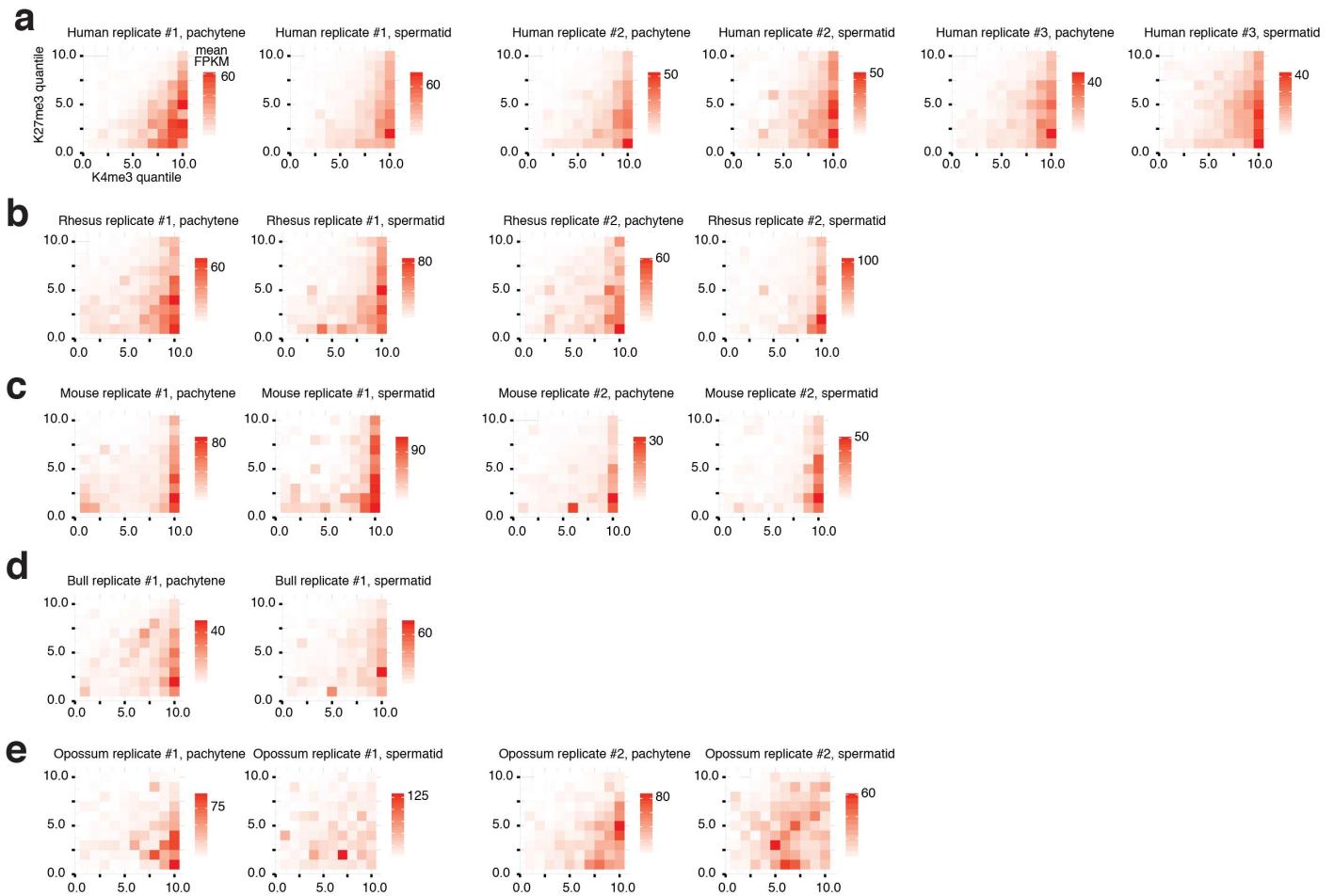


**Supplementary Figure 1**

**Assessment of sample purity and quality.**

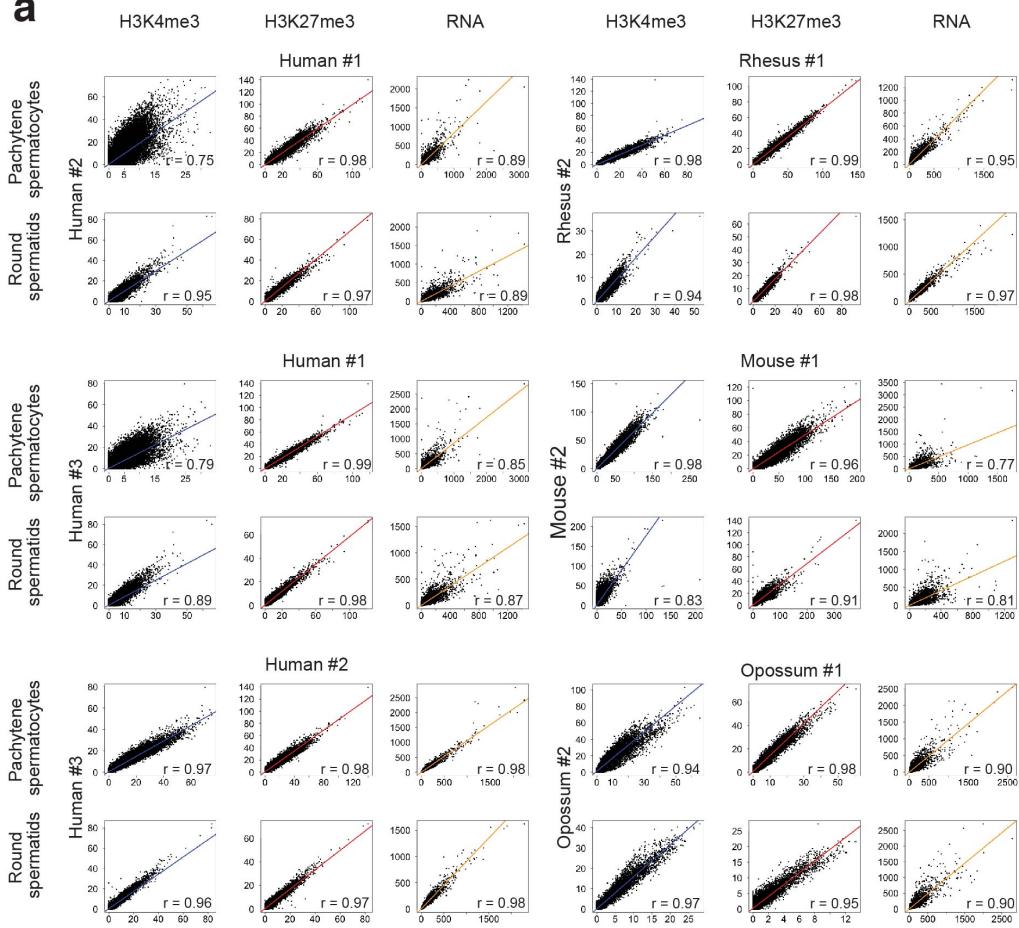
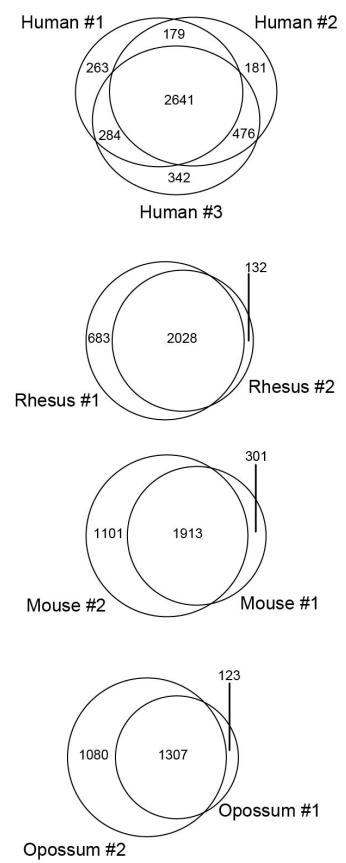
(a) Hematoxylin and eosin staining of formaldehyde-fixed, paraffin-embedded sections from a human testis biopsy collected concurrently with the sample used for cell sorting. Left, 10 $\times$  magnification; right, 40 $\times$  magnification of the boxed region at left. (b) Phase-contrast microscopy images of dissociated spermatogenic cells before StaPut and of sorted pachytene spermatocyte and round spermatid populations after StaPut. All images are shown at 40 $\times$  magnification. A small population of contaminating red blood cells, which do not contain chromatin, was present in the human round spermatid sample shown (arrows). (c) Numbers of sorted cells (of a total of 100) from the fractions shown that are identifiable as belonging to the reported cell type population in each fraction.



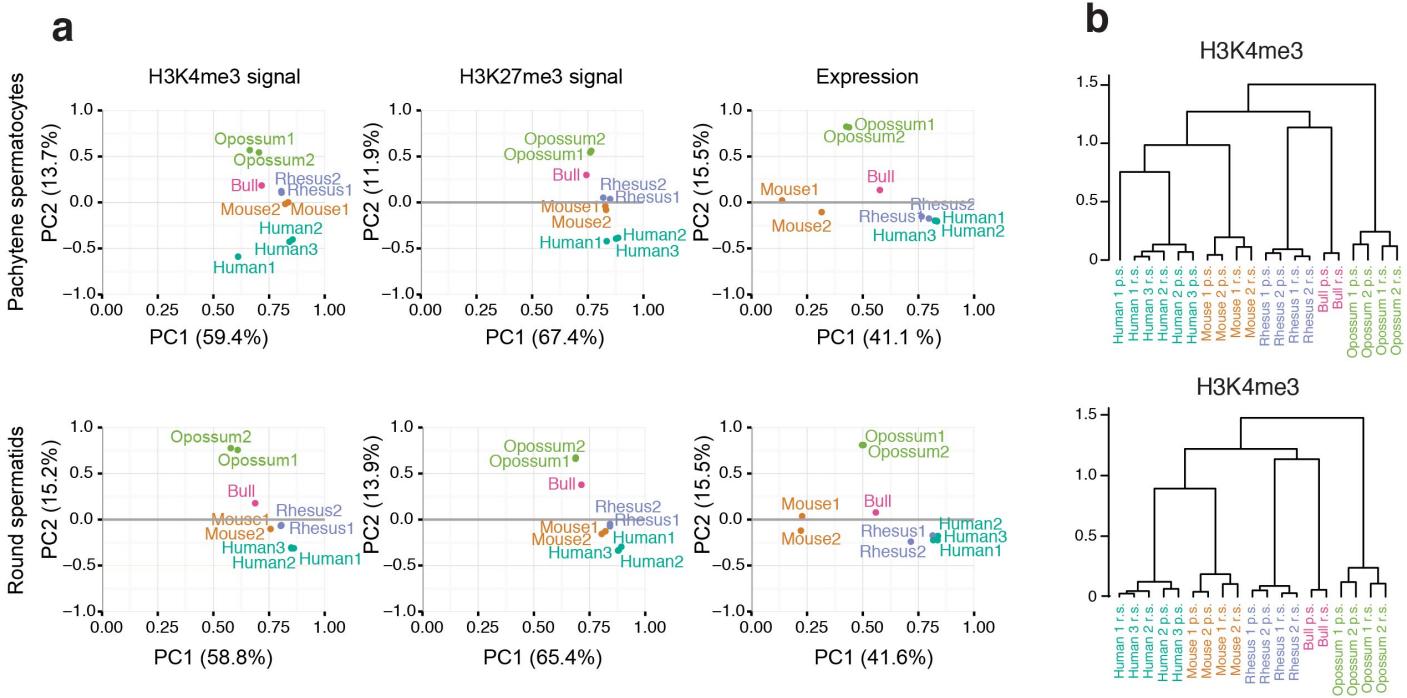
## Supplementary Figure 2

### Correspondence between H3K4me3, H3K27me3, and expression levels.

(a–e) Heat maps showing mean expression level as a function of H3K4me3 and H3K27me3 quantile, as shown in **Figure 1b**. Each pachytene spermatocyte and round spermatid sample is shown separately for human (a), rhesus (b), mouse (c), bull (d), and opossum (e). The plots for human replicate 2 are identical to those shown in **Figure 1b** and are reproduced here for comparison.

**a****b****Supplementary Figure 3****Correlations between biological replicates.**

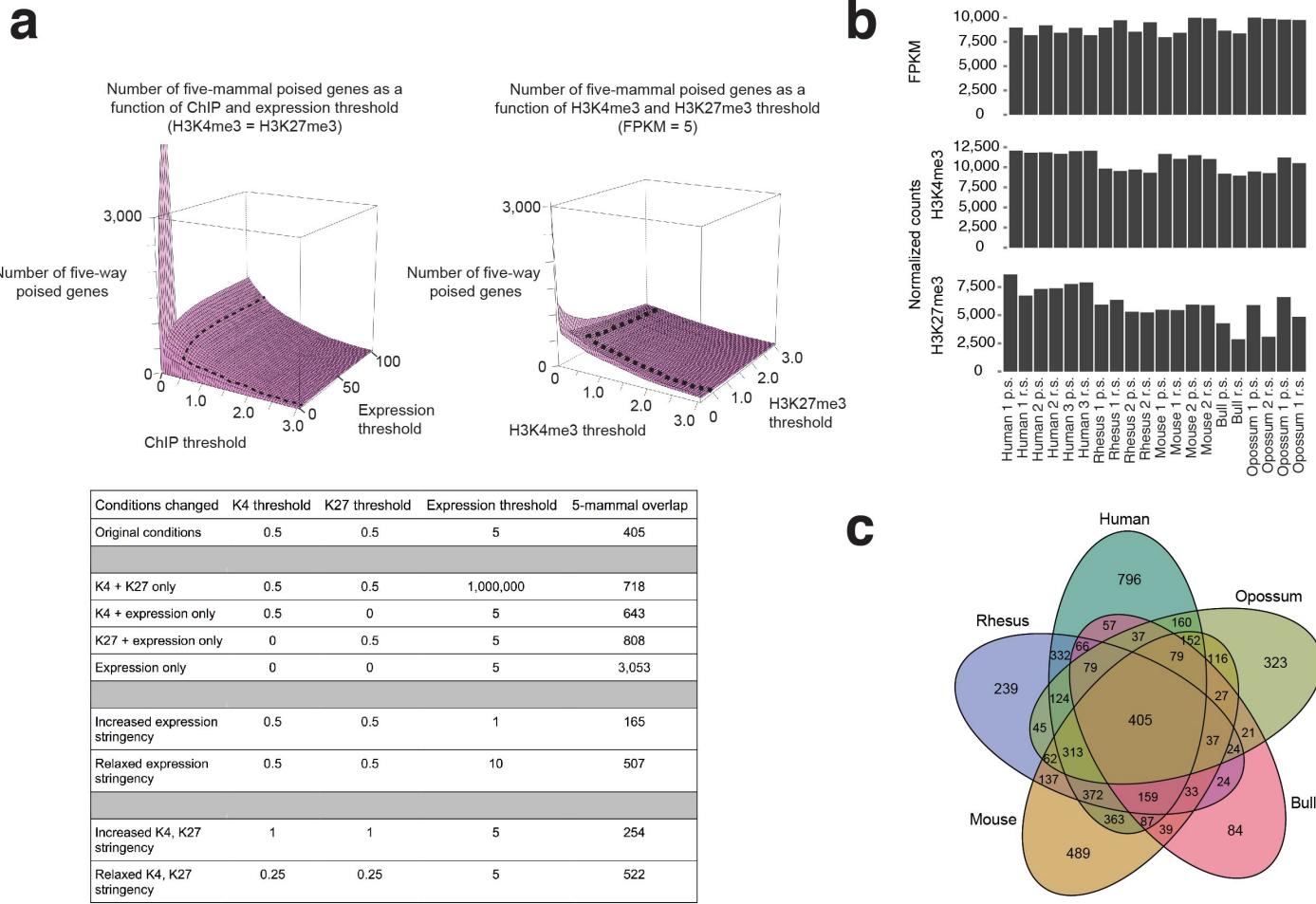
(a) Correlation of normalized, input-subtracted ChIP-seq signal or normalized RNA-seq counts by gene. Pearson's correlation coefficient ( $r$ ) is shown for each.  $P < 2.2 \times 10^{-16}$  for all correlations. (b) Overlap between poised gene sets called for each human, rhesus, mouse, or opossum replicate.



**Supplementary Figure 4**

**Principal-component and clustering analysis of H3K4me3, H3K27me3, and RNA-seq data.**

(a) Principal-component analysis. Axes represent the first two principal components, labeled with the percentage of total variance explained by each. Left, H3K4me3 signal; middle, H3K27me3 signal; right, expression (FPKM). (b) Clustering of H3K4me3 data with and without inclusion of an outlier data set. Top, clustering of H3K4me3 data reproduced from **Figure 1c**; bottom, the same clustering analysis after removal of the pachytene sample from human replicate 1 ("human 1 p.s."), the least well correlated of the biological replicates. Removal of the outlier does not alter the clustering result.



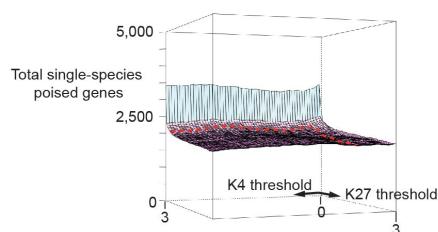
## Supplementary Figure 5

### Effect of changing ChIP and expression thresholds on the number of poised genes called.

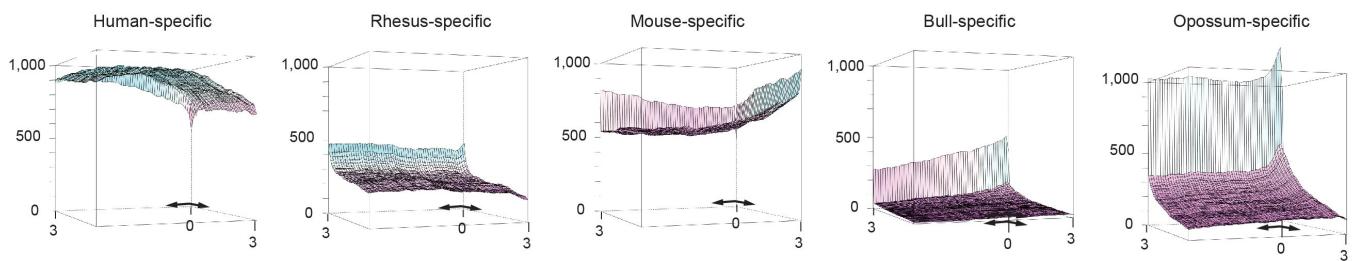
(a) Top left, numbers of five-mammal poised genes called as expression and ChIP thresholds vary; H3K4me3 and H3K27me3 thresholds are set as equal in all conditions. Top right, numbers of five-mammal poised genes called as H3K4me3 and H3K27me3 thresholds vary relative to each other; the expression threshold is held constant at FPKM  $\leq 5$ . Dashed lines show the gene space included by the criteria used in this study. (b) Numbers of genes meeting each threshold independently (expression, H3K4me3, or H3K27me3) in each sample. (c) Numbers of poised genes for each overlap condition among five mammalian species.

**a**

Total number of single-species poised genes as a function of H3K4me3 and H3K27me3 thresholds



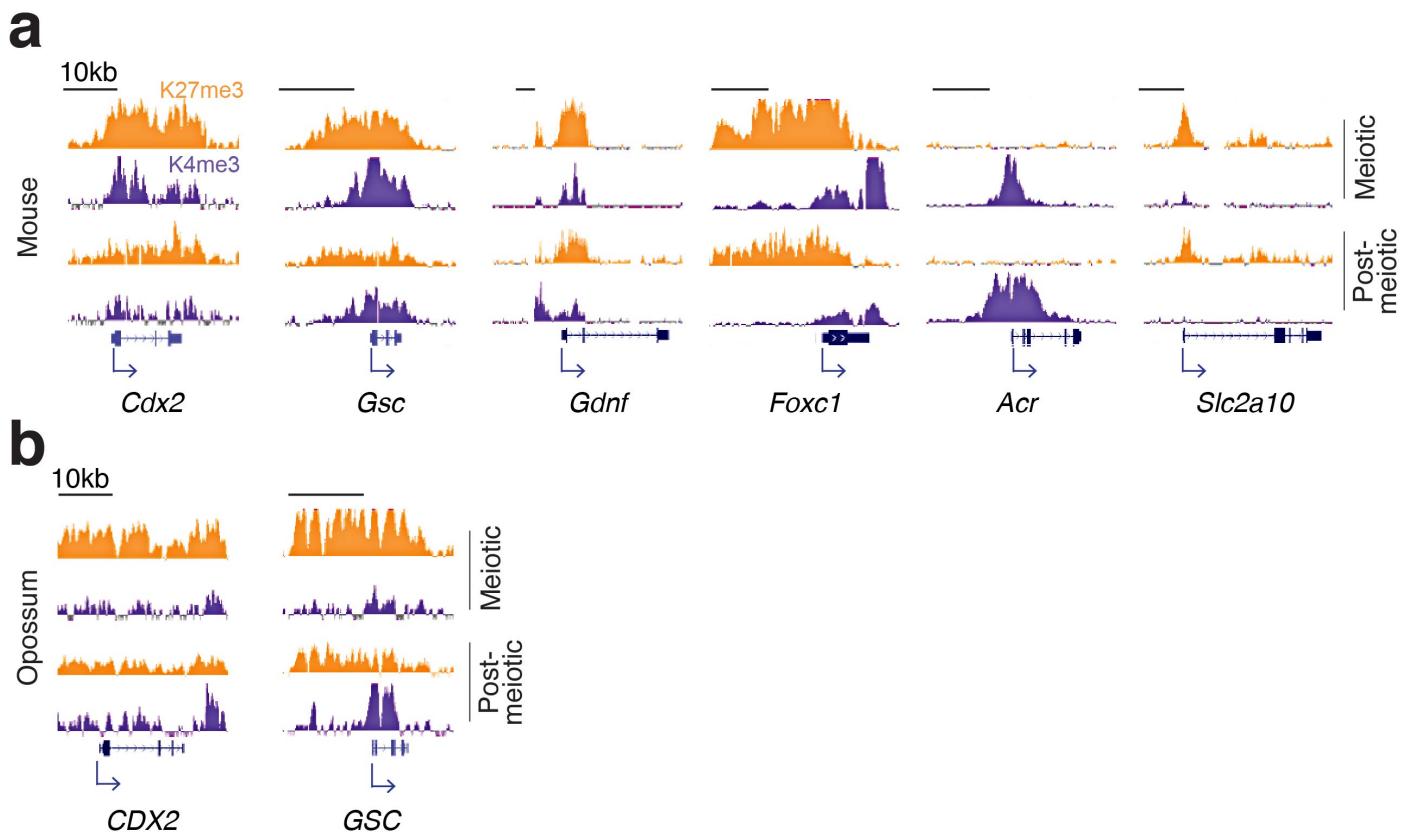
Conditions changed	K4 threshold	K27 threshold	Expression threshold	Total single-species poised genes
Original conditions	0.5	0.5	5	1,931
K4 + K27 only	0.5	0.5	1,000,000	2,403
K4 + expression only	0.5	0	5	3,057
K27 + expression only	0	0.5	5	2,016
Expression only	0	0	5	2,776
Increased expression stringency	0.5	0.5	1	1,561
Relaxed expression stringency	0.5	0.5	10	2,013
Increased K4, K27 stringency	1	1	5	1,806
Relaxed K4, K27 stringency	0.25	0.25	5	2,105
Relaxed K4, increased K27 stringency	0.25	1	5	1,825
Increased K4, relaxed K27 stringency	1	0.25	5	2,013

**b**

### Supplementary Figure 6

**Effect of changing ChIP and expression thresholds on the number of poised genes called in only one of five mammalian species.**

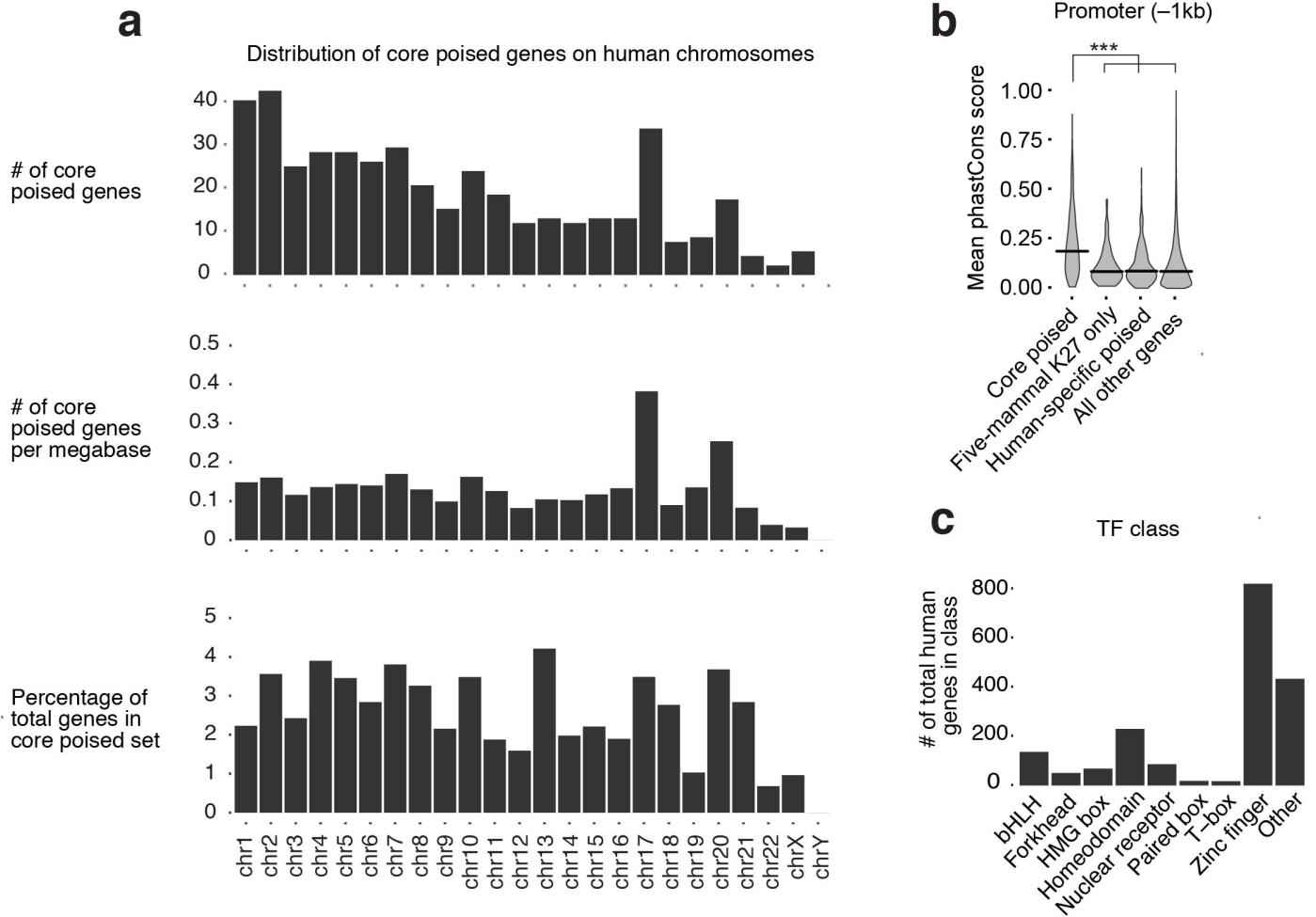
(a) Left, total numbers of species-specific poised genes called as H3K4me3 and H3K27me3 ChIP thresholds vary; the expression threshold is held constant at FPKM  $\leq 5$ . The red dashed line on the plot shows the gene space included by the criteria in this study. Right, sample values for numbers of total species-specific poised genes called using different combinations of H3K4me3, H3K27me3, and expression thresholds. (b) Numbers of species-specific poised genes called for individual species as H3K4me3 and H3K27me3 ChIP thresholds vary, with the expression threshold held constant at FPKM  $\leq 5$ .



**Supplementary Figure 7**

**ChIP signal at sequential ChIP–qPCR targets.**

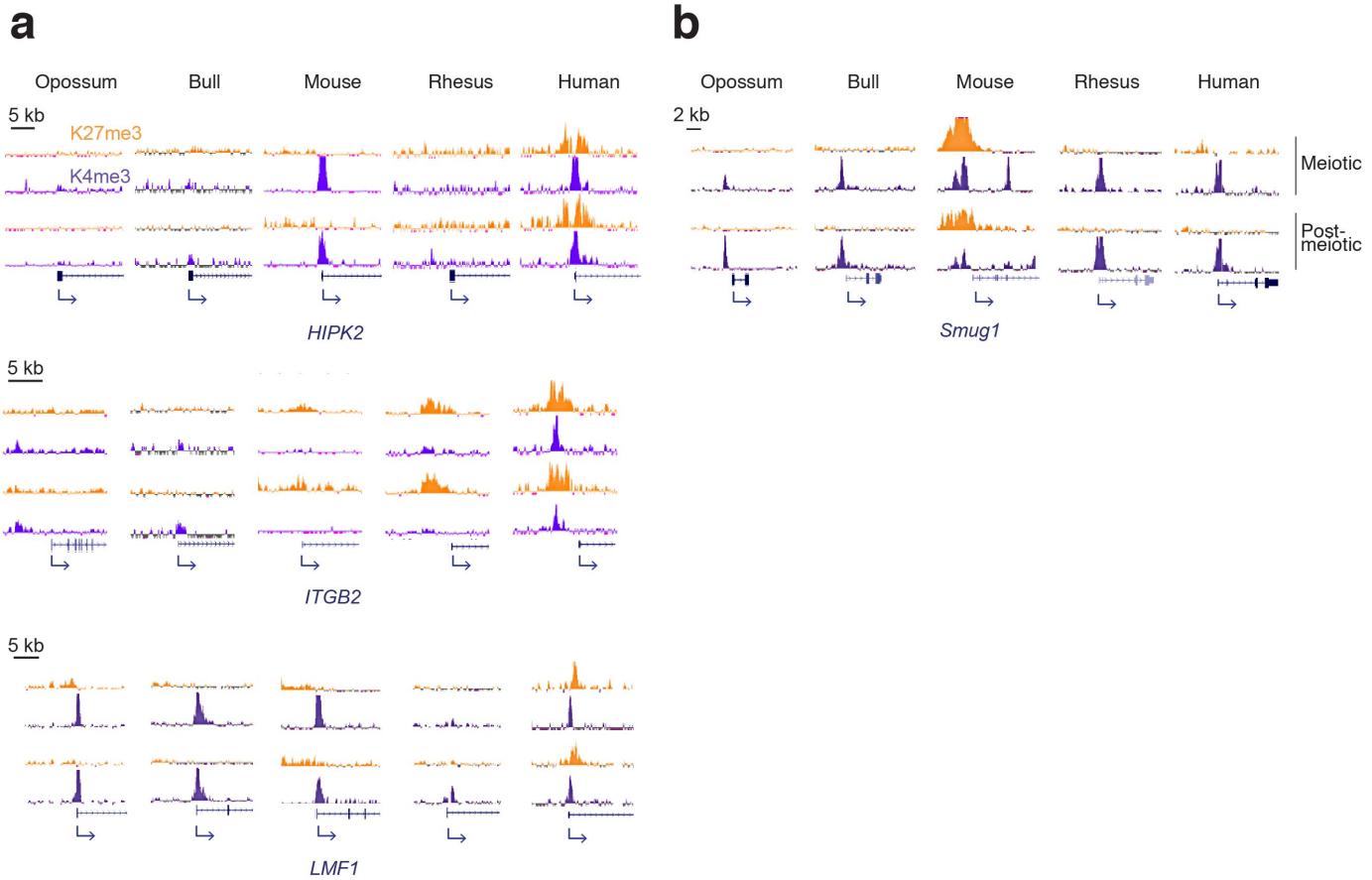
(a) Input-subtracted ChIP signal tracks for the promoters targeted in mouse sequential ChIP–qPCR experiments (Fig. 2c). (b) Input-subtracted ChIP signal tracks for the promoters targeted in opossum sequential ChIP–qPCR experiments (Fig. 2d).



**Supplementary Figure 8**

**Characteristics of core poised genes.**

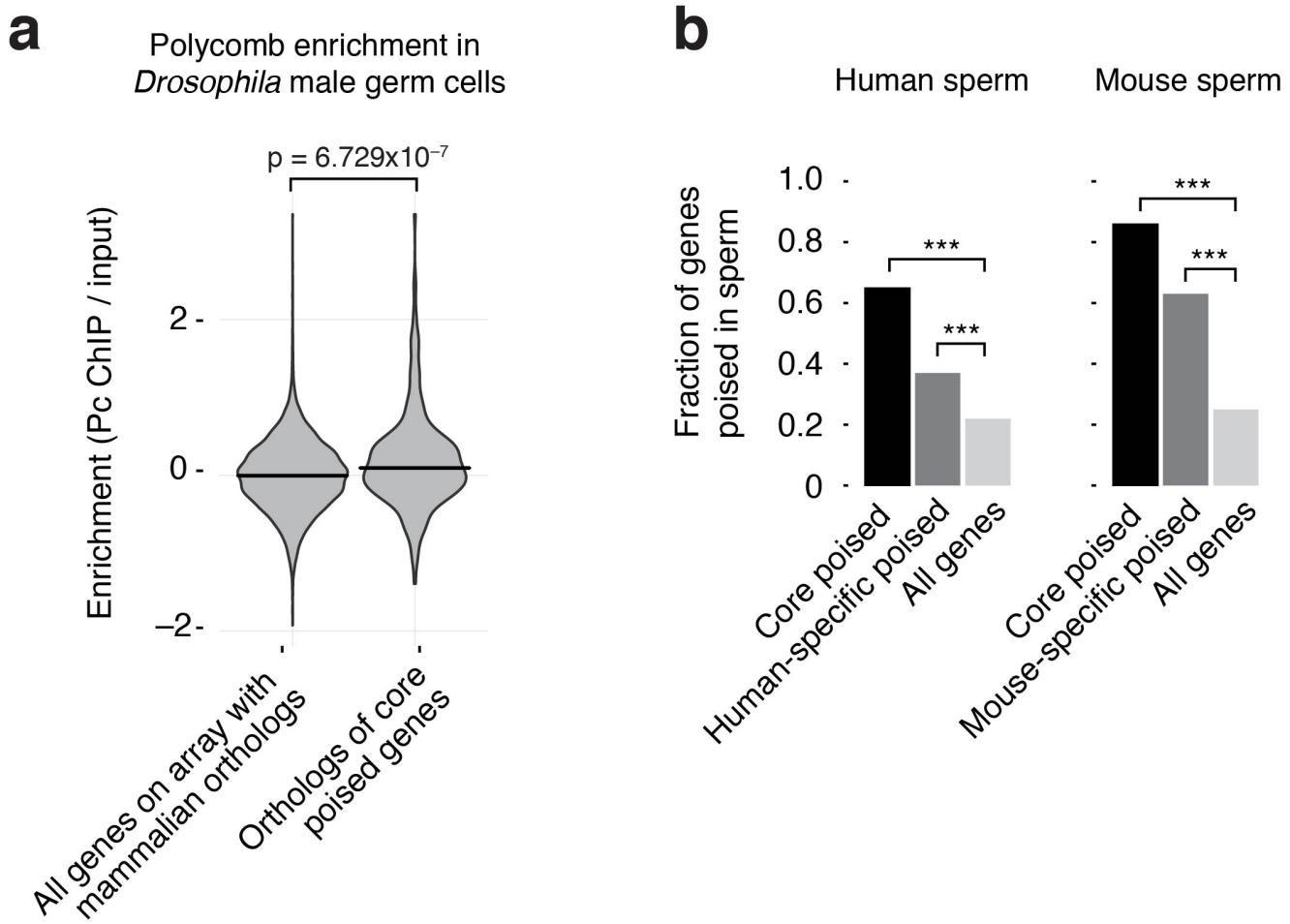
(a) Distribution of core poised genes on human chromosomes. Top, absolute numbers of poised genes on each human chromosome; middle, numbers of poised genes corrected for chromosome length; bottom, numbers of poised genes corrected for chromosome gene density, expressed as the percentage of all genes on the chromosome that are poised. (b) Conservation of promoter regions (1 kb upstream of the transcription start site) for human core poised genes, human genes with conserved retention of H3K27me3 only, human-specific poised genes, and all other genes. Horizontal bars represent the median. \*\*\* $P < 0.001$  by two-sided Welch  $t$  test. phastCons score is derived from multiple alignments of 99 vertebrate genomes to the human genome. (c) Class distribution of all transcription factors encoded in the human genome (compare to Fig. 3b).



**Supplementary Figure 9**

**Additional examples of genes poised specifically in only one of five mammalian species.**

(a) Human-specific poised genes. *HIPK2* has gained a human-specific enhancer that functions in early limb development<sup>8</sup>. *ITGB2* exhibits human-specific gain of active chromatin at its promoter in developing brain<sup>9</sup>. *LMF1* is expressed in human but not mouse or bull placenta<sup>32</sup>. (b) Mouse-specific poised gene. *Smug1* is expressed in mouse but not human or bovine placenta<sup>32</sup>. For *HIPK2*, *ITGB2*, and *LMF1*, only the region surrounding the TSS is shown to fit all tracks on the same page.



**Supplementary Figure 10**

**Poised genes in *Drosophila* germ cells and in human and mouse spermatozoa.**

(a) Polycomb enrichment at orthologs of five-mammal poised genes in the *Drosophila melanogaster* germ line. ChIP–microarray data are from ref. 50.  $P$  value was calculated by two-sided Welch t test. (b) Fraction of poised genes marked by both H3K4me3 and H3K27me3 in mature human and mouse spermatozoa. Data are from ref. 19 (human) and ref. 20 (mouse). \*\*\* $P < 1 \times 10^{-15}$  (Fisher's exact test).

## **Supplementary Note**

### **Evolution of histone modification states in pachytene spermatocytes and round spermatids**

We found that both H3K4me3 and H3K27me3 chromatin signals are evolving more rapidly in spermatids compared to spermatocytes, whereas the rate of evolution of gene expression did not differ dramatically between cell types (Figure 1d).

This observation is consistent with the developmental and evolutionary role of round spermatids relative to pachytene spermatocytes. Following meiosis, mammalian male germ cells experience dramatic changes in morphology and nuclear state as they undergo the transition to mature spermatozoa. These changes, which begin at the round spermatid stage, include transcriptional shutdown and widespread histone H4 acetylation, followed by replacement of most histones with protamines and extensive nuclear condensation.<sup>1</sup> Given these developmental changes, faster rates of evolution of histone methylation state in spermatids might result from (1) positive selection on regulation of nuclear architecture and compaction, or (2) relaxed negative selection on regulation of gene expression. The first possibility is supported by previous reports that rapid evolution of DNA packaging proteins expressed in spermatids is driven by positive selection.<sup>93,94</sup> On the other hand, round spermatids exhibit a transcriptionally permissive state, promiscuously expressing both genic and intergenic sequences<sup>95</sup>; in this context, there may be reduced selective pressure on histone methylation states involved in reinforcing transcriptional activity or repression. Indeed, we also observe that correlations between H3K4me3 and gene expression, and between H3K27me3 and gene repression, are weaker in spermatids than in spermatocytes, at least in some species (Supplementary Fig. 2).

## **References**

93. Wyckoff, G.J., Wang, W. & Wu, C.I. Rapid evolution of male reproductive genes in the descent of man. *Nature* **403**, 304-9 (2000).
94. Good, J.M. & Nachman, M.W. Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis. *Mol Biol Evol* **22**, 1044-52 (2005).
95. Soumillon, M. *et al.* Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. *Cell Rep* **3**, 2179-90 (2013).

**Supplementary Table 10: Primer sequences used for qPCR.**

Target	Forward primer (5'-3')	Reverse primer (5'-3')
mouse <i>Gdnf</i>	GTCTCCTGGATGGGATT CGG	TCCAGCACGTTCCCTCTCT
mouse <i>Foxc1</i>	CAGGACATCAAGACGGAGAAC	CTCTCGATTTGGGCAGTGT
mouse <i>Cdx2</i>	TAGCCATTCCAGTCCTCGC	TTTGTCA GTCCCTCCGCAGTA
mouse <i>Gsc</i>	TGGAGAACCTCTCCAGGA	AGGATCGCTCTGTCGTCT
mouse <i>Acr</i>	GAGGAGATGGGTTGGTGCA	GCAGGCCTCCTCGATTAAC
mouse <i>Slc2a10</i>	GGAAAGTTGTCCGCCACTG	CCGCGACTTACCCATAGCTA
opossum <i>Cdx2</i>	CCGGGGTTGAGTGTTGC	CGGCTGATTACCACTCACATC
opossum <i>Gsc</i>	TCAGCATCGACAATATTCTCGC	ACTGGCATAGAGAGAGTCCG