doi:10.1038/nature19468

DEVELOPMENTAL BIOLOGY

Panoramic views of the early epigenome

Four studies detail changes in how DNA is wrapped around histone proteins and in molecular modifications to histones that occur after fertilization. The results shed light on the early regulation of gene expression.

JUAN M. VAQUERIZAS
& MARIA-ELENA TORRES-PADILLA

The beginning of life, marked by the fertilization of an egg by a sperm, has been a major focus of research for decades. When the differentiated cells give rise to an embryo, there is a dramatic reprogramming of the epigenome — the collection of molecular modifications to DNA and associated histone proteins that alter gene expression without changing DNA sequence. But details of the genomic regions affected by epigenetic reprogramming have been lacking. Four papers in *Nature* (three now online[1–3] and one published in June[4]) now reveal striking and unexpected features of this process in the developing oocyte (the unfertilized egg) and the early mouse embryo.

Gene expression changes drastically during the generation of oocytes and sperm in mammals, and halts completely by the time these cells are fully mature. In mice, expression resumes shortly after fertilization, with a minor wave of gene activation (dubbed zygotic genome activation; ZGA). A second, major wave of ZGA occurs at the late two-cell stage, marking deployment of the developmental gene-expression program. Four divisions later, a cell population called the inner cell mass develops. These cells will form the embryo proper, and can be extracted to derive embryonic stem (ES) cells *in vitro*. The epigenomic state of ES cells has been thoroughly investigated, but that of earlier developmental stages has remained elusive, mostly owing to the minimal amount of material available for study.

The four current studies analysed the regions of the genome with which three histone modifications are associated in sperm and oocytes and in early mouse embryos. The authors adapted techniques to allow the analysis of just a few cells. First, Liu *et al.*[1], Dahl *et al.*[2] and Zhang *et al.*[3] studied modification of the amino-acid residue lysine 4 (K4) on histone H3 by three methyl groups (a modification referred to as H3K4me3). Second, Dahl *et al.* and Wu *et al.* [4] examined modification of lysine 27 (K27) by an acetyl group (H3K27ac). Third, Liu *et al.* and Wu *et al.* analysed trimethylation of K27 (H3K27me3). The studies differed in the number of cells analysed and how the DNA and associated proteins (collectively called chromatin) were treated before analysis, but the groups all reached similar conclusions.

In ES cells and mature cell types, H3K4me3 is primarily clustered around small DNA regions at which gene transcription begins, and is associated with gene activity. One of the most striking findings of the current papers is that, in oocytes, H3K4me3 is enriched at low levels across large genomic regions, spanning more than 10 kilobases, and is mostly distant from transcription start sites. This pattern of 'non-canonical' H3K4me3 persists in the fertilized oocyte and in embryos at the early two-cell stage (Fig. 1).

Dahl *et al.* and Zhang *et al.* delineate two groups of genes associated with non-canonical H3K4me3 in the mature oocyte[2,3]. First, non-canonical H3K4me3 is found close to genes that are expressed during oocyte growth — in agreement with previous work[5] suggesting that dynamic remodelling of H3K4 methylation occurs as the oocyte matures, and is coupled to changes in gene expression and in DNA methylation, which represses transcription. Second, the modification is found associated with genes expressed during major ZGA. Therefore, this atypical modification seems to provide an epigenetic memory of the transcriptional state of the oocyte that is inherited by the developing embryo.

More unexpected findings are also reported. First, Zhang and colleagues observed non-canonical H3K4me3 in regions enriched in certain repetitive sequences, some of which are highly active during early embryonic development[3]. H3K4me3 had previously been reported[6] to associate with only one such repetitive element in embryos, LINE-1. This is of particular interest because Wu *et al.* find[4] that large domains of DNA are accessible to transcription-factor binding before major ZGA. This accessibility is associated with the transcription of specific families of repetitive elements and nearby genes, highlighting the regulatory potential of repetitive elements.

Second, Zhang *et al.* found that removing methyl groups from H3K4me3 in oocytes resulted in increased — albeit

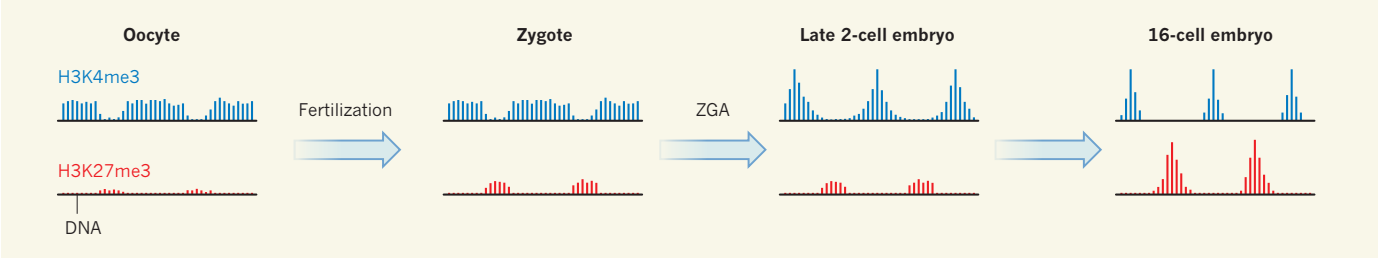

Figure 1 | Measuring methylation during early development. Four studies[1–4] analyse regions of the mouse genome that are associated with histone proteins that have been modified at amino-acid residues lysine 4 or lysine 27 by the addition of methyl groups (modifications called H3K4me3 or H3K27me3, respectively). In this simple schematic, the heights of vertical lines represent the level of modification associated with sequential regions of DNA. There is no gene transcription in mature, unfertilized eggs (oocytes), and H3K4me3 is distributed in broad regions across the genome, which is atypical for this modification. These domains are mostly maintained following fertilization (at the zygote stage) until the late two-cell stage, when a process called major zygotic genome activation (ZGA) leads to a wave of gene expression. From then on, the broad signature is replaced by narrow H3K4me3 regions associated with gene activation at sites at which transcription is initiated. By contrast, H3K27me3, which is associated with gene repression, is lowest in the oocyte and at early stages, and becomes increasingly abundant as development proceeds, showing a mutually exclusive distribution with narrow H3K4me3 regions until the 16-cell stage.

aberrant — transcriptional activation[3], and Dahl *et al.* reached similar conclusions using a complementary approach[2]. This surprising finding implies that non-canonical H3K4me3 can have a role in transcriptional silencing. It remains to be determined whether this effect is direct, or whether the presence of H3K4me3 somehow sends signals that lead to alterations in the levels of other histone modifications, regulating silencing indirectly.

The papers reveal that the histone demethylase enzyme KDM5B is crucial in limiting the genomic distribution of H3K4me3 during oocyte maturation and ZGA. And Liu *et al.* show[1] that, in the developing embryo, H3K4me3 domains become broader following the loss of KDM5B. Thus, keeping the levels of H3K4me3 in check seems to be essential for correctly establishing and deploying early-embryonic gene-expression programs.

Zhang *et al.* and Wu *et al.* also examine differences in the histone modifications inherited by the embryo from the father and the mother[3,4]. They find that parental differences in H3K4me3 distribution are retained in the two sets of chromosomes in the early embryo, supporting the idea that some epigenetic information is inherited.

Turning to other histone modifications, Liu and colleagues[1] compared changes in H3K27me3, which is associated with gene repression, with those of the activation-associated H3K4me3. In contrast to the other studies, the group focused on analysing these modifications only in regions close to ZGA genes that harbour the typical, 'canonical', high-level H3K4me3 signal around transcription start sites. They found that levels of canonical H3K4me3 increased from the late two-cell stage onward. This differs from the non-canonical H3K4me3 sites, where levels drop off after the two-cell stage.

Liu *et al.* found that the number of regions that contain canonical H3K4me3 but not H3K27me3 increased sharply at the late two-cell stage. By contrast, the number of H3K27me3-only regions increased gradually (Fig. 1). This probably reflects different dynamics, and hence different mechanisms, in establishing these two epigenetic marks. H3K4me3 and H3K27me3 are mutually exclusive up to the 16-cell stage, possibly because of the low levels of H3K27me3. By contrast, ES cells contain many domains marked by both such histone modifications. Thus, bivalent domains of modification are established at later stages of development. By having both 'active' and 'repressive' modifications, bivalent domains are thought to be crucial for the efficient expression of lineage-specific developmental programs as cells start to differentiate into mature lineages.

Finally, Dahl *et al.* found stage-specific H3K27ac domains[2], which are presumed to activate the expression of nearby genes. H3K27ac domains tended to be near genes associated with ZGA, and the authors used the domains to identify transcription factors that potentially bind to these nearby genes to regulate early, stage-specific developmental programs. Although this will certainly constitute a powerful resource, some of the factors identified differ from those documented in a paper published in June[7]. Further work will be required to determine the specific details of the mechanisms by which these transcription factors drive development.

Overall, the studies demonstrate a drastic epigenetic remodelling process in oocytes and sperm, and at early stages of embryonic development. This hints at mechanisms by which histone modifications are passed between two generations, playing a vital part in the activation of the newly formed genome. Further work will be necessary to characterize the precise molecular mechanisms that govern these transitions. ■

Juan M. Vaquerizas *is at the Max Planck Institute for Molecular Biomedicine, 48149 Münster, Germany.* **Maria-Elena Torres-Padilla** *is at the Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM U964, Strasbourg 67404, France, and at the Institute of Epigenetics and Stem Cells, Helmholtz Zentrum München, Munich, Germany.*
e-mail: torres-padilla@helmholtz-muenchen.de

1. Liu, X. *et al. Nature* http://dx.doi.org/10.1038/nature19362 (2016).
2. Dahl, J. A. *et al. Nature* http://dx.doi.org/10/1038/nature19360 (2016).
3. Zhang, B. *et al. Nature* http://dx.doi.org/10.1038/nature19361 (2016).
4. Wu. J. *et al. Nature* **534,** 652–657 (2016).
5. Stewart, K. R. *et al. Genes Dev.* **29,** 2449–2462 (2015).
6. Fadloun, A. *et al. Nature Struct. Mol. Biol.* **20,** 332–338 (2013).
7. Lu, F. *et al. Cell* **165,** 1375–1388 (2016).