Expression analysis suggests that DNMT3L is required for oocyte de novo DNA methylation only in *Muridae* and *Cricetidae* rodents

Lirik Behluli[1†], Alyssa M. Fontanilla[1†], Laura Andessner-Angleitner[1†], Nikolas Tolar[1], Julia M. Molina[1,2] and Lenka Gahurova[1*]

Abstract

Background During early mammalian development, DNA methylation undergoes two waves of reprogramming, enabling transitions between somatic cells, oocyte and embryo. The first wave of de novo DNA methylation establishment occurs in oocytes. Its molecular mechanisms have been studied in mouse, a classical mammalian model. Current model describes DNA methyltransferase 3A (DNMT3A) and its cofactor DNMT3L as two essential factors for oocyte DNA methylation—the ablation of either leads to nearly complete abrogation of DNA methylation. However, *DNMT3L* is not expressed in human oocytes, suggesting that the mechanism uncovered in mouse is not universal across mammals.

Results We analysed available RNA-seq data sets from oocytes of multiple mammals, including our novel data sets of several rodent species, and revealed that *Dnmt3l* is expressed only in the oocytes of mouse, rat and golden hamster, and at a low level in guinea pigs. We identified a specific promoter sequence recognised by an oocyte transcription factor complex associated with strong *Dnmt3l* activity and demonstrated that it emerged in the rodent clade *Eumuroida*, comprising the families *Muridae* (mice, rats, gerbils) and *Cricetidae* (hamsters). In addition, an evolutionarily novel promoter emerged in the guinea pig, driving weak *Dnmt3l* expression, likely without functional relevance. Therefore, *Dnmt3l* is expressed and consequently plays a role in oocyte de novo DNA methylation only in a small number of rodent species, instead of being an essential pan-mammalian factor. In contrast to somatic cells, where catalytically inactive DNMT3B interacts with DNMT3A, forming a heterotetramer, we did not find evidence for the expression of such inactive *Dnmt3b* isoforms in the oocytes of the tested species.

Conclusions The analysis of RNA-seq data and genomic sequences revealed that DNMT3L is likely to play a role in oocytes de novo DNA methylation only in mice, rats, gerbils and hamsters. The mechanism governing de novo DNA methylation in the oocytes of most mammalian species, including humans, occurs through a yet unknown mechanism that differs from the current model discovered in mouse.

†Lirik Behluli, Alyssa M. Fontanilla and Laura Andessner-Angleitner have contributed equally to this work.

*Correspondence:
Lenka Gahurova
lveselovska@prf.jcu.cz
[1] Department of Molecular Biology and Genetics, Faculty of Science, University of South Bohemia, Branisovska 1760, 37005 Ceske Budejovice, Czech Republic

[2] Department of Biological Sciences, Faculty of Sciences and Languages, São Paulo State University "Júlio de Mesquita Filho"- UNESP, Assis, São Paulo, Brazil

Behluli *et al. Epigenetics & Chromatin*   (2023) 16:43

Page 2 of 15

Introduction

DNA methylation is an epigenetic mark contributing to the determination and maintenance of cell identity through regulation of gene expression patterns and chromatin states. It is substantially reprogrammed during mammalian development, enabling the transition from differentiating somatic cells into germline, and from highly specialised oocyte and sperm into totipotent zygote, followed by the differentiation of individual cell lineages. This reprogramming comprises two waves of nearly complete DNA demethylation followed by de novo DNA methylation establishment [48].

In mammalian oocytes, DNA methylation is established on a largely blank slate during oocyte growth, with the exception of a number of repetitive elements retaining their methylation throughout oogenesis [46]. Current widely accepted model describes DNA methyltransferase 3A (DNMT3A) as the main de novo DNA methylation enzyme in oocytes, catalysing the transfer of the methyl group from S-adenosyl methionine onto unmethylated DNA. DNMT3L acts as an essential co-factor without catalytic activity [5, 25]. Two copies each of DNMT3A and DNMT3L assemble into a heterotetrameric complex [23], and deletion of either of them leads to nearly complete abrogation of oocyte de novo DNA methylation [27, 44, 47]. Despite its oocyte expression, another de novo DNA methyltransferase, DNMT3B, was shown to be dispensable for oocyte de novo DNA methylation [44], while *Muroidea* rodents-specific DNMT3C appears to be restricted to the male germline [3].

Recruitment of the DNMT3A–DNMT3L complex to target genomic regions subject to de novo DNA methylation appears to be exerted by specific post-translational modifications of histone tails [48]. DNMT3A and DNMT3L share the ATRX–DNMT3–DNMT3L (ADD) domain interacting with the N-terminal histone 3 (H3) tail. This interaction is inhibited by H3 lysine 4 di- and tri-methylation (H3K4me2/3) [32, 34, 68]. In contrast to DNMT3A, DNMT3L lacks a large part of the catalytic methyltransferase domain and a Pro–Trp–Trp–Pro motif (PWWP) domain that recognises H3K36me2/3. In vitro studies revealed that the DNMT3A PWWP domain guides DNA methylation to regions with H3K36me3 [14], a mark associated with gene bodies of actively transcribed genes. It was believed that H3K36me3 plays the same DNA methylation guidance role in oocytes, especially considering that the oocyte DNA methylation landscape largely matches the transcribed regions, while methylation is absent from gene bodies of silent genes or from intergenic regions [58]. Moreover, simultaneous depletion of H3K36me2/3 leads to global DNA hypomethylation [65]. However, it was demonstrated that the DNMT3A PWWP domain instead restricts DNA

methylation to H3K36me2/3 regions. When the association between DNMT3A and H3K36me2/3 is abolished by D329A substitution in the PWWP domain, oocytes are hypermethylated [26]. Instead, a recent study revealed that amino acid substitutions in the ADD domain of DNMT3A lead to mosaic loss of mouse oocyte DNA methylation. The ADD domain, therefore, likely plays a role in the efficiency and processivity of the DNMT3A–DNMT3L de novo DNA methylation complex [54]. These results suggest that oocyte de novo DNA methylation targeting actively transcribed regions involves additional mechanisms.

Oocyte de novo DNA methylation establishment has been studied in mouse, a classical mammalian model. DNMT3L has been described as an essential factor in oocyte DNA methylation establishment—initially only at the imprinted loci [5], and later, once the development of technologies allowed profiling the whole DNA methylation landscape, at the genome-wide level [27, 44, 47]. Despite *DNMT3L* being transcriptionally silent in human oocytes [36], the DNMT3L-based model described in mouse continues to be accepted as a general mammalian model [8, 11]. Nevertheless, some studies highlight the enigmatic status of DNMT3L due to the lack of expression in human oocytes [13, 42].

Here, we show that the key role of DNMT3L in oocyte DNA methylation establishment is restricted to a specific rodent lineage *Eumuroida* comprising mice, rats, gerbils and hamsters. Its oocyte expression in these species is due to the emergence of an oocyte-specific upstream promoter within the third intron of the upstream gene *Aire*. In other rodent and mammalian lineages, the absence of DNMT3L may have consequences for the interplay between histone modifications and the DNA methylation landscape in oocytes.

Results

Dnmt3l is transcriptionally silent in the oocytes of rodent naked mole-rat and non-rodent species

To analyse oocyte *Dnmt3l* expression beyond mouse and human, we quantified available RNA-seq data of rat, golden hamster, cow, pig, macaque rhesus and our novel data sets of naked mole-rat oocytes. We demonstrated that while *Dnmt3a* and *Dnmt3b* genes are expressed in all species (except golden hamster, where *Dnmt3b* is absent from gene annotation of the MesAur1.0 genome), the expression of *Dnmt3l* is restricted to mouse, rat and golden hamster (Fig. 1a). To confirm that the absence of *Dnmt3l* expression in the naked mole-rat, another rodent species, is not just due to incorrect *Dnmt3l* annotation, we performed a BLAST search of mouse, human and naked mole-rat *Dnmt3l* sequences against the naked mole-rat genome. The existing *Dnmt3l* annotation is

Behluli *et al. Epigenetics & Chromatin*   (2023) 16:43

Page 3 of 15

the only region that is significantly similar to mouse and human *Dnmt3l*, and no other *Dnmt3l*-like sequence is present in the naked mole-rat genome. Furthermore, we identified that the largest window of sequence identity between naked mole-rat *Dnmt3l* and other *Dnmt3* family genes is 13 bp, confirming that hypothetical reads originating from *Dnmt3l* would not incorrectly map to *Dnmt3a* or *Dnmt3b*.

In mouse, oocyte *Dnmt3l* transcription is initiated from an annotated oocyte-specific promoter localised upstream of the canonical promoter within the third intron of the *Aire* gene encoded on the opposite DNA strand [45] (Fig. 1b). There is an additional isoform of *Dnmt3l* with a transcriptional start site (TSS) within the fourth *Aire* exon and intron; however, it consists of only two exons and lacks a defined coding sequence (ENS-MUST00000135316.2 in the Ensembl genome browser). Visual inspection of mouse, rat and golden hamster oocyte RNA-seq data revealed that both isoforms are expressed in mouse and rat oocytes, while only the long protein-coding variant is expressed in golden hamster (although the latter case might be an artefact caused by lower sequencing depth and higher 3' end bias in golden hamster data). The mouse annotated TSS of the protein-coding oocyte *Dnmt3l* variant is at an adenosine (A) with a CT-rich sequence motif immediately upstream, a GA-rich motif immediately downstream, and a TTCTTAA sequence motif 32 bp upstream of the TSS (Fig. 1b). These sequences are conserved in rat and golden hamster (Fig. 1c) and match a published binding site of TATA box-binding protein-like 2 (TBPL2) assembled into a Transcription factor IIA (TFIIA) complex (Fig. 1d), suggesting that this complex is likely to drive oocyte *Dnmt3l* expression. The TBPL2/TFIIA complex has been previously shown to regulate transcription from oocyte-specific TSSs because of its specific TSS recognition code [66]. Quantification of *Dnmt3l* expression and its oocyte TSS activity using published RNA-seq and super-low input carrier–CAGE (SLIC–CAGE) data sets

[66], respectively, revealed that TBPL2 knock-out leads to significant downregulation of mouse oocyte *Dnmt3l* (Fig. 1e and Additional file 1: Fig. S1a for SLIC–CAGE results, while the results of RNA-seq can be found in [66]). It should be noted that SLIC–CAGE data confirm that both full length oocyte *Dnmt3l* and short noncoding transcript are expressed in mouse oocytes, but only the TSS of full-length *Dnmt3l* appears to be substantially regulated by the TBPL2/TFIIA complex, as it is statistically significantly downregulated in the TBPL2 knock-out SLIC–CAGE data set. Although the *TBPL2* gene and its oocyte expression are conserved in all tested mammalian species with oocyte RNA-seq data sets, the *Dnmt3l* oocyte promoter sequence motif is not conserved in species, where *Dnmt3l* is transcriptionally silent in oocytes (Fig. 1c).

To exclude the possibility of Dnmt3l expression from the identified promoter in other mammalian lineages, we analysed sequences of *Aire* intron equivalent to mouse intron 3 of additional mammalian species (10 primates, 10 Laurasiatheria species, 2 Afrotheria species and 1 Xenarthra species) with genome sequences available in the Ensembl genome database and with an annotated *Aire* gene (Additional file 6: Table S1), and confirmed the absence of TTCTTAA or a similar promoter sequence motif, as well as the TSS sequence. These results associate the evolutionary emergence of a sequence recognised by an oocyte TF complex with *Dnmt3l* expression in the oocytes, employing the oocyte-specific promoter.

Oocyte *Dnmt3l* expression is restricted to *Muridae* and *Cricetidae* rodents

To obtain more insights into *Dnmt3l* expression across rodents, we downloaded *Aire* intron sequences of all rodents available in the Ensembl genome database and of a species of tree shrew, which belongs to *Scandentia*, a sister clade to glires comprising rodents and lagomorphs (Additional file 6: Table S2). Based on the manual analysis of exon and intron sequences, we

(See figure on next page.)
Fig. 1 Oocyte *Dnmt3l* expression and promoter sequence. **a** Expression of *Dnmt3* genes across mammalian oocytes using available RNA-seq data sets. RPKM = reads per kilobase of transcript per million reads in the library (in paired-end libraries, fragments instead of reads were quantified). Dots represent individual data points (biological replicates), error bars show standard deviation. **b** Ensembl genome browser screenshot of the mouse *Dnmt3l* gene and its promoter region in single bp resolution. **c** Screenshot of reverse complement of *Aire* intron 3 multiple sequence alignment (only the sequence alignment around TSS and presumable TBPL2/TFIIA binding site is shown) in mammals with available oocyte RNA-seq data sets. The *Dnmt3l* oocyte TSS is highlighted in green, the upstream CT-rich region and downstream GA-rich region surrounding the TSS are highlighted in grey and the presumable TBPL2/TFIIA binding site is highlighted in yellow. Ro rodents, Pr primates, Un ungulates. **d** Visualisation of the TBPL2/TFIIA consensus binding site (adapted from [66] and matching sequences identified in mouse oocyte *Dnmt3l* TSS. Top sequence (I.)) represents the sequence with exactly matching positions upstream of TSS, bottom sequence (II.)) represents best fit for the binding site positioned 3 bp upstream of the exact match. The numbers indicate the position in the mouse sequence respective to the *Dnmt3l* TSS. **e** Screenshot from the program Seqmonk showing SLIC–CAGE read quantification of the first 200 bp from mouse *Dnmt3l* TSSs in control and TBPL2 knock-out oocytes [66]. *RPM* reads per million reads in the library
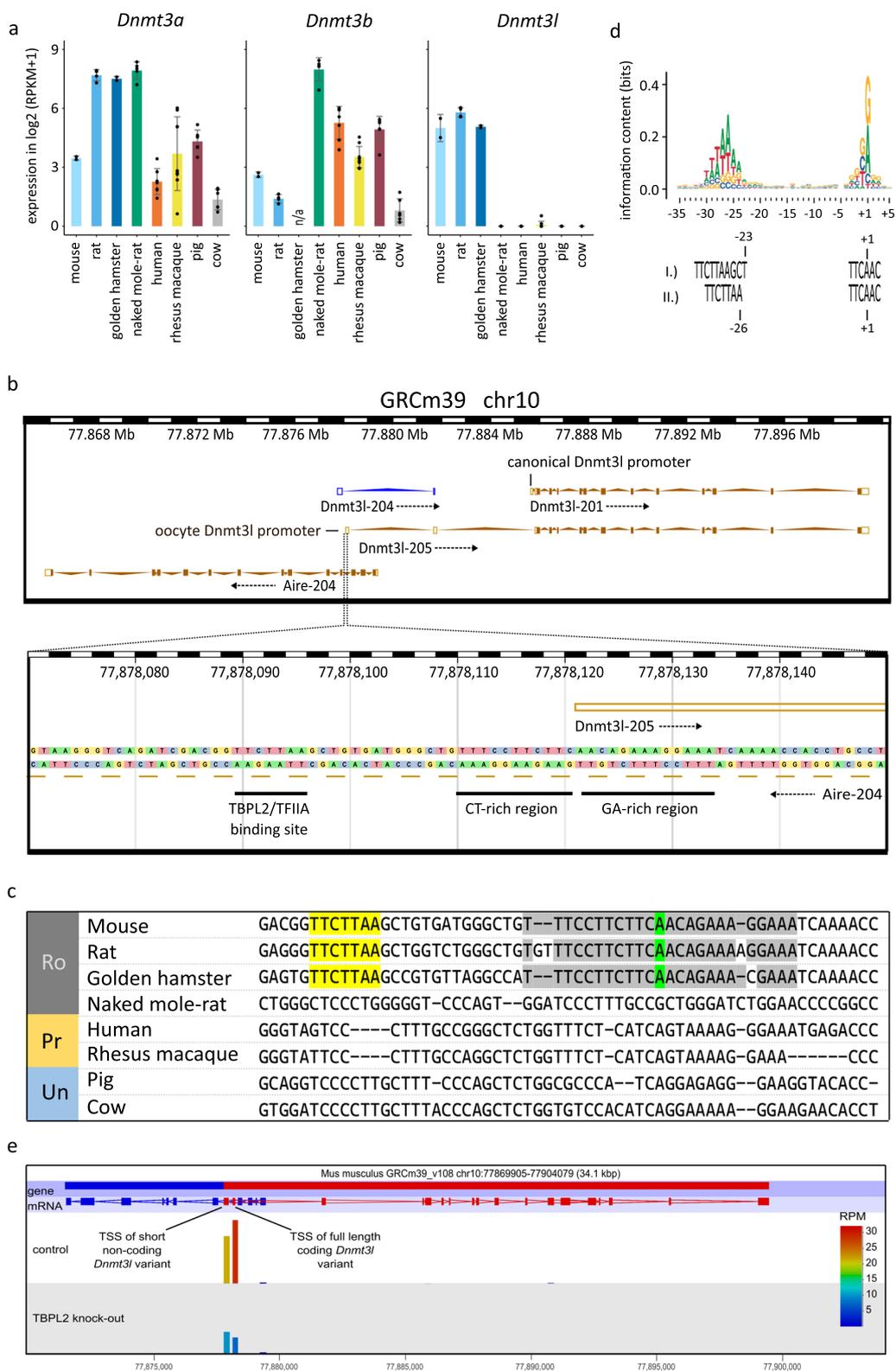
Fig. 1 (See legend on previous page.)

identified an intron corresponding to mouse intron 3 in each species except kangaroo rat. Moreover, we excluded guinea pig from the analysis, as the sequence of the corresponding intron is not determined at the nucleotide level (unknown nucleotides in the sequence). A multiple sequence alignment (Additional file 5) of these intron sequences revealed that the TTC TTAA promoter sequence motif is present only in the *Eumuroida* rodent clade, comprising families *Muridae* (mice, rats, gerbils) and *Cricetidae* (hamsters), with the

exception of the prairie vole, where a larger region surrounding the promoter sequence was deleted (Fig. 2a). The TSS, including upstream CT-rich and downstream GA-rich motif, is also conserved in the prairie vole (Additional file 1: Fig. S1b). In the *Spalacidae* family (exemplified by blind mole rat), a sister group to *Eumuroida* rodents within superfamily *Muroidea*, as well as in the family *Sciuridae* (squirrels), predicted TBPL2/TFIIA binding motif is conserved to some extent, although with multiple mismatches. Region
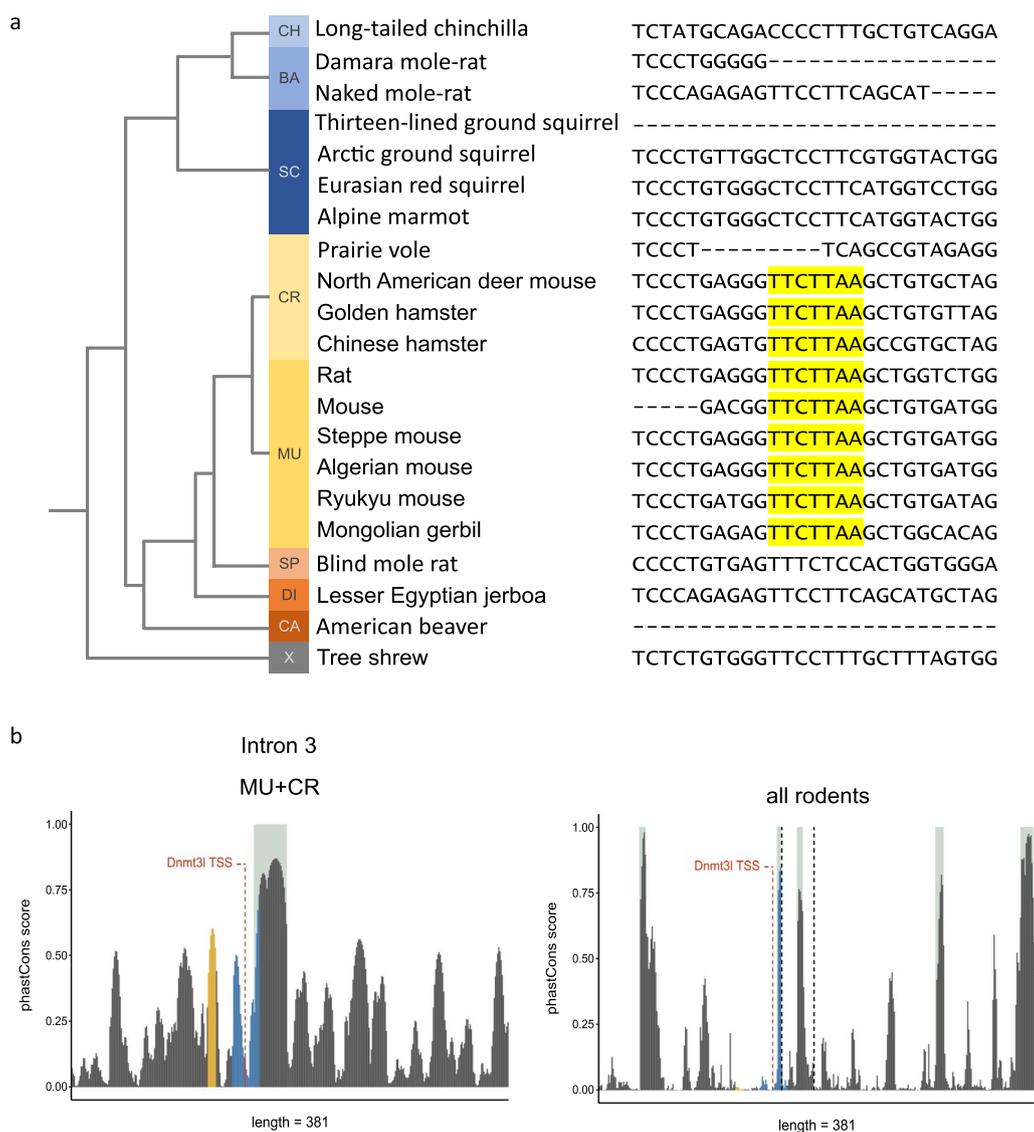

Fig. 2 *Dnmt3l* oocyte promoter sequence conservation in rodents. **a** Screenshot of a reverse complement of *Aire* intron 3 multiple sequence alignment with highlighted conserved promoter sequences in rodents with available sequences in Ensembl genome database. CH *Chinchillidae*, BA *Bathyergidae*, SC *Sciuridae*, CR *Cricetidae*, MU *Muridae*, SP *Spalacidae*, DI *Dipodidae*, CA *Castoridae*, X *Scandentia*. **b** PhastCons scores across mouse *Aire* intron 3 in *Muridae* and *Cricetidae* (except prairie vole) rodents on the left and all rodents with available sequence in the Ensembl genome database on the right. Regions highlighted in light grey indicate highly conserved regions identified by PhastCons. The presumed TBPL2/TFIIA binding motif is highlighted in yellow and CT-rich and GA-rich regions surrounding the TSS are highlighted in blue

Behluli *et al. Epigenetics & Chromatin*   (2023) 16:43

Page 6 of 15

surrounding potential TSS is conserved only partially. Moreover, in blind mole rat, there is a relatively large insertion (> 100 bp) between predicted binding site and region homologous to TSS. These sequences are probably not compatible with *Dnmt3l* oocyte transcription. For example, a similar sequence homologous to predicted TBPL2/TFIIA binding motif in naked mole-rat does not drive oocyte *Dnmt3l* expression. Nevertheless, this suggests that the ancestral sequence in rodents or a subgroup of rodents was favourable for the emergence of sequence motif recognised by an oocyte TF complex, driving oocyte *Dnmt3l* transcription.

To confirm that the oocyte *Dnmt3l* promoter region (particularly the presumable TBPL2/TFIIA binding motif and CT-rich and GA-rich regions upstream and downstream of the TSS, respectively) became evolutionarily conserved after its emergence, we analysed local conservation scores across the first five *Aire* introns in rodents using PhastCons [37]. We demonstrated that the presumable TBPL2/TFIIA binding motif forms a relatively conserved peak in *Muridae* and *Cricetidae* rodents with oocyte *Dnmt3l* expression, while there is no conservation of this region if all rodents are considered. The conservation is similar for the CT-rich region upstream of *Dnmt3l* oocyte TSS. In contrast, the GA-rich sequence downstream of the TSS is universally conserved across rodents (Fig. 2b). However, intron 3 and other *Aire* introns appear to contain other regions with similar or higher conservation scores, suggesting that the *Dnmt3l* promoter region is not unique in its high conservation (Additional file 2: Fig. S2).

Evolutionarily novel oocyte *Dnmt3l* promoter in guinea pigs drives weak oocyte expression

To further confirm the lack of oocyte *Dnmt3l* expression in rodents outside the *Muridae* and *Cricetidae* families, we performed RNA-seq on oocytes of guinea pig and coruro. Both species belong to the *Caviomorpha* lineage of *Hystricognathi* rodents, while naked mole-rat belongs to the *Phiomorpha* lineage. Guinea pig is representative of the *Caviidae* family within the *Cavioidea* superfamily, while coruro, closely related to degu, is representative of the *Octodontidae* family within the *Octodontoidea* superfamily. Because coruro genome assembly is not available, we mapped its data to degu *Dnmt3a*, *Dnmt3b* and *Dnmt3l* sequences with less stringent mapping parameters. All three genes appear to be under similar evolutionary pressures based on dN/dS ratios according to the best fitting model for each gene (Additional file 6: Table S3) [1], confirming that coruro reads should map to degu sequences with comparable probabilities. We showed relatively high expression of coruro *Dnmt3a* and *Dnmt3b*, while no reads were mapped to *Dnmt3l* (Fig. 3a). In contrast, mapping of guinea pig reads revealed very weak oocyte *Dnmt3l* expression (Fig. 3a) that was confirmed by qPCR using primers specific for *Dnmt3l* mRNA comprising exons 8–10 (Fig. 3b) and sequence verification of the PCR product.

According to the Ensembl annotation for the guinea pig Cavpor3.0 genome, the *Dnmt3l* TSS is located closely
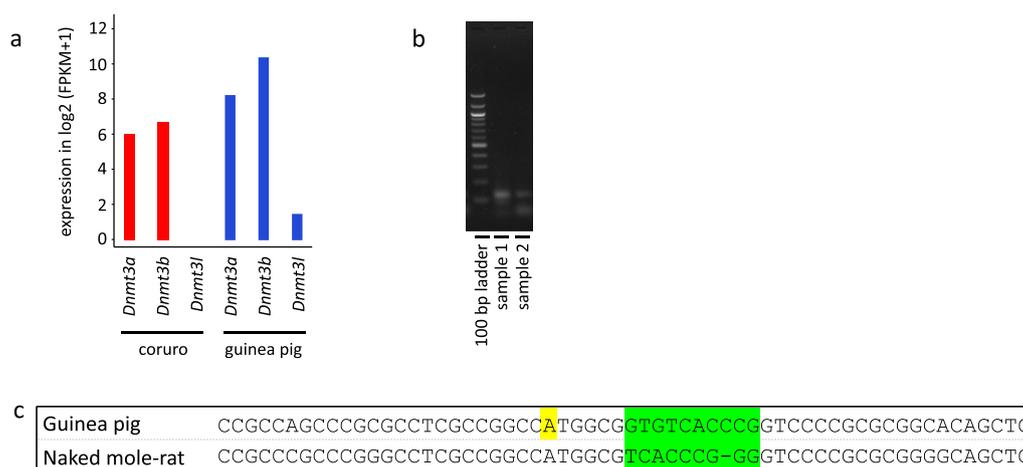

Fig. 3 *Dnmt3l* oocyte expression in guinea pig and coruro. **a** Expression of *Dnmt3* genes in coruro and guinea pig oocytes (single replicates). FPKM = fragments per kilobase of transcript per million fragments in the library. **b** PCR with reverse transcribed RNA from guinea pig oocytes amplifying region of exons 8–10 of the annotated *Dnmt3l* gene **c** Sequence alignment screenshot of a region surrounding Aire TSS in guinea pig and naked mole rat. The guinea pig *Dnmt3l* TSS is highlighted in yellow, the low conservation region in the otherwise highly conserved sequence is highlighted in green

Behluli *et al. Epigenetics & Chromatin*   (2023) 16:43

Page 7 of 15

upstream of the *Aire* TSS encoded on the opposite DNA strand. This appears unique across rodents and other mammals, as the canonical (embryonic) *Dnmt3l* TSS is located much more upstream of the *Aire* TSS in other species, while the TSS of the oocyte isoform is within the third *Aire* intron. As we mentioned previously, a large part of the corresponding guinea pig *Aire* intron is not deciphered at the nucleotide level. We wondered whether the oocyte *Dnmt3l* transcript could potentially start within this intron; however, sequencing of this region revealed that the oocyte promoter sequence is not conserved in guinea pig (Additional file 3: Fig. S3). To identify what could be driving weak oocyte *Dnmt3l* expression in guinea pig, we compared sequences ± 250 bp around annotated *Aire* TSS between guinea pig and naked mole-rat. The whole sequence was largely conserved, including the immediate region around the annotated *Dnmt3l* TSS in guinea pig, with the exception of a 9 bp region located 6 bp downstream, suggesting that this sequence could potentially regulate *Dnmt3l* expression in guinea pig (Fig. 3c). To confirm that this evolutionarily novel *Dnmt3l* promoter drives weak transcription in oocytes, we manually inspected RNA-seq and found out that no read maps to the annotated first exon. Moreover, we designed two primer pairs with forward primers specific for exon 1, but neither reaction produced specific products (data not shown). Therefore, we cannot confirm that weak guinea pig oocyte *Dnmt3l* expression initiates at the annotated TSS.

DNMT3L does not appear to play a substantial role in retrotransposon silencing

Due to the restricted rodent lineage specificity of both the *Dnmt3c* gene and oocyte *Dnmt3l* isoform, it is tempting to speculate about the similar functional role of proteins encoded by these genes in male and female germline, respectively. The *Muroidea* rodents-specific *Dnmt3c* gene plays a role in suppressing evolutionary young retrotransposons in the male germline and shows convergence with elements targeted by the piRNA pathway [3]. This gene is evolutionarily older than the oocyte *Dnmt3l* isoform, as it was also detected in the *Spalacidae* family [3], while we show that the oocyte *Dnmt3l* TSS appeared after splitting of *Muroidea* rodents into *Spalacidae* and *Eumuroida* lineages.

To obtain more insights into the potential association between retrotransposon activity and the emergence of oocyte *Dnmt3l* expression, we first compared the genomic content of all retrotransposons and evolutionarily young (and, therefore, still potentially active, or recently active) retrotransposons as estimated by Osmanski et al. [33] in rodents with and without predicted oocyte *Dnmt3l* expression. We detected significantly fewer SINE elements and significantly more LTR elements in species with oocyte *Dnmt3l*. We also noticed that species with oocyte *Dnmt3l* tend to have more evolutionarily young retrotransposons, especially of LINE and LTR classes (although this difference is not statistically significant) (Fig. 4a). However, there is no clear trend visible when visualised in the phylogenetic context (Fig. 4b). On the other hand, no such trend is associated with the emergence of *Dnmt3c* either (Fig. 4b).

In addition, we analysed a published RNA-seq data set from *Dnmt3l* knock-out (*Dnmt3l*-ko) fully grown mouse oocytes [44] to determine whether abrogation of DNMT3L-dependent DNA methylation leads to transcriptional reactivation of specific retrotransposons compared to control oocytes. No retrotransposon class was upregulated in *Dnmt3l*-ko oocytes (Additional file 4: Fig. S4a). The analysis of individual TE subfamilies using TEtranscripts [24] identified eight that are differentially expressed with probability 0.95 or higher (Additional file 4: Fig. S4b). However, seven of these are downregulated in *Dnmt3l*-ko oocytes. More detailed analysis of the single upregulated subfamily, ERVB4_3-LTR_MM, revealed that most of the reads originate in a single insertion that is at the 3' end of a longer unannotated transcript that is strongly upregulated in *Dnmt3l*-ko oocytes (data not shown). Therefore, we concluded that no TE subfamily appears to be derepressed if DNMT3L-dependent DNA methylation is abrogated. Retrotransposons with the highest activation in *Dnmt3c*-ko testes (L1Md_A, L1Md_T, L1Md_Gf, L1_Mm, MMERVK10C-int) [3] are silenced in both control and *Dnmt3l*-ko oocytes (data not shown). As mouse has one of the highest proportions of evolutionarily young LINE and LTR retrotransposons [33], we compared the expression of the most abundant subfamilies of these classes (LINE-L1, LTR–ERVK, LTR–ERVL, LTR–MaLR) and demonstrated that neither retrotransposon is substantially upregulated in *Dnmt3l*-ko oocytes. A relatively small number of specific insertions of various retrotransposon classes appear to be upregulated in *Dnmt3l*-ko oocytes; however, an equal or larger number of insertions of the same element is downregulated in *Dnmt3l*-ko oocytes (Additional file 1: Fig. S4c). Therefore, unlike *Dnmt3c*, which has a demonstrated role in silencing evolutionarily young retrotransposons in mouse testes, *Dnmt3l* does not appear to play a substantial role in retrotransposon silencing in mouse oocytes.

Absence of DNMT3L is not compensated by the expression of inactive DNMT3B isoforms in the oocytes of mammals outside Eumuroida rodents

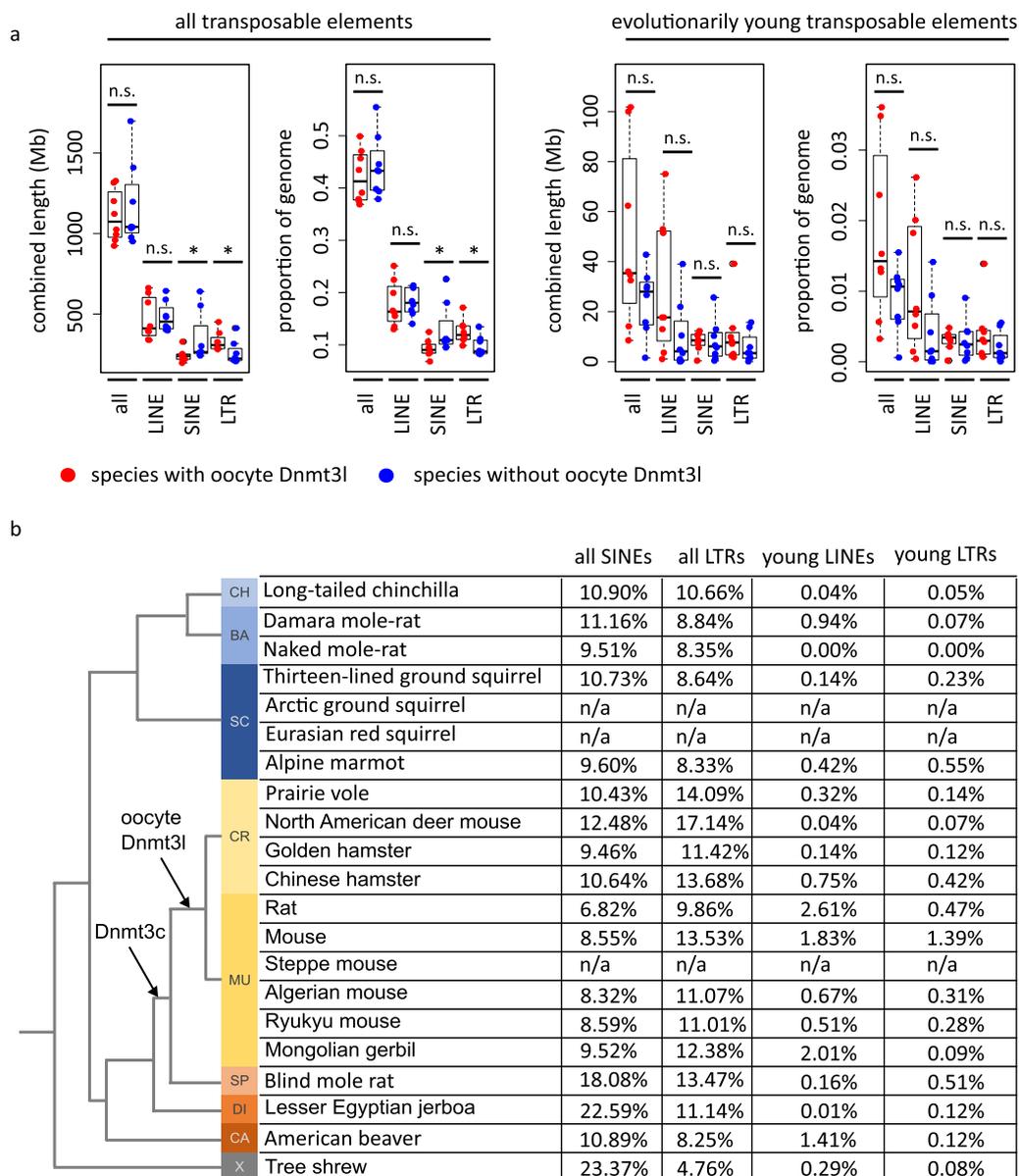It was previously suggested that DNMT3L may be substituted by an inactive isoform of DNMT3B in human

Fig. 4 Oocyte *Dnmt3l* does not play a substantial role in retrotransposon silencing. **a** Combined length of and proportion of genome represented by all transposons and selected retrotransposon classes for all transposons and evolutionary young transposons in rodent species with and without oocyte *Dnmt3l*. **b** Proportion of the genome represented by all SINEs, all LTRs, evolutionarily young LINEs and LTRs in individual species with and without oocyte *Dnmt3l*

oocytes [13], as was observed in somatic cells [15, 67]. This is supported by generally higher expression levels of *Dnmt3b* in species without oocyte *Dnmt3l* (Figs. 1a, 3a) and a higher *Dnmt3b* to *Dnmt3a* expression ratio (> 1 in all except cow and rhesus macaque) (Fig. 5a). However, our analysis of the RNA-seq data sets did not detect a correlation between the absence of oocyte *Dnmt3l* and increased expression of *Dnmt3b* isoforms homologous to the mouse inactive isoforms *Dnmt3b3*, *Dnmt3b4*

and *Dnmt3b7* (Fig. 5b), which are characterised by second and third exons from the 3' end being spliced out. Instead, such isoforms were detected at substantial expression levels only in mouse oocytes. However, this result could be affected by the technical differences across libraries from different species (Fig. 5b). Visual inspection of spliced reads across *Dnmt3b* in individual species did not reveal the presence of novel potentially inactive *Dnmt3b* isoforms. Therefore, it remains to be
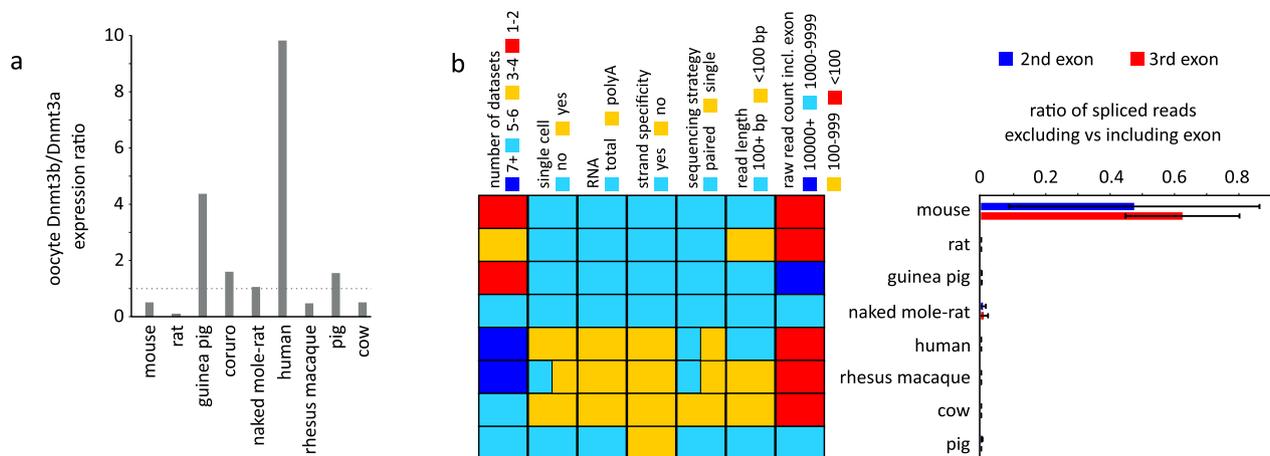
Fig. 5 Inactive DNMT3B isoform is not likely to replace DNMT3L. **a** *Dnmt3b/Dnmt3a* oocyte expression ratio across mammalian species; the dotted line marks a ratio of 1. **b** Right: ratio of spliced reads that splice out the two exons that are spliced out in mouse inactive *Dnmt3b* isoforms to spliced reads that include the two exons. Exons are marked in their order from the 3' end of the transcript. Left: technical information about data sets used for the analysis, from left to right—number of data sets analysed, whether the data sets are from single oocytes, whether total or polyA-selected RNA was used, whether the generated library is strand-specific, whether the library was sequenced single-end or paired-end, sequencing read length and average raw read count for spliced reads that include the two analysed exons. Two colours indicate different parameters if multiple sources of data sets were used (human and rhesus macaque)

elucidated what substitutes DNMT3L in a tetramer with DNMT3A, and whether both DNMT3A and DNMT3B are essential for oocyte de novo DNA methylation in species without oocyte DNMT3L.

Discussion

Here, we challenge the current model that DNMT3L is generally required for de novo DNA methylation in mammalian oocytes. Instead, we show that its expression in oocytes is driven by an oocyte-specific upstream promoter that emerged in the *Eumuroida* rodent lineage, restricting oocyte *Dnmt3l* activity to mice, rats, gerbils and hamsters. Our study demonstrates that results obtained by studying mouse as a model cannot be easily extrapolated to all mammals, as was also shown previously in other biological contexts [7, 19, 21, 22, 35, 43, 51, 56, 57], or even to all rodents. Moreover, it highlights the importance of examining non-model species to both elucidate the universality of mechanisms first identified in mouse and explore the variability across mammals.

As we did not detect the expression of inactive *Dnmt3b* isoforms in the oocytes of species outside *Eumuroida* rodents, we hypothesise that DNMT3A may act in a complex with active DNMT3B, or potentially form a homotetramer [16]. This interaction was previously shown in P19 somatic cells in the absence of DNMT3L; however, inactive DNMT3B was also found in the complex [28]. Long read sequencing of mammalian oocyte transcriptomes could shed more light on which *Dnmt3b* isoforms are expressed and, therefore, have the potential

to form a complex with DNMT3A. To date, only short read Illumina data sets are available for all species except mouse. When we attempted to analyse if any annotated exons are spliced out using the same approach, as shown in Fig. 5b, no exon in any species except mouse showed notable proportion of reads splicing the exon out compared to reads including the exon in the transcript (data not shown). This suggests that full length *Dnmt3b* isoforms are predominantly expressed irrespective of the presence or absence of oocyte DNMT3L.

Although it was originally believed that the DNMT3A–DNMT3L complex was targeted to the gene bodies of active genes in oocytes through DNMT3A PWWP domain-mediated recognition of H3K36me3 [48], recent findings showed that abolition of the association between DNMT3A and H3K36me2/3 in mouse oocytes does not result in hypomethylation of active gene bodies [26]. Instead, such oocytes are hypermethylated, particularly at transcriptionally inactive regions. Therefore, we hypothesise that the ADD domains of DNMT3L and/or DNMT3A target de novo DNA methyltransferase complex activity to regions lacking methylation at H3K4 (as was recently shown for DNMT3A [54]), while the DNMT3A PWWP domain further refines the preference by association with H3K36me2/3. Nevertheless, correct genomic landscapes of both H3K4 and H3K36 methylation are required for appropriate DNA methylation establishment [10, 19–21, 49, 65]. DNMT3B, including its inactive isoforms, contains the PWWP domain recognising H3K36me2/3, while this domain is absent

Behluli *et al. Epigenetics & Chromatin*   (2023) 16:43

Page 10 of 15

from DNMT3L. Additional PWWP domains in the DNMT3A–DNMT3B complex compared to DNMT3A–DNMT3L complex may, therefore, affect the recognition of the histone modification landscape permissive of de novo DNA methylation by enhancing the role of H3K36me2/3 in complex targeting. This might have consequences for the overall oocyte epigenomic landscape and could explain the differences in profiles of individual histone modifications between species with and without oocyte DNMT3L. In particular, large non-canonical H3K4me3 domains in mouse and rat oocytes might serve to exclude de novo DNA methylation targeted primarily by ADD domains of DNMT3A–DNMT3L complex. In contrast, these domains appear to be absent in human oocytes and less prominent in cow and pig oocytes [29], where de novo DNA methylation might be targeted primarily by H3K36me2/3 recognised by PWWP domains of the DNMT3A–DNMT3B complex.

In mouse, *Dnmt3l* is transcriptionally silent in primordial germ cells and its expression starts during oocyte growth [18]. Limitation of this study is that we cannot exclude the possibility that in species without oocyte *Dnmt3l* expression, *Dnmt3l* is expressed in primordial germ cells and then persists as a protein until oocyte growth when DNA methylation is established. Nevertheless, published proteomes of human and bovine oocytes do not detect DNMT3L protein [12, 31], while DNMT3B is detected. In mouse oocytes, DNMT3L protein was detected [59]. Moreover, *Dnmt3l* expression could occur earlier in the oogenesis and be no longer detectable at mRNA or protein level in fully grown oocytes while still playing a role in de novo DNA methylation establishment. Therefore, the role of DNMT3L in mammals outside *Eumuroida* rodents cannot be conclusively rejected or confirmed without generation of female germ line-specific DNMT3L-knock out individuals.

In mouse, it was shown that DNMT3L is essential for DNA methylation at transposable elements crucial for correct development of male germ cells [4]. *Dnmt3l* is transcribed from different promoters in male and female germline in mouse [45], suggesting independent regulation of expression. Therefore, the restriction of oocyte *Dnmt3l* expression to *Eumuroida* rodents cannot be extrapolated to male germline. Moreover, DNMT3L was shown to play a role in DNA methylation of extraembryonic ectoderm [2]. Although the requirement for DNMT3L in DNA methylation in male germline or placenta was not yet studied in any mammalian species outside mouse, it may be conserved, explaining the evolutionary stable presence of *Dnmt3l* gene in therian mammals.

We are aware that our conclusions are primarily supported by the identification of sequence motifs associated with *Dnmt3l* transcription in mouse and their absence across mammalian species outside *Eumuroida* rodents. The regulation of oocyte *Dnmt3l* TSS by TBPL2/TFIIA complex binding to the identified motif should be confirmed, for example, by TBPL2 chromatin immunoprecipitation. We do not exclude the possibility that *Dnmt3l* expression in other species could be regulated by different sequence motifs recognised by different oocyte TFs, or even evolutionarily more distant TBPL2/TFIIA complex—for example, in squirrels with relatively high similarity to mouse sequence motifs. Nevertheless, the analysis of available mammalian oocyte RNA-seq data sets strongly supports our conclusion that DNMT3L is not a universal mammalian de novo DNA methylation factor, as we could not detect oocyte *DNMT3L* expression in the oocytes of human, macaque rhesus, cow, pig and even another rodents, naked mole-rat and coruro, and only very weakly in guinea pig. Due to the very low expression level, it is unlikely that DNMT3L in guinea pig forms a tetramer complex with DNMT3A and plays an essential role in oocyte de novo DNA methylation establishment.

The development of low input and single cell approaches for profiling transcriptomes and epigenomes, together with high conservation of histone sequences and a high number of mammalian species with available genome sequences, facilitates interrogation of transcription, DNA methylation and histone modification landscapes in the oocytes of non-model species even when only scarce samples are available. This was already proven by mapping transcriptomes and epigenomes of the oocytes and early embryos of not only mouse and human, but also emerging mammalian models, such as cow and pig, revealing differences in oocyte epigenomes and different patterns of epigenome reprogramming after fertilisation [29]. Applying these approaches to non-model mammalian species would thus shed more light on the mechanisms of oocyte de novo DNA methylation establishment and its interplay with the histone modifications landscape, as well as the mechanisms of oocyte-to-zygotic transition at the transcriptome and epigenome levels. Importantly, it will also contribute to the understanding of the variability of these processes across mammals, reflecting different evolutionary histories and having functional consequences for the reproductive biology of individual species.

Conclusion

Our results suggest that the *Dnmt3l* oocyte-specific promoter emerged in the *Eumuroida* rodent lineage. Therefore, *Dnmt3l* is likely to be expressed and consequently to play a role as an essential de novo DNA methylation factor only in the oocytes of these species. In contrast

Behluli *et al. Epigenetics & Chromatin*   (2023) 16:43

Page 11 of 15

to *Dnmt3c,* which is required for silencing-specific evolutionarily young retrotransposons, *Dnmt3l* has a genome-wide role without a substantial effect on retrotransposon silencing. The mechanism governing de novo DNA methylation in the oocytes of most mammalian species including humans remains to be elucidated, as our results do not support the hypothesis that inactive isoforms of DNMT3B substitute DNMT3L in tetramers with DNMT3A.

Methods

Oocyte collection
Experimental procedures were carried out in accordance with EU directive 2010/63/EU (for animal experiments). Ovaries were dissected from 5 freshly sacrificed naked mole-rat females 1–8 years old, 3 guinea pig females 6–9 months old and one coruro female 2 years old. Oocytes (fully grown and at the late stage of growth) were isolated from ovaries by mechanical disruption of follicles in M2 medium (pre-warmed at 37 °C for a minimum of 2 h), and cumulus cells were removed by repeated mouth pipetting with a narrow glass pipette. Clean oocytes were washed through several 1X PBS drops.

RNA extraction and RNA-seq library preparation.

Immediately after oocyte collection, total RNA was extracted using an ARCTURUS PicoPure RNA Isolation Kit (Thermo Fisher Scientific; Cat. No. KIT0204) following the manufacturer's protocol. Due to the number of isolated oocytes, naked mole-rat samples were processed individually for each female, while guinea pig oocytes were merged together from all three females. Ribosomal RNA was depleted with NEBNext rRNA Depletion Kit v2 (Human/Mouse/Rat) with sample purification beads (NEB, E7405) according to the manufacturer's instructions, including DNase treatment by TURBO DNase (Thermo Fisher Scientific, AM2238) followed by library preparation using NEBNext Ultra II Directional RNA Library Prep with sample purification beads (NEB, E7765S) according to the manufacturer's instructions. The library was sequenced as 150 bp paired-end reads on an Illumina NovaSeq 6000 SP.

Data processing and mapping
In addition to our newly generated RNA-seq data set from naked mole-rat, guinea pig and coruro oocytes, we downloaded and processed additional publicly available data sets from oocytes of multiple mammalian species: mouse (GSE70116, GO2 and FGO only) [58], rat (GSE112622) [6], golden hamster (GSE86470) [17], cow (GSE61717) [41], pig (GSE108900) [53], rhesus macaque (GSE103313, GSE86938) [9, 60] and human (GSE36552, GSE101571) [62, 63]. In addition, we downloaded and processed SLIC–CAGE data sets from

control and TBPL2 knock-out oocytes (E-MTAB-8866) [66] and RNA-seq data sets from control and *Dnmt3l-*ko oocytes (DRA000570). Publicly available data sets were trimmed with Trim Galore! v0.4.1 or Trim Galore! V0.6.2 (E-MTAB-8866, DRA000570) with default parameters, specifying whether the reads were generated in single end or paired end mode. Naked mole-rat data were trimmed with Trim Galore! v0.6.2 in paired end mode. Trimmed RNA-seq data were mapped with Hisat2 v2.0.5 to respective genomes: GRCm38 and GRCm39 for mouse, Rnor_6.0 for rat, MesAur1.0 for golden hamster, Cavpor3.0 for guinea pig, HetGla_female_1.0 for naked mole-rat, Mmul_8 for rhesus macaque, GRCh38 for human, UMD3.1 for cow and Sscrofa11.1 for pig. SLIC–CAGE data sets (E-MTAB-8866) were mapped with Bowtie2 v2.5.1. Coruro data were mapped to degu Dnmt3a (XM_023711114.1), Dnmt3b (XM_004630806.2) and Dnmt3l (XM_023716073.1) sequences downloaded from the NCBI Nucleotide database with Bowtie2 v2.5.1 specifying the local alignment option. Quantified RNA-seq data sets from control and TBPL2 knock-out oocytes were analysed as published [66]. When analysing TE expression in mouse oocytes at the level of whole families, reads were mapped to artificially merged repeat sequences for the GRCm39 mouse genome downloaded from the UCSC genome browser separated by the NNNNN sequence using Bowtie2 v2.5.1.

Data analysis
The expression of DNMTs was visually inspected and quantified in SeqMonk v1.47.2 (https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/). To confirm that *Dnmt3l* is annotated at the correct position in the naked mole-rat genome, *Dnmt3l* sequences from the NCBI database for mouse, human and naked mole-rat were blasted to the naked mole-rat HetGla_female_1.0 genome using NCBI Genome Workbench. A custom Python script was prepared to identify the length of regions with continuous sequence identity between *Dnmt3l* vs *Dnmt3a* and *Dnmt3b*. Expression levels were $\log_2(RPKM+1)$-transformed in R and plotted using ggplot2 [61].

Aire intron sequences for the genome versions listed above and in Additional file 6: Tables S1 and S2 were downloaded from the Ensembl genome database and were converted to their reverse complements. Multiple sequence alignments of the original sequences and reverse complements were performed using MUltiple Sequence Comparison by Log-Expectation (MUSCLE) tool [30], with ClustalW output format and default parameters. The alignment and individual sequences of species without identified homologous sequence for predicted TBPL2/TFIIA binding site were further manually

checked for the presence for such sequence according to the sequence similarity and sequence similarity of the surrounding sequences. PhastCons analyses were performed using the PhastWeb interface [37], with expected length = 7 and default target coverage and rho values. Multiple sequence alignments and Newick trees generated by ClustalW were used as input files, and the mouse sequence was consistently used as a reference sequence. PhastCons scores were visualised using ggplot2 [61].

dN/dS ratios were computed for each of the *Dnmt3* genes using the CODEML software of PAML v4.9j [64] by running a null model (M0) that assumes equal dN/dS ratios across all branches of the phylogeny, and two branch models that assume differences in dN/dS ratios between foreground and background branches. For the first branch model *Hystricognathi* were designated as foreground branches, whereas only *Fukomys damarensis* and *Heterocephalus glaber* were set as foreground branches for the second. Chi-squared tests were conducted at 1% and 5% significance levels using the LRT statistics from each model [1]. The phylogenies and multiple sequence alignments used as CODEML input were generated using phytools [40] and MACSE v2.07 [38, 39], respectively. Rooted and unrooted consensus trees were created based on a subset of 1,000 node-dated phylogenetic trees downloaded from VertLife [55]. Sequences included in the alignments (listed in Additional file 6: Table S4) were downloaded from the NCBI Nucleotide database. For the alignment runs, internal frameshifts were replaced with "---", while internal stop codons were replaced with "NNN". The alignments were also trimmed, so that they start and end with at least 80% of the sequences having a nucleotide.

Expression of TE subfamilies (raw counts) was quantified using TEtranscripts v2.2.3 [24], using GRCm39 mouse genome and corresponding TE annotation available on the TEtranscripts website (https://labshare.cshl.edu/shares/mhammelllab/www-data/TEtranscripts/TE_GTF/). This was followed by the identification of differentially expressed TEs using NOISeq-sim within NOISeq [52]. For quantification of individual TE insertion expression, only elements with > 50 insertions outside gene annotation with > 5 reads (mapped using Hisat2 v2.0.5 reporting random one position for multimapping reads) in either condition were considered. TE annotation was obtained from UCSC genome browser for GRCm38.

DNA extraction, PCR and sanger sequencing
To obtain the missing guinea pig *Aire* intron 3 sequence, DNA was extracted from guinea pig liver using ethanol precipitation. PCR on the region of interest was performed using Phusion High-Fidelity PCR Master Mix with HF buffer (Thermo Fisher Scientific, F531S) and primers with sequences: forward AAGACTCCTGTACTGCCACC and reverse GCTGCAACTCCGAATTACCC. A single PCR product was confirmed by gel electrophoresis, purified by a Monarch PCR & DNA Cleanup Kit (NEB, T1030L) and sequenced by Sanger sequencing.

To confirm guinea pig *Dnmt3l* oocyte expression, RNA extracted from oocytes was treated with TURBO DNase (Thermo Fisher Scientific, AM2238), reverse transcribed with SuperScript III Reverse Transcriptase (Thermo Fisher Scientific, 18080093) and amplified by PCR using AMPIGENE Taq Mix (Enzo, ENZ-NUC100-0200) with primers-specific exons 8 + 10 (forward AGAGCTGATGAGTTTGGGCT, reverse GCCGTACACCAGGTCAAATG) of the annotated guinea pig *Dnmt3l* transcript ENSCPOT00000004115.3. PCR product was purified by a Monarch PCR & DNA Cleanup Kit (NEB, T1030L) and sequenced by Sanger sequencing.

Supplementary Information

Additional file 1: Figure S1. *Dnmt3l* regulation by the TBPL2/TFIIA complex and its TSS conservation across rodents. (a) Mouse oocyte SLIC–CAGE read quantification of the first 200 bp of mRNAs in control and TBPL2 knock-out oocytes [66]. The TSS of full-length oocyte *Dnmt3l* is highlighted in red. (b) Multiple sequence alignment of the reverse complements of *Aire* intron 3 sequences in rodents with available sequences in the Ensembl genome database. TSS and surrounding CT-rich and GA-rich sequences are highlighted in species with known or predicted *Dnmt3l* oocyte expression. Asterisk (*) and dollar ($) mark positions in blind mole rat and lesser Egyptian jerboa with insertions (without homologous sequences in other species) that are not visualised. Unser the alignment, the lengths of the insertions are indicated. The simplified phylogeny (in light grey) to the left of the species names is based on [50].

Additional file 2: Figure S2. PhastCons results for Aire introns 1–5. PhastCons scores across mouse *Aire* introns 1–5 in *Muridae* and *Cricetidae* (except prairie vole) rodents. In intron 3, the region highlighted in light grey represents the highly conserved region identified by PhastCons. The promoter sequence is highlighted in yellow and CT-rich and GA-rich regions surrounding the TSS are highlighted in blue.

Additional file 3: Figure S3. Oocyte *Dnmt3l* promoter within *Aire* intron 3 is not conserved in guinea pig. (a) Multiple sequence alignment of the *Aire* intron 3 sequence, inclu3ding the newly sequenced guinea pig sequence. The presumed TBPL2/TFIIA binding motif is highlighted in yellow. GP guinea pig, NMR naked mole-rat.

Additional file 4: Figure S4. Expression changes of retrotransposon insertions in *Dnmt3l*-ko mouse oocytes. (a) Percentage of reads mapped to retrotransposon classes in *Dnmt3l*-ko mouse oocytes compared to the control. (b) TE subfamilies that were identified as differentially changing in *Dnmt3l*-ko (ko) oocytes compared to wild type (wt). (c) Proportions of insertions that are > twofold upregulated (up), > twofold downregulated (down) or with smaller change in expression (same) in *Dnmt3l*-ko mouse oocytes compared to the control.

Additional file 5. Multiple sequence alignment of rodent *Aire* intron 3 sequences.

Additional file 6: Table S1. List of additional mammalian species with genome sequences in the Ensembl database and with annotated Aire gene that were used for analysis for the presence of a potential *Dnmt3l* oocyte promoter within the *Aire* intron equivalent to mouse intron 3. **Table S2.** List of all rodents with genome sequences in the Ensembl

Behluli *et al. Epigenetics & Chromatin*   (2023) 16:43

Page 13 of 15

database and with annotated *Aire* gene and a tree shrew that were used for analysis for the presence of a potential *Dnmt3l* oocyte promoter within the *Aire* intron equivalent to mouse intron 3. **Table S3.** Results of dN/dS ratio analysis for *Dnmt3* genes in rodents. **Table S4.** List of *Dnmt3* sequences used for computing dN/dS ratios. Sequences were downloaded from the NCBI Nucleotide database.

Availability of data and materials
The RNA-seq data sets supporting the conclusion of this article are available in the Gene Expression Omnibus (GEO) repository with accession number GSE236457. Custom scripts are available upon request.

Declarations

Competing interests
The authors declare no competing interests.

References
1. Álvarez-Carretero S, Kapli P, Yang Z. Beginner's guide on the use of PAML to detect positive selection. Mol Biol Evolut. 2023;40(4):msad041. https://doi.org/10.1093/molbev/msad041.
2. Andrews S, Krueger C, Mellado-Lopez M, Hemberger M, Dean W, Perez-Garcia V, Hanna CW. Mechanisms and function of de novo DNA methylation in placental development reveals an essential role for DNMT3B. Nat Commun. 2023;14(1):371. https://doi.org/10.1038/s41467-023-36019-9.
3. Barau J, Teissandier A, Zamudio N, Roy S, Nalesso V, Hérault Y, Guillou F, Bourc'his D. The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. Science. 2016;354(6314):909–12. https://doi.org/10.1126/science.aah5143.
4. Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature. 2004;431(7004):96–9. https://doi.org/10.1038/nature02886.
5. Bourc'his D, Xu G-L, Lin C-S, Bollman B, Bestor Timothy H. Dnmt3L and the establishment of maternal genomic imprints. Science. 2001;294(5551):2536–9. https://doi.org/10.1126/science.1065848.
6. Brind'Amour J, Kobayashi H, Richard Albert J, Shirane K, Sakashita A, Kamio A, Bogutz A, Koike T, Karimi MM, Lefebvre L, Kono T, Lorincz MC. LTR retrotransposons transcribed in oocytes drive species-specific and heritable changes in DNA methylation. Nat Commun. 2018;9(1):3331–3331. https://doi.org/10.1038/s41467-018-05841-x.
7. Carter AM. Animal models of human pregnancy and placentation: alternatives to the mouse. Reproduction. 2020;160(6):R129–43. https://doi.org/10.1530/REP-20-0354.
8. Chen Z, Zhang Y. Role of mammalian DNA methyltransferases in development. Annu Rev Biochem. 2020;89(1):135–58. https://doi.org/10.1146/annurev-biochem-103019-102815.
9. Chitwood JL, Burruel VR, Halstead MM, Meyers SA, Ross PJ. Transcriptome profiling of individual rhesus macaque oocytes and preimplantation embryos†. Biol Reprod. 2017;97(3):353–64. https://doi.org/10.1093/biolre/iox114.
10. Ciccone DN, Su H, Hevi S, Gay F, Lei H, Bajko J, Xu G, Li E, Chen T. KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. Nature. 2009;461(7262):415–8. https://doi.org/10.1038/nature08315.
11. Clarke HJ, Vieux K-F. Epigenetic inheritance through the female germline: the known, the unknown, and the possible. Semin Cell Dev Biol. 2015;43:106–16. https://doi.org/10.1016/j.semcdb.2015.07.003.
12. Dang Y, Zhu L, Yuan P, Liu Q, Guo Q, Chen X, Gao S, Liu X, Ji S, Yuan Y, Lian Y, Li R, Yan L, Wong CCL, Qiao J. Functional profiling of stage-specific proteome and translational transition across human pre-implantation embryo development at a single-cell resolution. Cell Discov. 2023;9(1):10. https://doi.org/10.1038/s41421-022-00491-2.
13. Demond H, Kelsey G. The enigma of DNA methylation in the mammalian oocyte [version 1; peer review: 4 approved]. F1000 Res. 2020;9:146. https://doi.org/10.12688/f1000research.21513.1.
14. Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, Jeltsch A. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation*. J Biol Chem. 2010;285(34):26114–20. https://doi.org/10.1074/jbc.M109.089433.
15. Duymich CE, Charlet J, Yang X, Jones PA, Liang G. DNMT3B isoforms without catalytic activity stimulate gene body methylation as accessory proteins in somatic cells. Nat Commun. 2016;7(1):11453. https://doi.org/10.1038/ncomms11453.
16. Emperle M, Bangalore DM, Adam S, Kunert S, Heil HS, Heinze KG, Bashtrykov P, Tessmer I, Jeltsch A. Structural and biochemical insight into the mechanism of dual CpG site binding and methylation by the DNMT3A DNA methyltransferase. Nucleic Acids Res. 2021;49(14):8294–308. https://doi.org/10.1093/nar/gkab600.
17. Franke V, Ganesh S, Karlic R, Malik R, Pasulka J, Horvat F, Kuzman M, Fulka H, Cernohorska M, Urbanova J, Svobodova E, Ma J, Suzuki Y, Aoki F, Schultz RM, Vlahovicek K, Svoboda P. Long terminal repeats power evolution of genes and gene expression programs in mammalian oocytes and zygotes. Genome Res. 2017. https://doi.org/10.1101/gr.216150.116.
18. Gahurova L, Tomizawa S-I, Smallwood SA, Stewart-Morgan KR, Saadeh H, Kim J, Andrews SR, Chen T, Kelsey G. Transcription and chromatin determinants of de novo DNA methylation timing in oocytes. Epigenet Chromatin. 2017;10:25–25. https://doi.org/10.1186/s13072-017-0133-5.
19. Hanna CW, Demond H, Kelsey G. Epigenetic regulation in development: is the mouse a good model for the human? Hum Reprod Update. 2018;24(5):556–76. https://doi.org/10.1093/humupd/dmy021.
20. Hanna CW, Huang J, Belton C, Reinhardt S, Dahl A, Andrews S, Stewart AF, Kranz A, Kelsey G. Loss of histone methyltransferase SETD1B in oogenesis results in the redistribution of genomic histone 3 lysine 4 trimethylation. Nucleic Acids Res. 2022;50(4):1993–2004. https://doi.org/10.1093/nar/gkac051.
21. Hanna CW, Taudt A, Huang J, Gahurova L, Kranz A, Andrews S, Dean W, Stewart AF, Colomé-Tatché M, Kelsey G. MLL2 conveys transcription-independent H3K4 trimethylation in oocytes. Nat Struct Mol Biol. 2018;25(1):73–82. https://doi.org/10.1038/s41594-017-0013-5.
22. Inoue A. Noncanonical imprinting: intergenerational epigenetic inheritance mediated by Polycomb complexes. Curr Opin Genet Dev. 2023;78:102015. https://doi.org/10.1016/j.gde.2022.102015.
23. Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. Nature. 2007;449(7159):248–51. https://doi.org/10.1038/nature06146.
24. Jin Y, Tam OH, Paniagua E, Hammell M. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. Bioinformatics. 2015;31(22):3593–9. https://doi.org/10.1093/bioinformatics/btv422.
25. Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, Sasaki H. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature. 2004;429(6994):900–3. https://doi.org/10.1038/nature02633.

26. Kibe K, Shirane K, Ohishi H, Uemura S, Toh H, Sasaki H. The DNMT3A PWWP domain is essential for the normal DNA methylation landscape in mouse somatic cells and oocytes. PLoS Genet. 2021;17(5):e1009570. https://doi.org/10.1371/journal.pgen.1009570.

27. Kobayashi H, Sakurai T, Imai M, Takahashi N, Fukuda A, Yayoi O, Sato S, Nakabayashi K, Hata K, Sotomaru Y, Suzuki Y, Kono T. Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. PLoS Genet. 2012;8(1):e1002440. https://doi.org/10.1371/journal.pgen.1002440.

28. Li J-Y, Pu M-T, Hirasawa R, Li B-Z, Huang Y-N, Zeng R, Jing N-H, Chen T, Li E, Sasaki H, Xu G-L. Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog. Mol Cell Biol. 2007;27(24):8748–59. https://doi.org/10.1128/MCB.01380-07.

29. Lu X, Zhang Y, Wang L, Wang L, Wang H, Xu Q, Xiang Y, Chen C, Kong F, Xia W, Lin Z, Ma S, Liu L, Wang X, Ni H, Li W, Guo Y, Xie W. Evolutionary epigenomic analyses in mammalian early embryos reveal species-specific innovations and conserved principles of imprinting. Sci Adv. 2021;7(48):eabi6178. https://doi.org/10.1126/sciadv.abi6178.

30. Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, Madhusoodanan N, Kolesnikov A, Lopez R. Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res. 2022. https://doi.org/10.1093/nar/gkac240.

31. Memili E, Peddinti D, Shack LA, Nanduri B, McCarthy F, Sagirkaya H, Burgess SC. Bovine germinal vesicle oocyte and cumulus cell proteomics. Reproduction. 2007;133(6):1107–20. https://doi.org/10.1530/REP-06-0149.

32. Ooi SKT, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin S-P, Allis CD, Cheng X, Bestor TH. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature. 2007;448(7154):714–7. https://doi.org/10.1038/nature05987.

33. Osmanski AB, Paulat NS, Korstian J, Grimshaw JR, Halsey M, Sullivan KAM, Moreno-Santillán DD, Crookshanks C, Roberts J, Garcia C, Johnson MG, Densmore LD, Stevens RD, Zoonomia C, Rosen J, Storer JM, Hubley R, Smit AFA, Dávalos LM, Karlsson EK, Lindblad-Toh K, Ray DA, Andrews G, Armstrong JC, Bianchi M, Birren BW, Bredemeyer KR, Breit AM, Christmas MJ, Clawson H, Damas J, Di Palma F, Diekhans M, Dong MX, Eizirik E, Fan K, Fanter C, Foley NM, Forsberg-Nilsson K, Garcia CJ, Gatesy J, Gazal S, Genereux DP, Goodman L, Grimshaw J, Halsey MK, Harris AJ, Hickey G, Hiller M, Hindle AG, Hubley RM, Hughes GM, Johnson J, Juan D, Kaplow IM, Karlsson EK, Keough KC, Kirilenko B, Koepfli K-P, Korstian JM, Kowalczyk A, Kozyrev SV, Lawler AJ, Lawless C, Lehmann T, Levesque DL, Lewin HA, Li X, Lind A, Lindblad-Toh K, Mackay-Smith A, Marinescu VD, Marques-Bonet T, Mason VC, Meadows JRS, Meyer WK, Moore JE, Moreira LR, Moreno-Santillan DD, Morrill KM, Muntané G, Murphy WJ, Navarro A, Nweeia M, Ortmann S, Osmanski A, Paten B, Paulat NS, Pfenning AR, Phan BN, Pollard KS, Pratt HE, Ray DA, Reilly SK, Rosen JR, Ruf I, Ryan L, Ryder OA, Sabeti PC, Schäffer DE, Serres A, Shapiro B, Smit AFA, Springer M, Srinivasan C, Steiner C, Storer JM, Sullivan KAM, Sullivan PF, Sundström E, Supple MA, Swofford R, Talbot J-E, Teeling E, Turner-Maier J, Valenzuela A, Wagner F, Wallerman O, Wang C, Wang J, Weng Z, Wilder AP, Wirthlin ME, Xue JR, Zhang X. Insights into mammalian TE diversity through the curation of 248 genome assemblies. Science. 2023;380(6643):1430. https://doi.org/10.1126/science.abn1430.

34. Otani J, Nankumo T, Arita K, Inamoto S, Ariyoshi M, Shirakawa M. Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX–DNMT3–DNMT3L domain. EMBO Rep. 2009;10(11):1235–41. https://doi.org/10.1038/embor.2009.218.

35. Perlman RL. Mouse models of human disease: an evolutionary perspective. Evolut Med Public Health. 2016;2016(1):170–6. https://doi.org/10.1093/emph/eow014.

36. Petrussa L, Van de Velde H, De Rycke M. Dynamic regulation of DNA methyltransferases in human oocytes and preimplantation embryos after assisted reproductive technologies. Mol Hum Reprod. 2014;20(9):861–74. https://doi.org/10.1093/molehr/gau049.

37. Ramani R, Krumholz K, Huang Y-F, Siepel A. PhastWeb: a web interface for evolutionary conservation scoring of multiple sequence alignments using phastCons and phyloP. Bioinformatics. 2019;35(13):2320–2. https://doi.org/10.1093/bioinformatics/bty966.

38. Ranwez V, Douzery EJP, Cambon C, Chantret N, Delsuc F. MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. Mol Biol Evol. 2018;35(10):2582–4. https://doi.org/10.1093/molbev/msy159.

39. Ranwez V, Harispe S, Delsuc F, Douzery EJP. MACSE: multiple alignment of coding SEquences accounting for frameshifts and stop codons. PLoS ONE. 2011;6(9):e22594. https://doi.org/10.1371/journal.pone.0022594.

40. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evolut. 2012;3(2):217–23. https://doi.org/10.1111/j.2041-210X.2011.00169.x.

41. Reyes JM, Chitwood JL, Ross PJ. RNA-Seq profiling of single bovine oocyte transcript abundance and its modulation by cytoplasmic polyadenylation. Mol Reprod Dev. 2015;82(2):103–14. https://doi.org/10.1002/mrd.22445.

42. Sendžikaitė G, Kelsey G. The role and mechanisms of DNA methylation in the oocyte. Essays Biochem. 2019;63(6):691–705. https://doi.org/10.1042/EBC20190043.

43. Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker HV, Xu W, Richards DR, McDonald-Smith GP, Gao H, Hennessy L, Finnerty CC, López CM, Honari S, Moore EE, Minei JP, Cuschieri J, Bankey PE, Johnson JL, Sperry J, Nathens AB, Billiar TR, West MA, Jeschke MG, Klein MB, Gamelli RL, Gibran NS, Brownstein BH, Miller-Graziano C, Calvano SE, Mason PH, Cobb JP, Rahme LG, Lowry SF, Maier RV, Moldawer LL, Herndon DN, Davis RW, Xiao W, Tompkins RG, Null N, Abouhamze A, Balis UGJ, Camp DG, De AK, Harbrecht BG, Hayden DL, Kaushal A, O'Keefe GE, Kotz KT, Qian W, Schoenfeld DA, Shapiro MB, Silver GM, Smith RD, Storey JD, Tibshirani R, Toner M, Wilhelmy J, Wispelwey B, Wong WH. Genomic responses in mouse models poorly mimic human inflammatory diseases. Proc Natl Acad Sci. 2013;110(9):3507–12. https://doi.org/10.1073/pnas.1222878110.

44. Shirane K, Toh H, Kobayashi H, Miura F, Chiba H, Ito T, Kono T, Sasaki H. Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of Non-CpG methylation and role of DNA methyltransferases. PLoS Genet. 2013;9(4):e1003439. https://doi.org/10.1371/journal.pgen.1003439.

45. Shovlin TC, Bourchis D, La Salle S, O'Doherty A, Trasler JM, Bestor TH, Walsh CP. Sex-specific promoters regulate Dnmt3L expression in mouse germ cells. Hum Reprod. 2007;22(2):457–67. https://doi.org/10.1093/humrep/del379.

46. Smallwood SA, Kelsey G. De novo DNA methylation: a germ cell perspective. Trends Genet. 2012;28(1):33–42. https://doi.org/10.1016/j.tig.2011.09.004.

47. Smallwood SA, Tomizawa S-I, Krueger F, Ruf N, Carli N, Segonds-Pichon A, Sato S, Hata K, Andrews SR, Kelsey G. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. Nat Genet. 2011;43(8):811–4. https://doi.org/10.1038/ng.864.

48. Stewart KR, Veselovska L, Kelsey G. Establishment and functions of DNA methylation in the germline. Epigenomics. 2016;8(10):1399–413. https://doi.org/10.2217/epi-2016-0056.

49. Stewart KR, Veselovska L, Kim J, Huang J, Saadeh H, Tomizawa S-I, Smallwood SA, Chen T, Kelsey G. Dynamic changes in histone modifications precede de novo DNA methylation in oocytes. Genes Dev. 2015;29(23):2449–62. https://doi.org/10.1101/gad.271353.115.

50. Swanson MT, Oliveros CH, Esselstyn JA. A phylogenomic rodent tree reveals the repeated evolution of masseter architectures. Proc Royal Soc B Biol Sci. 2019;286(1902):20190672. https://doi.org/10.1098/rspb.2019.0672.

51. Swearengen JR. Choosing the right animal model for infectious disease research. Animal Models Exp Med. 2018;1(2):100–8. https://doi.org/10.1002/ame2.12020.

52. Tarazona S, Furió-Tarí P, Turrà D, Pietro AD, Nueda MJ, Ferrer A, Conesa A. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. Nucleic Acids Res. 2015;43(21):e140–e140. https://doi.org/10.1093/nar/gkv711.

53. Tsai T-S, Tyagi S, St John JC. The molecular characterisation of mitochondrial DNA deficient oocytes using a pig model. Hum Reprod. 2018;33(5):942–53. https://doi.org/10.1093/humrep/dey052.

54. Uehara R, Au Yeung WK, Toriyama K, Ohishi H, Kubo N, Toh H, Suetake I, Shirane K, Sasaki H. The DNMT3A ADD domain is required for efficient de novo DNA methylation and maternal imprinting in mouse oocytes. PLoS Genet. 2023;19(8):e1010855. https://doi.org/10.1371/journal.pgen.1010855.

55. Upham NS, Esselstyn JA, Jetz W. Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. PLoS Biol. 2019;17(12):e3000494. https://doi.org/10.1371/journal.pbio.3000494.

56. van der Worp HB, Howells DW, Sena ES, Porritt MJ, Rewell S, O'Collins V, Macleod MR. Can animal models of disease reliably inform human studies? PLoS Med. 2010;7(3):e1000245. https://doi.org/10.1371/journal.pmed.1000245.

57. Vandamme T. Use of rodents as models of human diseases [Review Article]. J Pharm Bioallied Sci. 2014;6(1):2–9. https://doi.org/10.4103/0975-7406.124301.

58. Veselovska L, Smallwood SA, Saadeh H, Stewart KR, Krueger F, Maupetit-Méhouas S, Arnaud P, Tomizawa S-I, Andrews S, Kelsey G. Deep sequencing and de novo assembly of the mouse oocyte transcriptome define the contribution of transcription to the DNA methylation landscape. Genome Biol. 2015;16:209–209. https://doi.org/10.1186/s13059-015-0769-z.

59. Wang S, Kou Z, Jing Z, Zhang Y, Guo X, Dong M, Wilmut I, Gao S. Proteome of mouse oocytes at different developmental stages. Proc Natl Acad Sci. 2010;107(41):17639–44. https://doi.org/10.1073/pnas.1013185107.

60. Wang X, Liu D, He D, Suo S, Xia X, He X, Han J-DJ, Zheng P. Transcriptome analyses of rhesus monkey preimplantation embryos reveal a reduced capacity for DNA double-strand break repair in primate oocytes and early embryos. Genome Res. 2017;27(4):567–79. https://doi.org/10.1101/gr.198044.115.

61. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016.

62. Wu J, Xu J, Liu B, Yao G, Wang P, Lin Z, Huang B, Wang X, Li T, Shi S, Zhang N, Duan F, Ming J, Zhang X, Niu W, Song W, Jin H, Guo Y, Dai S, Hu L, Fang L, Wang Q, Li Y, Li W, Na J, Xie W, Sun Y. Chromatin analysis in human early development reveals epigenetic transition during ZGA. Nature. 2018;557(7704):256–60. https://doi.org/10.1038/s41586-018-0080-8.

63. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, Liu P, Lian Y, Zheng X, Yan J, Huang J, Li M, Wu X, Wen L, Lao K, Li R, Qiao J, Tang F. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol. 2013;20(9):1131–9. https://doi.org/10.1038/nsmb.2660.

64. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24(8):1586–91. https://doi.org/10.1093/molbev/msm088.

65. Yano S, Ishiuchi T, Abe S, Namekawa SH, Huang G, Ogawa Y, Sasaki H. Histone H3K36me2 and H3K36me3 form a chromatin platform essential for DNMT3A-dependent DNA methylation in mouse oocytes. Nat Commun. 2022;13(1):4440. https://doi.org/10.1038/s41467-022-32141-2.

66. Yu C, Cvetesic N, Hisler V, Gupta K, Ye T, Gazdag E, Negroni L, Hajkova P, Berger I, Lenhard B, Müller F, Vincent SD, Tora L. TBPL2/TFIIA complex establishes the maternal transcriptome through oocyte-specific promoter usage. Nat Commun. 2020;11(1):6439. https://doi.org/10.1038/s41467-020-20239-4.

67. Zeng Y, Ren R, Kaur G, Hardikar S, Ying Z, Babcock L, Gupta E, Zhang X, Chen T, Cheng X. The inactive Dnmt3b3 isoform preferentially enhances Dnmt3b-mediated DNA methylation. Genes Dev. 2020;34(21–22):1546–58.

68. Zhang Y, Jurkowska R, Soeroes S, Rajavelu A, Dhayalan A, Bock I, Rathert P, Brandt O, Reinhardt R, Fischle W, Jeltsch A. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. Nucleic Acids Res. 2010;38(13):4246–53. https://doi.org/10.1093/nar/gkq147.

Publisher's Note