bring the lncRNA *Kcnq1ot1* to its target genes [50]. Similarly, at the imprinted IGF2 receptor (*Igf2r*) gene on mouse chromosome 17, and the *Dlk1–Dio3* locus on mouse chromosome 12, ICR-controlled lncRNAs mediate silencing at close-by genes in a tissue-specific manner [51–53].

Transcription-associated acquisition of DNA methylation occurs also in the early embryo, at several somatic **differentially methylated regions** (DMRs) of imprinted domains. Similarly for methylation acquisition in oocytes, this process involves allelic transcription through these DMR regions [54–56].

In X-inactivation in the mouse, similarly, the *Xist* promoter becomes DNA-methylated on its non-activated allele through overlapping transcription of an antisense lncRNA called *Tsix* [57,58]. Several other lncRNAs transcribed from the **X-inactivation center** (*Xic*) contribute to the complex control of *Xist* expression as well [7,59].

Transcriptional upregulation of *Xist* on one of the two X chromosomes occurs in the preimplantation mouse embryo [7]. On this randomly chosen allele, the initiation of X-inactivation reduces the probability that inactivation occurs on the opposite X chromosome. This negative feedback is proposed to involve a reduced expression dosage of the X-encoded RNF12, an E3 ubiquitin ligase which influences the level of *Xist* expression [60].

ncRNA transcription is functionally important at *AgR* loci as well. Induced premature termination of transcriptional elongation across the J cluster of the *TCRα* locus demonstrated that transcription through the RSSs renders them accessible to the RAG complex [61]. Specific recognition of transcription-associated H3K4me3 by a 'plant-homeodomain' of RAG2 is also essential for efficient V(D)J recombination [62,63]. ncRNA transcription initially occurs at the **recombination centers** (RCs) before, and following, D–J recombination (Box 2). This is controlled by specific enhancers whose deletion severely affects V(D)J recombination, with

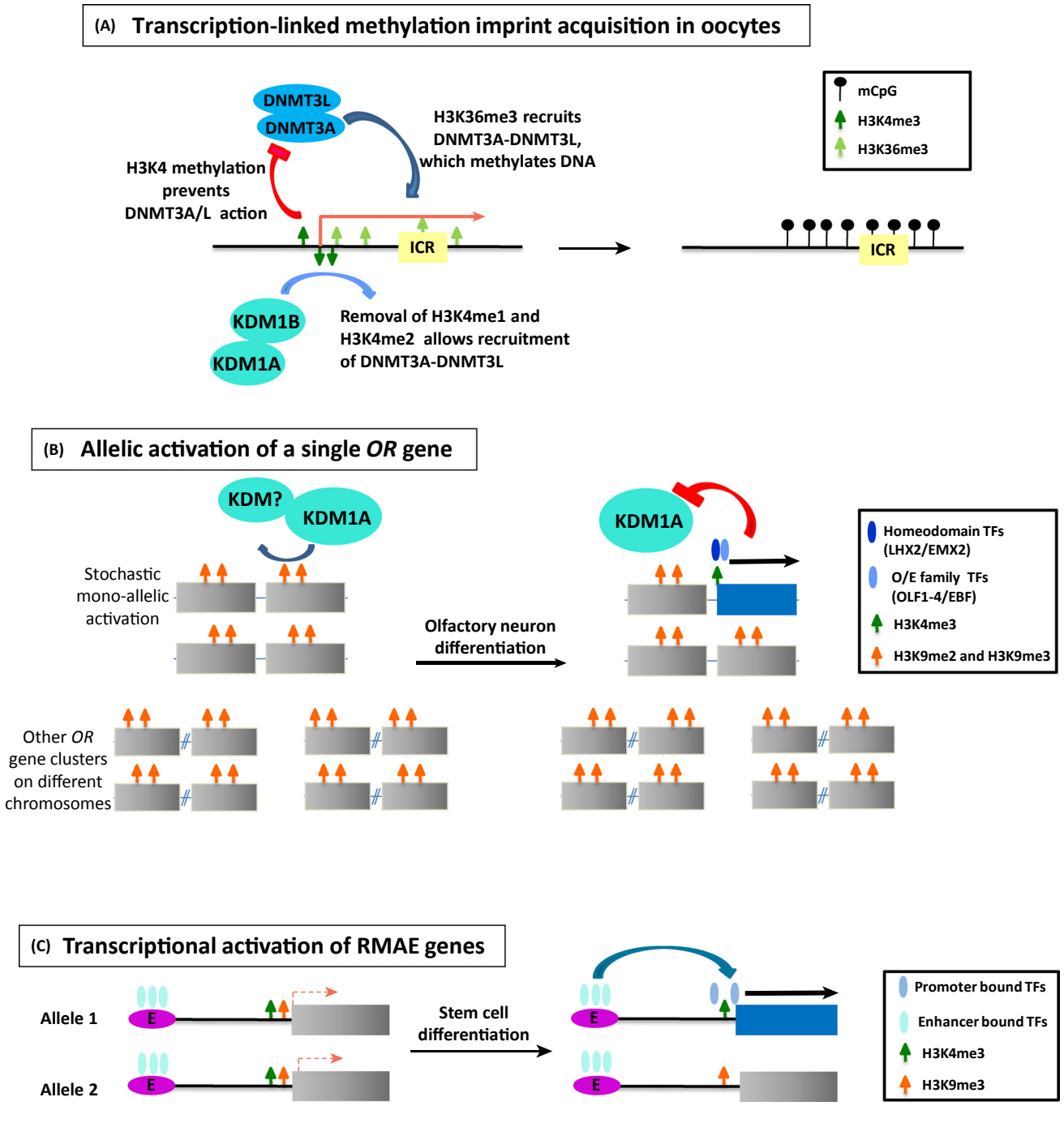


Figure 1. Transcriptional Events in the Initiation of Monoallelic Expression (MAE). (A) Acquisition of maternal methylation imprinting in growing oocytes at an imprinting control regions (ICR; yellow rectangle) that shows transcription-associated histone H3 trimethylation on lysine 36 (H3K36me3), which recruits the DNMT3A–DNMT3L *de novo* methyltransferase complex. H3K4me prevents this complex from binding to chromatin, and is removed by KDM1A/KDM1B to allow subsequent methylation acquisition. (B) Activation of a single olfactory receptor (*OR*) allele during olfactory neuron differentiation involves chromatin derepression by KDM1A, enhancer-mediated recruitment of OLF1/EBF and homeodomain transcription factors (TFs), and acquisition of active chromatin. All other *OR* genes remain repressed by H3K9me2/3 (orange arrows). (C) Developmental random activation of a MAE allele, with loss of H3K9me3, TF recruitment, and enhancer activity.

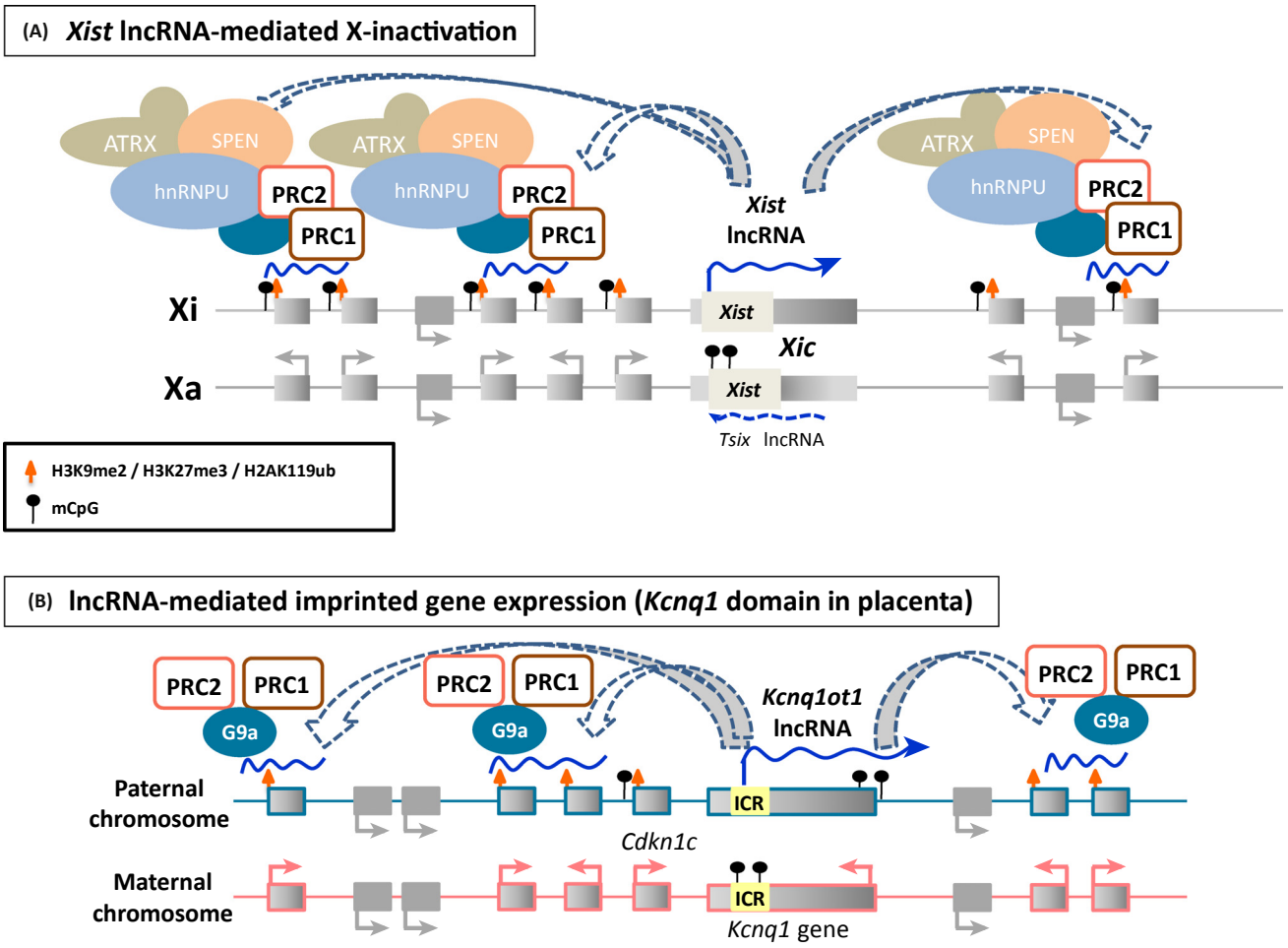


Figure 2. lncRNA-Mediated Allelic Chromatin Repression in X-Inactivation and Genomic Imprinting. (A) The X inactivation center (*Xic*) controls X-inactivation in females. In the early mouse embryo, *Xist* lncRNA expression is randomly upregulated on one allele. It accumulates in *cis* on the chromosome and recruits chromatin proteins and regulatory complexes (including PRC1, PRC2, SPEN, ATRX, and hnRNPU), which brings about diverse epigenetic alterations including DNA methylation, ultimately leading to gene repression. Some genes on the inactive X (Xi) escape X-inactivation. On the X that remains active (Xa), transcription of the antisense lncRNA *Tsix* and its associated histone H3 trimethylated on lysine 36 (H3K36me3) mediate repressive DNA methylation at the *Xist* promoter during stem cell differentiation. (B) At the *Kcnq1* domain on mouse chromosome 7, the lncRNA *Kcnq1ot1* is expressed from the unmethylated paternal ICR only. This induces repression in *cis* at multiple genes in the placenta, which involves recruitment of PRC1, PRC2, and lysine methyltransferase KMT1C, which generate H2AK119ub, H3K27me3, and H3K9me2, respectively. Except for the promoter of one gene, *Cdkn1c*, at which the lncRNA is required to maintain CpG methylation [118], lncRNA-mediated *cis* repression does not involve DNA methylation [45,46,49]. Abbreviation: ub, ubiquitin.

loss of ncRNA transcription and active chromatin modifications [23,64,65]. Following D–J recombination, transcription across the newly formed D–J RCs likely influences the capture of a single V segment for V–DJ recombination (Box 2). It also ensures that the DJ RCs are available for V–DJ recombination on the second allele in case of defective rearrangement on the first allele [66]. Within the *Igκ* variable region, importantly, each B cell displays allele-specific chromatin marks and transcription such that the V segments on one allele are differently marked than the V segments on the other [67].

Multiple sense and antisense transcripts are also generated throughout the *IgH* variable region [68,69], and are downregulated following a productive rearrangement [69]. In contrast to relatively short V sense transcripts, antisense transcripts are long and extend across multiple V genes and intergenic regions. They modify the entire variable region before the onset of V–DJ recombination, through recruitment of the chromatin remodeling complex SWI/SNF, removal/loss of repressive histone modifications, and acquisition of active modifications [69,70].

Transcriptional activation is essential also in the acquisition of a single active *OR* allele (Figure 1B). *OR* genes are organized in multiple clusters that include enhancers. Deletion of individual enhancers reduces the expression of nearby *OR* genes. Strong interchromosomal interactions occur between *OR* enhancers as well as with the single stochastically activated *OR* allele. This 'multi-enhancer hub' and its engagement in *trans*-interactions is likely involved in the 'choice' of a single *OR* gene [33,34,71–73]. Recruitment of TFs into the multi-enhancer hub ensures stable and high-level transcription of the single present allele [72]. Translation of the mRNA of the transcribed *OR* allele ultimately leads to inhibition of lysine demethylase KDM1A, and this prevents aberrant activation of additional OR genes [35,36].

## DNA and Histone Methylation Are Associated with Allelic Gene Repression

Following fertilization, the germline-acquired DNA methylation imprints at ICRs are maintained in all somatic lineages [16]. The way they bring about monoallelic gene expression during development differs between clusters, however, and is often tissue-specific [9,15,52]. A striking example is provided by the *GRB10* locus. This negative growth regulator is expressed from the maternal chromosome in the placenta, whereas in adult brain, where the gene controls behavior, expression occurs on the paternal chromosome only [74,75].

Essential to the somatic maintenance of methylation imprints at ICRs is the KRAB-domain zinc-finger protein, ZFP57, which recognizes a sequence motif present at ICRs, but only when this is methylated. ZFP57 recruits TRIM28, a platform protein for histone modifiers, including lysine methyltransferase SetDB1, which regulates H3K9me3 at ICRs [16]. These and other histone-modifying enzymes (Table 2) contribute to protection of ICRs against somatic loss of DNA methylation [15].

X-inactivation in the early mouse embryo also involves acquisition of repressive chromatin modifications at the promoters of most X-linked genes. Following its allelic transcriptional upregulation, *Xist* lncRNA associates with and 'coats' the X chromosome in *cis*. This is linked to the formation of a nuclear compartment into which genes are recruited and from which RNA polymerase II is excluded. There is *Xist* lncRNA-linked acquisition of repressive H3K27me3, H2AK119ub, and H3K9me2, and, ultimately, CpG methylation at gene promoters on the inactive X [7] (Figure 2A). Gene bodies, however, are more highly methylated on the active X, which reflects the maintained gene expression on this X chromosome [76].

Various proteins interact with *Xist* lncRNA and contribute to its repressive effects in *cis* [11,77,78]. These include the nuclear ribonucleoproteins hnRNPU and hnRNPK (Table 2), which bind to conserved repeats at the 5′ end of the lncRNA, and the transcriptional repressor SPEN [79–83]. ATRX interacts with the 5′ repeats as well, and this chromatin remodeler facilitates recruitment of Polycomb repressive complex-2 (PRC2) and the histone variants H3.3 and macro-H2A1 [84,85].

Table 2. Chromatin Regulatory Proteins Involved in MAE

| Chromatin regulators | Process | Activity | Refs |
|---|---|---|---|
| PRC2-type complex (EZH2 lysine methyltransferase) | X-inactivation, genomic imprinting, *AgR* loci | Mediates H3K27me3 and gene repression | [46,94,125–129] |
| PRC1-type complex (RING1B ubiquitylase) | X-inactivation, genomic imprinting, *AgR* loci (?) | Mediates H2AK119ub and gene repression | [46,83,127,129,130] |
| KMT1E (SETDB1, ESET) | Genomic imprinting, X inactivation | Mediates H3K9me3 | [131,132] |
| KMT1C (G9A, EHMT2) | Genomic imprinting, *AgR* loci [V(D)J recombination at the *Igλ* locus] | Mediates H3K9me2<br>Linked to lamina-associated domains | [47,51,133] |
| KDM1A (LSD1)<br>KDM1B (LSD2) | OR expression, genomic imprinting | Demethylate H3K4me1/2 and H3K9me1/2 | [36,43,134] |
| DNMT3A/B DNMT3L | X-inactivation, genomic imprinting | Acquisition of *de novo* DNA methylation | [44,135,136] |
| SMCHD1 | X-inactivation, genomic imprinting, RMAE, including protocadherin genes | Hinge-domain protein involved in maintenance of DNA methylation and gene repression | [137,138] |
| ATRX | X inactivation, genomic imprinting | Chromatin remodeling complex<br>Incorporates H3.3, maintains allelic repression | [84,139] |
| hnRNPU (SAF-A)<br>hnRNPK<br>SPEN (SHARP) | X-inactivation | (*Xist*-) RNA binding proteins that contribute to transcriptional silencing | [77,78,80–83] |
| CTCF<br>Cohesins | X-inactivation, *AgR* loci, genomic imprinting, protocadherin genes, *OR* genes (?) | Structure higher-order chromatin<br>Mediate functional DNA interactions | [11–13,140–142] |

X-inactivation maintenance requires continued *Xist* lncRNA expression, particularly in extra-embryonic lineages [58]. Although DNA methylation provides stability in the embryo, it is less important for imprinted X-inactivation in the mouse trophoblast [86], possibly because of additional stabilizing chromatin modifications in this lineage.

RMAE at individual genes is associated with repressive histone methylation at promoters (Figure 1C), but seems to be independent of DNA methylation [20,87]. This mode of repression is less stable than X-inactivation and imprinting, and thus may allow allelic derepression or reversal to occur during the lifespan of cellular clones. Whereas in undifferentiated ES cells only small numbers of genes show RMAE, in neural progenitor cells hundreds of genes acquire relatively stable RMAE. Their specific allelic activation seems to be regulated at the promoter level rather than by changes at enhancers. Many genes with RMAE show allele-specific H3K4me3 and H3K9me3 at their promoters on the transcriptionally activated and repressed alleles, respectively [87]. Whether this is a cause or a consequence of the allelic expression remains unclear. Gene promoters also show increased chromatin accessibility on the expressed allele. This persists during mitosis, which may indicate somatic heritability of TF binding [87]. Most single genes with RMAE are expressed at only low levels, and percentile ratios of cells that express one, the other, or both parental alleles are suggestive of a stochastic activation process controlled by TFs [20,87].

At *AgR* loci, coordinated chromatin changes direct the accessibility, activation, and, ultimately, monoallelic recombination. Broadly outlined, the germline *IgH* and *TCRβ* loci are uniformly

marked with repressive H3K9me2 and H3K27me3, and there is extensive DNA methylation. The developmentally controlled locus activation is associated with highly localized, bi-allelic recruitment of SWI/SNF complexes and acquisition of active histone modifications, including H3K9ac, H3K4me3, and H3K36me3, mostly at the RCs. Following D–J recombination, DNA at the newly formed DJ RCs becomes demethylated in an enhancer-dependent manner through an unknown mechanism [66,88,89]. Significantly, the *Igκ* locus was shown to undergo B cell-specific, monoallelic DNA demethylation across the J–C region, and the demethylated allele is preferentially targeted for recombination [90].

The *IgH* variable region covers 2.5 Mb and shows distinct histone modifications. The distal portion is transcribed and enriched in H3K36me3 and H3K4me2, whereas the proximal domain is enriched in H3K27me3. The ncRNAs of the distal region could play a role in this process by preventing the repressive functions of the PRC2 complex [91] at this part of the variable region. How H3K27me3 and other repressive histone methylations are removed, and active marks acquired, remains unclear [14,70,92], but this may involve recruitment of TFs. At *IgH*, for instance, loss of H3K9me2 at the V region before V–DJ recombination is dependent on PAX5 [93]. Binding of STAT5 to an essential *Igκ* enhancer leads to EZH2-mediated H3K27me3 and repression of non-coding transcription [94].

Chromatin repression also plays a role in *OR* regulation (Figure 1B). In multipotent neural progenitor cells, the different *OR* gene clusters are enriched in H3K9me2 and H4K20me2. In the main olfactory epithelium, there is additional acquisition of H3K9me3 and H4K20me3, both typical of heterochromatin. This configuration is characteristic of all repressed *OR* genes, but is absent from the single expressed allele [95], whose activation is associated with acquisition of H3K4me3 [33,36,96].

## Evolutionary Parallels between MAE Systems

Genomic imprinting arose about 180 million years ago [97], when placentation and other maternal investments became of growing importance for reproduction. This has evoked different evolutionary theories [98]. Mechanistically, the emergence of imprinting in mammals is linked to the acquisition of repressive DNA methylation at specific regions in the male and female germlines [97]. This finding does not exclude that imprinting initially arose independently of DNA methylation. At many genes there is acquisition of H3K27me3 during oogenesis [18,19] (Figure 3A), and this maternal chromatin imprint induces a transient, partial, repression of the maternal allele in the early embryo [18]. Oocyte-acquired H3K27me3 is functionally important in fruit fly embryos as well [99], and could represent an ancestral imprinting mechanism. Data on one of the evolutionarily oldest imprinted genes in mammals – called *Slc38a4* – suggest that DNA methylation was coopted to confer more stable somatic maintenance of the maternal histone methylation imprint [17,97,100,101].

In mice, X-inactivation in the extra-embryonic lineages is imprinted and controlled by oocyte-acquired H3K27me3 at the *Xist* promoter (Figure 3B). This H3K27me3 imprint prevents *Xist* activation on the maternal chromosome, and X-inactivation therefore occurs on the paternal X in the early preimplantation embryo [18]. The imprinted X-inactivation is maintained subsequently in the trophoblast only [86,102], similarly to H3K27me3-controlled imprinting at some autosomal loci [18]. In marsupials, however, X-inactivation is imprinted both in the embryo and the trophoblast, and is controlled by an lncRNA that is unrelated to *Xist* [103]. Combined, the recent findings suggest that X-inactivation and imprinting evolved in part from ancestral, DNA methylation-independent, mechanisms that involved H3K27me3.

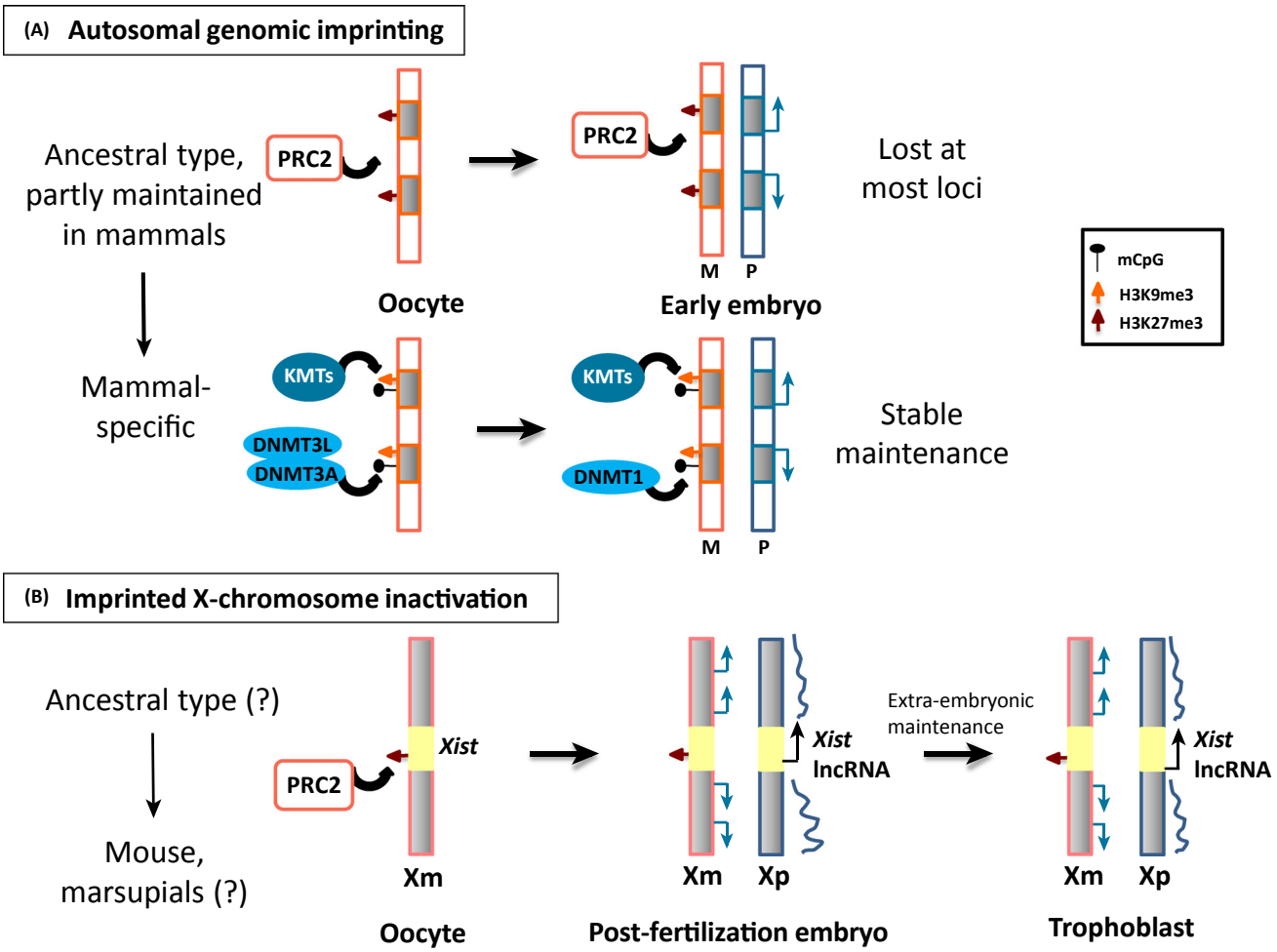


Figure 3. Evolutionary Parallels between Imprinted X-Inactivation and Genomic Imprinting. (A) Oocyte-acquired histone H3 lysine 27 trimethylation (H3K27me3) imprints mediate paternal allele-biased gene expression in the early embryo, apparently independently of DNA methylation. This ancestral type of imprinting likely evolved into a more stable imprinting mechanism involving both histone (H3K27me3 or H3K9me3) and DNA methylation [17]. (B) In imprinted X-inactivation, oocyte-acquired H3K27me3 contributes to the silencing the maternal *Xist* allele (mouse) and possibly also its functional equivalent in marsupials [18]. The resulting inactivation of the paternal X is maintained in the extra-embryonic lineages (and is lost in the embryo proper, where random X-inactivation is initiated subsequently) in the mouse, and in a constitutive manner in marsupials [86,102]. Abbreviations: M/m, maternal; P/p, paternal.

MAE of *AgR* loci in B and T lymphocytes arose >500 million years ago [31]. Following the appearance of V(D)J recombination, likely through a transposable element containing RAG1-like and RAG2-like genes [104,105], major differences evolved between species with regards to the organization and function of *AgR* loci. For instance, the ordered rearrangement (bi-allelic D–J, then monoallelic V–DJ) that is central to feedback regulation in mice and humans is not found in other mammals. In the rabbit, interestingly, almost half the non-expressed *IgH* alleles do not undergo D–J recombination, which suggests that the kinetics of D–J rearrangements may be the rate-limiting step [106].

The most striking divergence is observed in cartilaginous fish, where ordered rearrangements are completely absent (V–D can occur before D–J). In sharks, once an *IgH* locus is activated, all its gene segments immediately rearrange to completion. Activation is monoallelic and the allele

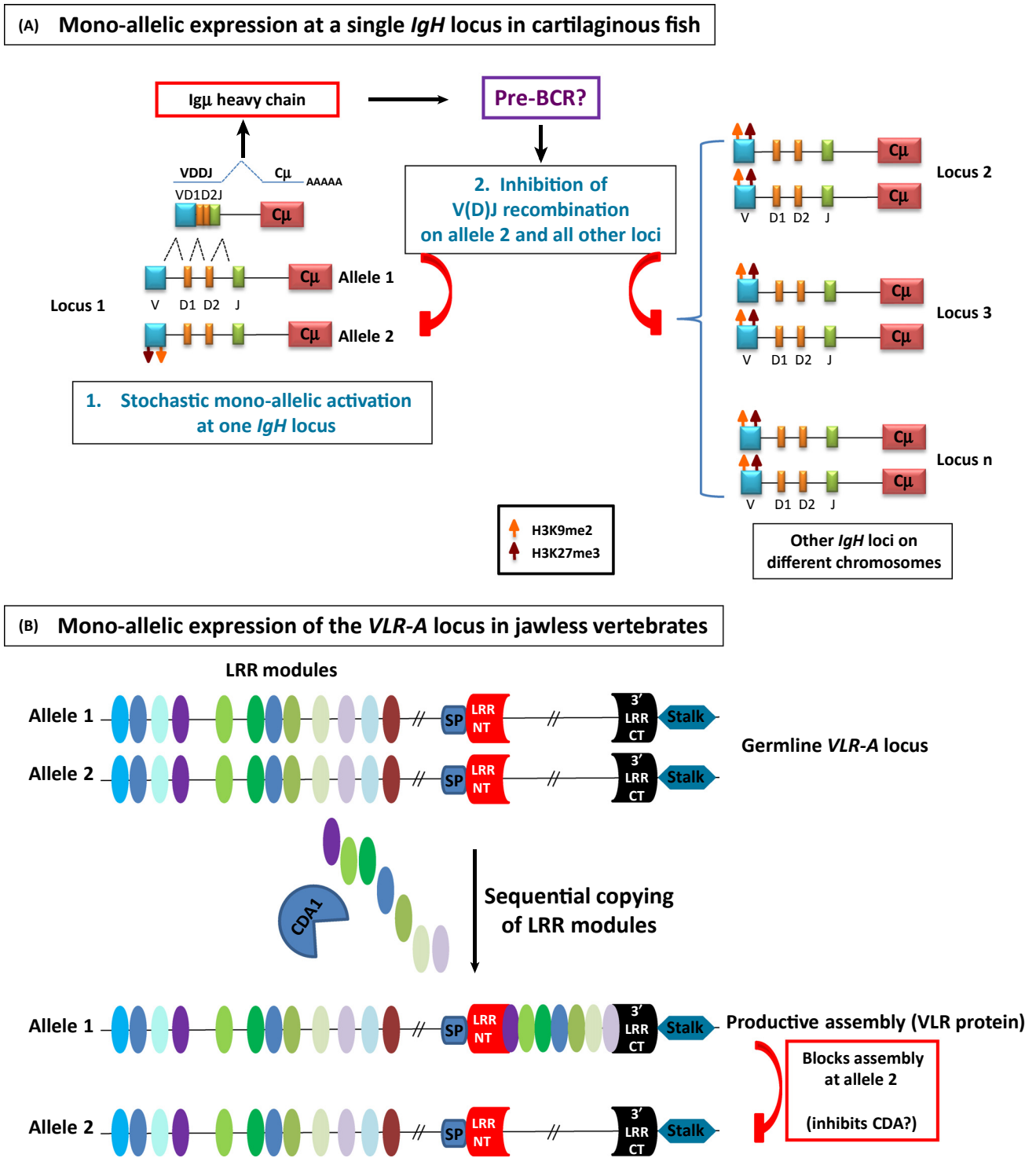


Figure 4. Structure of Antigen Receptor (*AgR*) Loci in Cartilaginous and Jawless Fish. (A) Cartilaginous fish have many *IgH* loci on different chromosomes [109]. During B cell development, one allele is randomly activated. If its rearrangement gives a functional Igμ product, this leads to the formation of a pre-BCR and

*(Figure legend continued on the bottom of the next page.)*

of a (productively or non-productively) rearranged gene retains its configuration [107,108]. Cartilaginous fish, interestingly, have multiple *Ig* gene clusters on different chromosomes [109], an organization reminiscent of the *OR* gene clusters. Tens to hundreds of alleles are prevented from undergoing recombination in developing B cells (Figure 4A).

In jawless fish (Agnathans), a different type of adaptive immunity evolved mediated by leucine-rich receptors known as 'variable lymphocyte receptors' (VLRs). Three *VLR* genes have been identified. *VLRA* and *VLRC* are expressed in T cell-like lymphocytes, whereas *VLRB* is expressed in B cell-like lymphocytes. Agnathans are devoid of *BCR* and *TCR* loci, and of RAG, and therefore do not undergo V(D)J recombination [110]. Instead, they diversify their AgRs through a conversion-like mechanism involving two members of the AID/APOBEC cytosine deaminase family, CDA1 and CDA2, which mediates insertion of leucine-rich repeats into the *VLR* genes [111]. Importantly, the expression of functional VLRs is largely monoallelic (Figure 4B). In the vast majority (>90%) of cases where one allele is productively assembled, the other allele retains a germline configuration. Where bi-allelic assembly was detected, only one was productive [112,113]. These figures are strongly suggestive of an operational and effective feedback inhibition.

Whether silencing of the excluded alleles in cartilaginous and jawless fish involves repressive histone methylation, as at *OR* gene clusters, remains unknown. Notwithstanding, feedback inhibition as an efficient mechanism to ensure allelic exclusion clearly arose early during evolution and was coopted by various species for different biological purposes.

## Concluding Remarks and Future Perspectives

Different epigenetic and genetic modes of MAE exist in mammals. We have presented examples of how transcription and lncRNAs control the acquisition and/or maintenance of MAE, but many aspects of RNA biology remain to be explored. Tremendous progress has been made on the intricacies of *Xist*-mediated chromatin repression during X-inactivation. Less is known about chromatin repression and RNA regulatory functions in other systems. Further mechanistic studies will be necessary to better understand how deregulation of MAE affects development and causes human diseases. Although not the emphasis of this review, 'loss of imprinting' (biallelic expression) is causally involved in many congenital disease syndromes, deregulated X-inactivation affects X-linked gene dosage and gives embryonic lethality, perturbations in V(D)J recombination and AgR expression affect immune cell function and may lead to cancer or autoimmune diseases, and perturbations in other RMAE systems can have developmental effects [6,9,16,20,30,114–116].

For random mechanisms (MAE systems other than genomic imprinting), a key question where progress has been difficult is how the expression of one allele initiates before that of the other. Also largely unexplored is the kinetics of the events that take place between the initial selection of one allele, and feedback inhibition of the opposite allele(s). A recent study, based on *IgH* locus engineering, live cell imaging, and modeling of large-scale locus movements, highlights

feedback inhibition of recombination on the other allele and all other *IgH* loci [107,108]. Non-activated alleles presumably retain their repressive histone H3 lysine 27 trimethylation (H3K27me3) and H3K9me2. (B) The germline *VLR* genes of jawless vertebrates (lamprey and hagfish) are flanked by hundreds of leucine-rich repeats (LRRs), and are composed of a conserved signal peptide (SP), highly variable LRR modules including an N-terminal LRR module (LRRNT), a C-terminal module (3′LRRCT), and a conserved stalk region (for membrane anchorage). Assembly of a *VLR* gene involves a gene conversion-like mechanism mediated by CDA enzymes (CDA1 for *VLRA* and *VLRC*, CDA2 for *VLRB*) which use the flanking LRRs as templates. Fragments of LRRs are sequentially copied at the intervening, non-coding, sequence between LRRNT and 3′LRRCT such that the intervening sequence is ultimately replaced by parts of LRR modules. If the assembly is productive, surface expression of the VLR protein signals inhibition of assembly on the second allele [111–113].

the importance of spatial confinement for the kinetics of V(D)J recombination [117]. Similar studies will help to elucidate the mechanisms that trigger the onset of allelic exclusion and how these stabilize MAE and confer functional identity.

The various modes of MAE evolved because of diverse selective pressures and serve different biological purposes. Nonetheless, recent research has unraveled strong mechanistic similarities. Combined with comparative studies between species, this suggests that that they may have evolved from analogous ancestral processes.

## Outstanding Questions

Which factors control stochastic gene activation in different types of MAE?

What controls the lineage-specificity of *cis* effects of (nc)RNA expression?

What is the causal relationship between allelic exclusion and chromatin alterations?

Are all aspects of allelic exclusion cell type-specific, or is there epigenetic priming?

Does the inclusion of somatic genetic recombination provide a selective advantage?

To what extent were ancestral MAE mechanisms independent of DNA methylation?

## Acknowledgments

The authors apologize to colleagues whose research could not be cited owing to space limitations. They thank Dr Daniel Fisher (IGMM) for reading the manuscript, and the anonymous reviewers for their valuable comments and suggestions. R. F. acknowledges grant funding from Fondation Recherche Médicale (FRM; DEQ2015033170), the Agence Nationale de Recherche (ANR-13-BSV2-0014), and Labex 'EpiGenMed', an ANR-funded 'investissement d'avenir' programme (ANR10-LABX-12-01). A.A.K. is grant funded by the Institut National du Cancer (INCA-9363/PLBIO15-134), the ANR (ANR-16-CE12-0017), and the Fondation ARC pour la Recherche sur le Cancer (PJA 20141201647).

## References

1. Jung, D. *et al.* (2006) Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu. Rev. Immunol.* 24, 541–570
2. Schatz, D.G. and Ji, Y. (2011) Recombination centres and the orchestration of V(D)J recombination. *Nat. Rev. Immunol.* 11, 251–263
3. Chess, A. *et al.* (1994) Allelic inactivation regulates olfactory receptor gene expression. *Cell* 78, 823–834
4. Monahan, K. and Lomvardas, S. (2015) Monoallelic expression of olfactory receptors. *Annu. Rev. Cell Dev. Biol.* 31, 721–740
5. Chen, W.V. and Maniatis, T. (2013) Clustered protocadherins. *Development* 140, 3297–3302
6. Chess, A. (2016) Monoallelic gene expression in mammals. *Annu. Rev. Genet.* 50, 317–327
7. Gendrel, A.V. and Heard, E. (2014) Noncoding RNAs and epigenetic mechanisms during X-chromosome inactivation. *Annu. Rev. Cell Dev. Biol.* 30, 561–580
8. Dupont, C. and Gribnau, J. (2013) Different flavors of X-chromosome inactivation in mammals. *Curr. Opin. Cell Biol.* 25, 314–321
9. Peters, J. (2014) The role of genomic imprinting in biology and disease: an expanding view. *Nat. Rev. Genet.* 15, 517–530
10. Ferguson-Smith, A.C. (2011) Genomic imprinting: the emergence of an epigenetic paradigm. *Nat. Rev. Genet.* 12, 565–575
11. da Rocha, S.T. and Heard, E. (2017) Novel players in X inactivation: insights into Xist-mediated gene silencing and chromosome conformation. *Nat. Struct. Mol. Biol.* 24, 197–204
12. Robert Finestra, T. and Gribnau, J. (2017) X chromosome inactivation: silencing, topology and reactivation. *Curr. Opin. Cell Biol.* 46, 54–61
13. Jegu, T. *et al.* (2017) The X chromosome in space. *Nat. Rev. Genet.* 18, 377–389
14. Bossen, C. *et al.* (2012) Chromatin topology and the regulation of antigen receptor assembly. *Annu. Rev. Immunol.* 30, 337–356
15. Sanli, I. and Feil, R. (2015) Chromatin mechanisms in the developmental control of imprinted gene expression. *Int. J. Biochem. Cell Biol.* 67, 139–147
16. Kelsey, G. and Feil, R. (2013) New insights into establishment and maintenance of DNA methylation imprints in mammals. *Philos. Trans. R. Soc. Lond. Biol. Sci.* 368, 20110336
17. Pathak, R. and Feil, R. (2017) Oocyte-derived histone H3 lysine 27 methylation controls gene expression in the early embryo. *Nat. Struct. Mol. Biol.* 24, 685–686
18. Inoue, A. *et al.* (2017) Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature* 547, 419–424
19. Inoue, A. *et al.* (2017) Genomic imprinting of Xist by maternal H3K27me3. *Genes Dev.* 31, 1927–1932
20. Gendrel, A.V. *et al.* (2016) Random monoallelic expression of genes on autosomes: parallels with X-chromosome inactivation. *Semin. Cell Dev. Biol.* 56, 100–110
21. Savova, V. *et al.* (2016) Genes with monoallelic expression contribute disproportionately to genetic diversity in humans. *Nat. Genet.* 48, 231–237
22. Eckersley-Maslin, M.A. and Spector, D.L. (2014) Random monoallelic expression: regulating gene expression one allele at a time. *Trends Genet.* 30, 237–244
23. Levin-Klein, R. and Bergman, Y. (2014) Epigenetic regulation of monoallelic rearrangement (allelic exclusion) of antigen receptor genes. *Front. Immunol.* 5, 625
24. Magklara, A. and Lomvardas, S. (2013) Stochastic gene expression in mammals: lessons from olfaction. *Trends Cell Biol.* 23, 449–456
25. Proudhon, C. *et al.* (2015) Long-range regulation of V(D)J recombination. *Adv. Immunol.* 128, 123–182
26. Mostoslavsky, R. *et al.* (2001) Asynchronous replication and allelic exclusion in the immune system. *Nature* 414, 221–225
27. Farago, M. *et al.* (2012) Clonal allelic predetermination of immunoglobulin-kappa rearrangement. *Nature* 490, 561–565
28. Alves-Pereira, C.F. *et al.* (2014) Independent recruitment of Igh alleles in V(D)J recombination. *Nat. Commun.* 5, 5623
29. Rivera-Mulia, J.C. *et al.* (2018) Allele-specific control of replication timing and genome organization during development. *Genome Res.* 28, 800–811
30. Vettermann, C. and Schlissel, M.S. (2010) Allelic exclusion of immunoglobulin genes: models and mechanisms. *Immunol. Rev.* 237, 22–42
31. Hsu, E. (2016) Assembly and expression of shark Ig genes. *J. Immunol.* 196, 3517–3523
32. Agata, Y. *et al.* (2007) Regulation of T cell receptor beta gene rearrangements and allelic exclusion by the helix-loop-helix protein, E47. *Immunity* 27, 871–884
33. Clowney, E.J. *et al.* (2012) Nuclear aggregation of olfactory receptor genes governs their monogenic expression. *Cell* 151, 724–737
34. Markenscoff-Papadimitriou, E. *et al.* (2014) Enhancer interaction networks as a means for singular olfactory receptor expression. *Cell* 159, 543–557

35. Dalton, R.P. *et al.* (2013) Co-opting the unfolded protein response to elicit olfactory receptor feedback. *Cell* 155, 321–332
36. Lyons, D.B. *et al.* (2013) An epigenetic trap stabilizes singular olfactory receptor expression. *Cell* 154, 325–336
37. Ferreira, T. *et al.* (2014) Silencing of odorant receptor genes by G protein betagamma signaling ensures the expression of one odorant receptor per olfactory sensory neuron. *Neuron* 81, 847–859
38. Serizawa, S. *et al.* (2003) Negative feedback regulation ensures the one receptor-one olfactory neuron rule in mouse. *Science* 302, 2088–2094
39. Kobayashi, H. *et al.* (2012) Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. *PLoS Genet.* 8, e1002440
40. Veselovska, L. *et al.* (2015) Deep sequencing and *de novo* assembly of the mouse oocyte transcriptome define the contribution of transcription to the DNA methylation landscape. *Genome Biol.* 16, 209
41. Chotalia, M. *et al.* (2009) Transcription is required for establishment of germline methylation marks at imprinted genes. *Genes Dev.* 23, 105–117
42. Joh, K. *et al.* (2018) Growing oocyte-specific transcription-dependent *de novo* DNA methylation at the imprinted Zrsr1-DMR. *Epigenetics Chromatin* 11, 28
43. Ciccone, D.N. *et al.* (2009) KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* 461, 415–418
44. Ooi, S.K. *et al.* (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature* 448, 714–717
45. Umlauf, D. *et al.* (2004) Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat. Genet.* 36, 1296–1300
46. Terranova, R. *et al.* (2008) Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. *Dev. Cell* 15, 668–679
47. Wagschal, A. *et al.* (2008) G9a histone methyltransferase contributes to imprinting in the mouse placenta. *Mol. Cell. Biol.* 28, 1104–1113
48. Pandey, R.R. *et al.* (2008) Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell* 32, 232–246
49. Lewis, A. *et al.* (2004) Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat. Genet.* 36, 1291–1295
50. Mohammad, F. *et al.* (2008) Kcnq1ot1/Lit1 noncoding RNA mediates transcriptional silencing by targeting to the perinucleolar region. *Mol. Cell. Biol.* 28, 3713–3728
51. Nagano, T. *et al.* (2008) The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* 322, 1717–1720
52. Andergassen, D. *et al.* (2017) Mapping the mouse allelome reveals tissue-specific regulation of allelic expression. *eLife* 6, e25125
53. Sanli, I. *et al.* (2018) Meg3 non-coding RNA expression controls imprinting by preventing transcriptional upregulation in *cis*. *Cell Rep.* 23, 337–348
54. Santoro, F. *et al.* (2013) Imprinted Igf2r silencing depends on continuous Airn lncRNA expression and is not restricted to a developmental window. *Development* 140, 1184–1195
55. Mehta, S. *et al.* (2015) Transcription driven somatic DNA methylation within the imprinted Gnas cluster. *PLoS One* 10, e0117378
56. Greenberg, M.V. *et al.* (2017) Transient transcription in the early embryo sets an epigenetic state that programs postnatal growth. *Nat. Genet.* 49, 110–118
57. Ohhata, T. *et al.* (2015) Histone H3 lysine 36 trimethylation is established over the Xist promoter by antisense Tsix transcription and contributes to repressing Xist expression. *Mol. Cell. Biol.* 35, 3909–3920
58. Ohhata, T. *et al.* (2011) Lineage-specific function of the non-coding Tsix RNA for Xist repression and Xi reactivation in mice. *Genes Dev.* 25, 1702–1715
59. Furlan, G. *et al.* (2018) The Ftx noncoding locus controls X chromosome inactivation independently of its RNA products. *Mol. Cell* 70, 462–472
60. Jonkers, I. *et al.* (2009) RNF12 is an X-encoded dose-dependent activator of X chromosome inactivation. *Cell* 139, 999–1011
61. Abarrategui, I. and Krangel, M.S. (2006) Regulation of T cell receptor-alpha gene recombination by transcription. *Nat. Immunol.* 7, 1109–1115
62. Liu, Y. *et al.* (2007) A plant homeodomain in RAG-2 that binds hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. *Immunity* 27, 561–571
63. Matthews, A.G. *et al.* (2007) RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* 450, 1106–1110
64. Kumari, G. and Sen, R. (2015) Chromatin interactions in the control of immunoglobulin heavy chain gene assembly. *Adv. Immunol.* 128, 41–92
65. Outters, P. *et al.* (2015) Long-range control of V(D)J recombination & allelic exclusion: modeling views. *Adv. Immunol.* 128, 363–413
66. Puget, N. *et al.* (2015) Insertion of an imprinted insulator into the IgH locus reveals developmentally regulated, transcription-dependent control of V(D)J recombination. *Mol. Cell. Biol.* 35, 529–543
67. Levin-Klein, R. *et al.* (2017) Clonally stable Vkappa allelic choice instructs Igkappa repertoire. *Nat. Commun.* 8, 15575
68. Yancopoulos, G.D. and Alt, F.W. (1985) Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. *Cell* 40, 271–281
69. Bolland, D.J. *et al.* (2004) Antisense intergenic transcription in V(D)J recombination. *Nat. Immunol.* 5, 630–637
70. Stubbington, M.J. and Corcoran, A.E. (2013) Non-coding transcription and large-scale nuclear organisation of immunoglobulin recombination. *Curr. Opin. Genet. Dev.* 23, 81–88
71. Fuss, S.H. *et al.* (2007) Local and *cis* effects of the H element on expression of odorant receptor genes in mouse. *Cell* 130, 373–384
72. Monahan, K. *et al.* (2017) Cooperative interactions enable singular olfactory receptor expression in mouse olfactory neurons. *eLife* 6, e28620
73. Nishizumi, H. *et al.* (2007) Deletion of the core-H region in mice abolishes the expression of three proximal odorant receptor genes in *cis*. *Proc. Natl. Acad. Sci. U. S. A.* 104, 20067–20072
74. Monk, D. *et al.* (2009) Reciprocal imprinting of human GRB10 in placental trophoblast and brain: evolutionary conservation of reversed allelic expression. *Hum. Mol. Genet.* 18, 3066–3074
75. Sanz, L.A. *et al.* (2008) A mono-allelic bivalent chromatin domain controls tissue-specific imprinting at Grb10. *EMBO J.* 27, 2523–2532
76. Hellman, A. and Chess, A. (2007) Gene body-specific methylation on the active X chromosome. *Science* 315, 1141–1143
77. Mira-Bontenbal, H. and Gribnau, J. (2016) New Xist-interacting proteins in X-chromosome inactivation. *Curr. Biol.* 26, 1383
78. McHugh, C.A. *et al.* (2015) The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* 521, 232–236
79. Moindrot, B. *et al.* (2015) A pooled shRNA screen identifies Rbm15, Spen, and Wtap as factors required for Xist RNA-mediated silencing. *Cell Rep.* 12, 562–572

80. Chu, C. *et al.* (2015) Systematic discovery of Xist RNA binding proteins. *Cell* 161, 404–416
81. Monfort, A. *et al.* (2015) Identification of Spen as a crucial factor for Xist function through forward genetic screening in haploid embryonic stem cells. *Cell Rep.* 12, 554–561
82. Hasegawa, Y. *et al.* (2010) The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev. Cell* 19, 469–476
83. Pintacuda, G. *et al.* (2017) hnRNPK recruits PCGF3/5-PRC1 to the Xist RNA B-repeat to establish Polycomb-mediated chromosomal silencing. *Mol. Cell* 68, 955–969
84. Sarma, K. *et al.* (2014) ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. *Cell* 159, 869–883
85. Ratnakumar, K. *et al.* (2012) ATRX-mediated chromatin association of histone variant macroH2A1 regulates alpha-globin expression. *Genes Dev.* 26, 433–438
86. Sado, T. *et al.* (2000) X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. *Dev. Biol.* 225, 294–303
87. Xu, J. *et al.* (2017) Landscape of monoallelic DNA accessibility in mouse embryonic stem cells and neural progenitor cells. *Nat. Genet.* 49, 377–386
88. Selimyan, R. *et al.* (2013) Localized DNA demethylation at recombination intermediates during immunoglobulin heavy chain gene assembly. *PLoS Biol.* 11, e1001475
89. Subrahmanyam, R. *et al.* (2012) Localized epigenetic changes induced by DH recombination restricts recombinase to DJH junctions. *Nat. Immunol.* 13, 1205–1212
90. Goldmit, M. *et al.* (2005) Epigenetic ontogeny of the Igk locus during B cell development. *Nat. Immunol.* 6, 198–203
91. Wang, X. *et al.* (2017) Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. *Nat. Struct. Mol. Biol.* 24, 1028–1038
92. Xu, C.R. *et al.* (2008) Reciprocal patterns of methylation of H3K36 and H3K27 on proximal vs: distal IgVH genes are modulated by IL-7 and Pax5. *Proc. Natl. Acad. Sci. U. S. A.* 105, 8685–8690
93. Johnson, K. *et al.* (2004) B cell-specific loss of histone 3 lysine 9 methylation in the V(H) locus depends on Pax5. *Nat. Immunol.* 5, 853–861
94. Mandal, M. *et al.* (2011) Epigenetic repression of the Igk locus by STAT5-mediated recruitment of the histone methyltransferase Ezh2. *Nat. Immunol.* 12, 1212–1220
95. Magklara, A. *et al.* (2011) An epigenetic signature for monoallelic olfactory receptor expression. *Cell* 145, 555–570
96. Armelin-Correa, L.M. *et al.* (2014) Nuclear compartmentalization of odorant receptor genes. *Proc. Natl. Acad. Sci. U. S. A.* 111, 2782–2787
97. Renfree, M.B. *et al.* (2013) The origin and evolution of genomic imprinting and viviparity in mammals. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368, 20120151
98. Patten, M.M. *et al.* (2014) The evolution of genomic imprinting: theories, predictions and empirical tests. *Heredity* 113, 119–128
99. Zenk, F. *et al.* (2017) Germ line-inherited H3K27me3 restricts enhancer function during maternal-to-zygotic transition. *Science* 357, 212–216
100. Xu, D. *et al.* (2018) Polymorphic imprinting of SLC38A4 gene in bovine placenta. *Biochem. Genet.* Published online May 21, 2018. http://dx.doi.org/10.1007/s10528-018-9866-5
101. Auclair, G. *et al.* (2016) EHMT2 directs DNA methylation for efficient gene silencing in mouse embryos. *Genome Res.* 26, 192–202
102. Sado, T. *et al.* (2004) De novo DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation. *Development* 131, 975–982
103. Grant, J. *et al.* (2012) Rsx is a metatherian RNA with Xist-like properties in X-chromosome inactivation. *Nature* 487, 254–258
104. Thompson, C.B. (1995) New insights into V(D)J recombination and its role in the evolution of the immune system. *Immunity* 3, 531–539
105. Carmona, L.M. and Schatz, D.G. (2017) New insights into the evolutionary origins of the recombination-activating gene proteins and V(D)J recombination. *FEBS J.* 284, 1590–1605
106. Tunyaplin, C. and Knight, K.L. (1997) IgH gene rearrangements on the unexpressed allele in rabbit B cells. *J. Immunol.* 158, 4805–4811
107. Malecek, K. *et al.* (2008) Immunoglobulin heavy chain exclusion in the shark. *PLoS Biol.* 6, e157
108. Zhu, C. *et al.* (2011) The multiple shark Ig H chain genes rearrange and hypermutate autonomously. *J. Immunol.* 187, 2492–2501
109. Hinds, K.R. and Litman, G.W. (1986) Major reorganization of immunoglobulin VH segmental elements during vertebrate evolution. *Nature* 320, 546–549
110. Flajnik, M.F. (2014) Re-evaluation of the immunological big bang. *Curr. Biol.* 24, R1060–R1065
111. Rogozin, I.B. *et al.* (2007) Evolution and diversification of lamprey antigen receptors: evidence for involvement of an AID-APOBEC family cytosine deaminase. *Nat. Immunol.* 8, 647–656
112. Guo, P. *et al.* (2009) Dual nature of the adaptive immune system in lampreys. *Nature* 459, 796–801
113. Kishishita, N. *et al.* (2010) Regulation of antigen-receptor gene assembly in hagfish. *EMBO Rep.* 11, 126–132
114. Schulz, E.G. (2017) X-chromosome dosage as a modulator of pluripotency, signalling and differentiation? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 372, 20160366
115. Brady, B.L. *et al.* (2010) Antigen receptor allelic exclusion: an update and reappraisal. *J. Immunol.* 185, 3801–3808
116. Alt, F.W. *et al.* (2013) Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell* 152, 417–429
117. Lucas, J.S. *et al.* (2014) 3D trajectories adopted by coding and regulatory DNA elements: first-passage times for genomic interactions. *Cell* 158, 339–352
118. Mohammad, F. *et al.* (2012) Long noncoding RNA-mediated maintenance of DNA methylation and transcriptional gene silencing. *Development* 139, 2792–2803
119. Degner, J.F. *et al.* (2009) Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. *Bioinformatics* 25, 3207–3212
120. DeVeale, B. *et al.* (2012) Critical evaluation of imprinted gene expression by RNA-Seq: a new perspective. *PLoS Genet.* 8, e1002600
121. Proudhon, C. and Bourc'his, D. (2010) Identification and resolution of artifacts in the interpretation of imprinted gene expression. *Brief. Funct. Genomics* 9, 374–384
122. Schubeler, D. (2015) Function and information content of DNA methylation. *Nature* 517, 321–326
123. Gendrel, A.V. *et al.* (2014) Developmental dynamics and disease potential of random monoallelic gene expression. *Dev. Cell* 28, 366–380
124. Hirayama, T. and Yagi, T. (2017) Regulation of clustered protocadherin genes in individual neurons. *Semin. Cell Dev. Biol.* 69, 122–130
125. Mager, J. *et al.* (2003) Genome imprinting regulated by the mouse Polycomb group protein Eed. *Nat. Genet.* 33, 502–507
126. da Rocha, S.T. *et al.* (2014) Jarid2 is implicated in the initial Xist-induced targeting of PRC2 to the inactive X chromosome. *Mol. Cell* 53, 301–316
127. Brockdorff, N. (2017) Polycomb complexes in X chromosome inactivation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 372, 20170021
128. Su, I.H. *et al.* (2003) Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat. Immunol.* 4, 124–131

129. Almeida, M. *et al.* (2017) PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation. *Science* 356, 1081–1084

130. Busslinger, M. and Tarakhovsky, A. (2014) Epigenetic control of immunity. *Cold Spring Harbor Perspect. Biol.* 6, a019307

131. Girardot, M. *et al.* (2014) PRMT5-mediated histone H4 arginine-3 symmetrical dimethylation marks chromatin at G + C-rich regions of the mouse genome. *Nucleic Acids Res.* 42, 235–248

132. Keniry, A. *et al.* (2016) Setdb1-mediated H3K9 methylation is enriched on the inactive X and plays a role in its epigenetic silencing. *Epigenetics Chromatin* 9, 16

133. Thomas, L.R. *et al.* (2008) Functional analysis of histone methyltransferase g9a in B and T lymphocytes. *J. Immunol.* 181, 485–493

134. Stewart, K.R. *et al.* (2015) Dynamic changes in histone modifications precede de novo DNA methylation in oocytes. *Genes Dev.* 29, 2449–2462

135. Bourc'his, D. *et al.* (2001) Dnmt3L and the establishment of maternal genomic imprints. *Science* 294, 2536–2539

136. Hata, K. *et al.* (2002) Dnmt3L cooperates with the Dnmt3 family of *de novo* DNA methyltransferases to establish maternal imprints in mice. *Development* 129, 1983–1993

137. Gendrel, A.V. *et al.* (2013) Epigenetic functions of smchd1 repress gene clusters on the inactive X chromosome and on autosomes. *Mol. Cell. Biol.* 33, 3150–3165

138. Chen, K. *et al.* (2015) Genome-wide binding and mechanistic analyses of Smchd1-mediated epigenetic regulation. *Proc. Natl. Acad. Sci. U. S. A.* 112, E3535–E3544

139. Voon, H.P. *et al.* (2015) ATRX plays a key role in maintaining silencing at interstitial heterochromatic loci and imprinted Genes. *Cell Rep.* 11, 405–418

140. Franco, M.M. *et al.* (2014) The role of CCCTC-binding factor (CTCF) in genomic imprinting, development, and reproduction. *Biol. Reprod.* 91, 125

141. Guo, C. *et al.* (2011) CTCF-binding elements mediate control of V(D)J recombination. *Nature* 477, 424–430

142. Monahan, K. *et al.* (2012) Role of CCCTC binding factor (CTCF) and cohesin in the generation of single-cell diversity of protocadherin-alpha gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 109, 9125–9130