

1 Widespread haploid-biased gene expression in
2 mammalian spermatogenesis associated with frequent
3 selective sweeps and evolutionary conflict

4 Kunal Bhutani^{1,†}, Katherine Stansifer^{1,†}, Simina Ticau^{1,†}, Lazar Bojic¹,
5 Chloe Villani², Joanna Slisz¹, Claudia Cremers¹, Christian Roy¹, Jerry
6 Donovan¹, Brian Fiske¹ and Robin Friedman^{*1,†}

7 ¹Ohana Biosciences, Cambridge, Massachusetts

8 ²Center for Cancer Research, Massachusetts General Hospital, Boston,
9 Massachusetts

10 [†]These authors contributed equally

11 **1 Abstract**

12 Mendel's first law dictates that alleles segregate randomly during meiosis and are dis-
13 tributed to offspring with equal frequency, requiring sperm to be functionally independent
14 of their genetic payload. Developing mammalian spermatids have been thought to accom-
15 plish this by freely sharing RNA from virtually all genes through cytoplasmic bridges,
16 equalizing allelic gene expression across different genotypes. Applying single cell RNA

*Corresponding author: rfriedman@ohanabio.com

17 sequencing to developing spermatids, we identify a large class of mammalian genes whose
18 allelic expression ratio is informative of the haploid genotype, which we call genoinforma-
19 tive markers (GIMs). 29% of spermatid-expressed genes in mice and 47% in non-human
20 primates are not uniformly shared, and instead show a confident allelic expression bias
21 of at least 2-fold towards the haploid genotype. This property of GIMs was significantly
22 conserved between individuals and between rodents and primates. Consistent with the
23 interpretation of specific RNA localization resulting in incomplete sharing through cyto-
24 plasmic bridges, we observe a strong depletion of GIM transcripts from chromatoid bodies,
25 structures involved in shuttling RNA across cytoplasmic bridges, and an enrichment for
26 3' UTR motifs involved in RNA localization. If GIMs are translated and functional in the
27 context of fertility, they would be able to violate Mendel's first law, leading to selective
28 sweeps through a population. Indeed, we show that GIMs are enriched for signatures of
29 positive selection, accounting for dozens of recent mouse, human, and primate selective
30 sweeps. Intense selection at the sperm level risks evolutionary conflict between germline
31 and somatic function, and GIMs show evidence of avoiding this conflict by exhibiting
32 more testis-specific gene expression, paralogs, and isoforms than expression-matched con-
33 trol genes. The widespread existence of GIMs suggests that selective forces acting at the
34 level of individual mammalian sperm are much more frequent than commonly believed.

35 **2 Author's summary**

36 Mendel's first law dictates that alleles are distributed to offspring with equal frequency,
37 requiring sperm carrying different genetics to be functionally equivalent. Despite a small
38 number of known exceptions to this, it is widely believed that sharing of gene products
39 through cytoplasmic bridges erases virtually all differences between haploid sperm. Here,
40 we show that a large class of mammalian genes are not completely shared across these
41 bridges, therefore causing sperm phenotype to correspond partly to haploid genotype. We
42 term these genes "genoinformative markers" (GIMs) and show that their identity tends

43 to be conserved from rodents to primates. Because some GIMs can link sperm genotype
44 to function, they can be thought of as selfish genetic elements which lead to natural se-
45 lection between sperm rather than between organisms, a violation of Mendel’s first law.
46 We find evidence of this biased inheritance, showing that GIMs are strongly enriched for
47 selective sweeps that spread alleles through mouse and human populations. For genes
48 expressed both in sperm and in somatic tissues, this can cause a conflict because opti-
49 mizing gene function for sperm may be detrimental to its other functions. We show that
50 there is evolutionary pressure to avoid this conflict, as GIMs are strongly enriched for
51 testis-specific gene expression, testis-specific paralogs, and testis-specific isoforms. There-
52 fore, GIMs and sperm-level natural selection may provide an elegant explanation for the
53 peculiarity of testis gene expression patterns, which are an extreme outlier relative to all
54 other tissues.

55 **3 Introduction**

56 In diploid organisms, Mendel’s First Law dictates equal transmission of alleles to the next
57 generation, with strong selective pressure maintaining this 50:50 ratio (Crow 1979). In
58 mammalian spermatogenesis, a long stage of haploid development raises the possibility
59 of allele-biased gene expression and extensive functional variation between mature sperm
60 (Immler 2008). This could be deleterious, for example for important gene products en-
61 coded on the X chromosome that would be missing from Y-bearing sperm. However,
62 haploid sperm precursors are equipped with a mechanism for sharing of gene products:
63 cytoplasmic bridges connecting neighboring cells (Braun et al. 1989). Therefore, mature
64 mammalian sperm are thought to be functionally diploid with very rare exceptions.

65 Most examples of transmission ratio distortion (TRD), i.e. known exceptions to
66 Mendelian inheritance, are attributable to factors other than sperm heterogeneity. How-
67 ever, a handful of sperm functional differences linked to genotype have been reported.
68 The mouse *t* haplotype, a selfish genetic element transmitted at a rate of up to 99% from

69 heterozygotes, is the best understood case. The mechanism for its TRD is post-meiotic
70 expression and a lack of sharing of *t complex responder* gene products across cytoplas-
71 mic bridges, resulting in differential motility (Véron et al. 2009). Likewise, *Spam1* gene
72 products have been shown to be retained in haploid spermatids, underlying TRD in mice
73 carrying certain Robertsonian translocations (Zheng, Deng, and P. Martin-DeLeon 2001).
74 In a mouse model for Niemann-Pick disease, heterozygous knockouts of *Smpd1* have
75 sperm with functional differences in mitochondrial membrane potential associated with
76 their genotype (Butler et al. 2007). Recently, TLR7/8 inhibitors have been reported to
77 differentially affect sperm with the X or Y chromosome (Umehara, Tsujita, and Shimada
78 2019). Nevertheless, it is widely assumed that most gene products are shared between
79 mammalian gametes, erasing any allelic expression bias.

80 If, however, sperm functional variation were linked to genotype more often than com-
81 monly believed, it might provide an elegant explanation for some peculiar evolutionary
82 phenomena. Testes and spermatids in particular are extreme evolutionary outliers, hav-
83 ing far more unique tissue-specific expression patterns, tissue-specific paralogs, alternative
84 isoforms, and selective sweeps compared to other tissues (Kleene 2005). Sexual selection
85 and intragenomic conflict is often invoked to explain this bias, but haploid selection on
86 genes with transmission ratio distortion could easily have contributed (Joseph and Kirk-
87 patrick 2004). For example, alleles with beneficial effects in mature sperm might have
88 deleterious effects in somatic cells, which could drive avoidance of this conflict by evolving
89 sperm-specific paralogs or isoforms. Widespread transmission ratio distortion would be
90 difficult to observe directly due to rapid fixation of beneficial alleles and depletion of dele-
91 terious ones, but might leave traces over evolutionary timescales, altering the properties
92 of testis-expressed genes.

93 TRD enabled by retention of haploid gene products in spermatids would require
94 specific RNA localization rather than free diffusion across cytoplasmic bridges. Recent
95 methodological advances in RNA detection have revealed widespread asymmetric mRNA

96 distributions in a wide variety of cell types, including up to 70% of mRNAs during *D.*
97 *melanogaster* development (Lécuyer et al. 2007; Buxbaum, Haimovich, and Singer 2015).

98 We therefore hypothesized that many endogenous mRNAs would be transcribed in
99 haploid spermatids and incompletely shared across cytoplasmic bridges, resulting in al-
100 lelic expression bias correlating to the sperm genotype (Fig. 1A). Since mature sperm are
101 transcriptionally and translationally silent, allelic biases in mature sperm protein corre-
102 lated with the haploid genotype would have to correspond to mRNA expression biases
103 at the haploid spermatid stage. We therefore performed single cell RNA sequencing in
104 spermatids (Fig. 1B) from hybrid mice and cynomolgus macaques, quantifying allele-
105 specific biases in expression. We found surprisingly widespread chromosome-scale biases
106 in single cells allowing confident identification of genes with strong allelic expression links
107 to the genotype, which we term genoinformative markers (GIMs). We show evidence
108 for subcellular localization patterns that help explain their lack of sharing across cyto-
109 plasmic bridges, as well as evolutionary consequences consistent with sperm-level natural
110 selection.

111 4 Results

112 4.1 Many genes have allelic expression bias reflecting the hap- 113 loid genotype in spermatids

114 We first set out to identify cases of incomplete sharing of RNA across cytoplasmic bridges
115 in haploid spermatids (Fig. 1A). This would result in shared information (i.e. correlation)
116 between the allelic expression of a gene and the haploid genotype of the cell, which we
117 call genoinformative expression. Most single cell RNAseq experiments are poorly suited
118 to quantifying allele-specific expression because they do not sequence samples from fully
119 phased individuals, they only sequence a short tag from each RNA molecule (which may
120 not contain a heterozygous site), and they do so with relatively low capture efficiency. To

121 maximize the accuracy of our allele-specific quantification, we used an F1 hybrid (therefore
122 fully phased) of distantly-related inbred mouse models, C57BL/6 and PWK/PhJ, having
123 over 20 million heterozygous SNPs, compared to roughly 3 million in a human genome
124 (Fig. 1B). We digested testis tissue to isolate single cells from their cytoplasmic bridges,
125 enriched for haploid cells by flow cytometry, and performed full-length single cell RNA
126 sequencing using a slightly modified SmartSeq2 protocol optimized for sensitive RNA
127 capture (Methods).

128 Of 144 cells obtained from a single male mouse having successful RNA amplification,
129 126 passed filters as likely singlets with substantial read counts. Principal Components
130 Analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (t-SNE) revealed a
131 mixture of three cell types expressing marker genes for spermatids, spermatocytes, and
132 spermatogonia, respectively (Fig. S1A-C). Focusing on the 95 haploid spermatids, we
133 used diffusion mapping (Angerer et al. 2016) to define a pseudotime space covering their
134 differentiation process. The pseudotime ranges from early round spermatids up until the
135 point that the number of genes expressed decreases rapidly at the elongation stage, when
136 transcription arrests (Fig 1C, Fig. S1D). Late spermatid markers such as *PRM3* increase
137 in expression over this pseudotime, while spermatocyte markers such as *SYCP3* decrease
138 (Fig. 1C).

139 10,991 genes passed filters for calculation of genoinformative expression, including hav-
140 ing at least one heterozygous site and having comparable mean expression of each allele
141 (see Methods). We first focused on autosomes rather than sex chromosomes, because
142 we could use the two alleles as an internal control, yielding an easily quantifiable allelic
143 expression ratio within each cell. Visualizing allelic expression in individual haploid cells,
144 we observed strong biases across large stretches of chromosomes, but no consistent bias
145 in diploid controls (Fig. 2A, S2A). Across all haploid autosomes, there was a significant
146 correlation of allelic ratios between neighboring genes that gradually decreased with chro-
147 mosomal distance, and this correlation was completely absent in diploid controls (Fig.

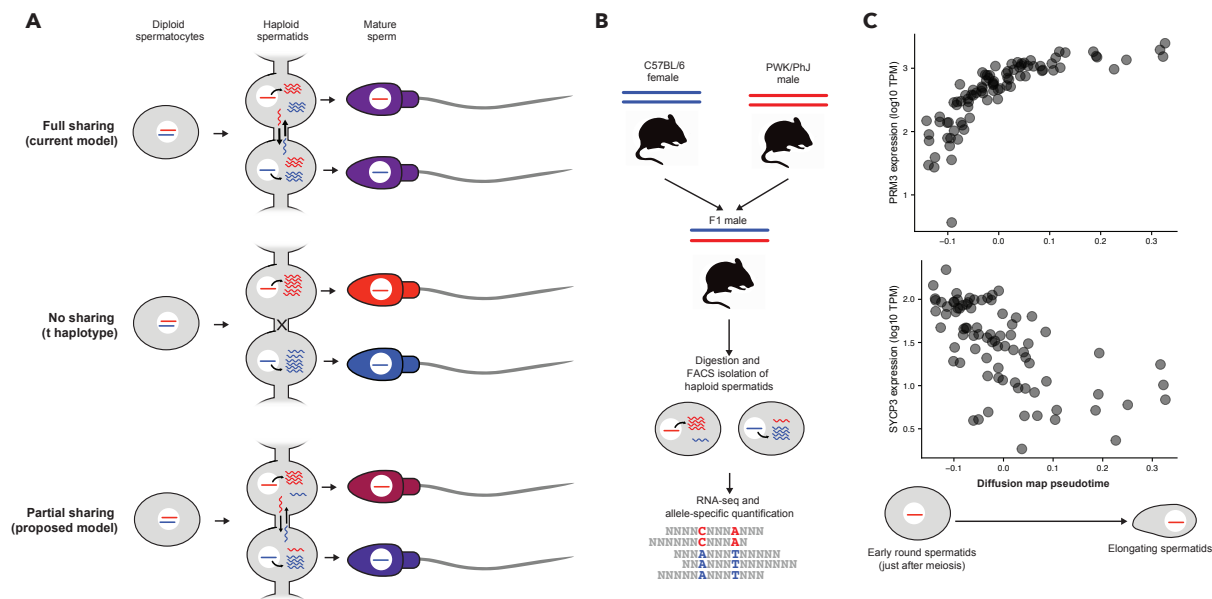
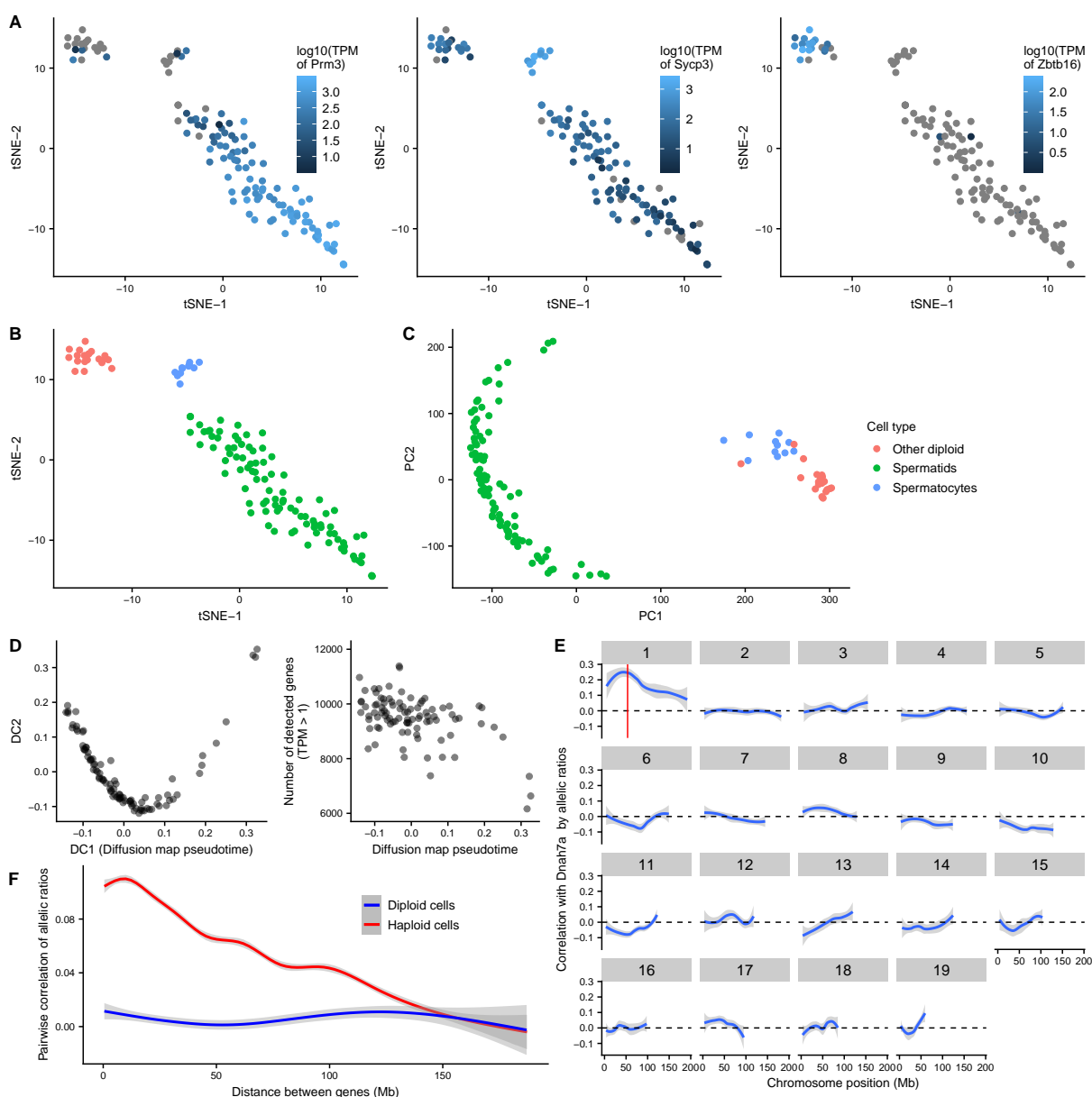


Figure 1: Single cell sequencing of haploid spermatids for assessing allelic bias. **A**) Models for allelic expression bias informative of the haploid genotype (genoinformative expression). The null hypothesis predicts complete sharing between spermatids, erasing any systematic allelic expression differences in mature sperm (top). Selfish genetic elements like the mouse *t* haplotype have virtually no sharing and lead to dramatic allelic differences in mature sperm (center), but incomplete sharing of transcripts would also lead to genoinformative expression (bottom). DNA is represented as straight lines with color representing an allele, and RNA is represented as wavy lines. Sperm color represents the degree of functional links to the allelic genotype. **B**) Experimental setup for single cell RNAseq. We crossed distantly related inbred mouse strains, digested single cells from the testis and enriched for haploid spermatids, and performed full-length RNA-seq and allele-specific quantification. **C**) Pseudotime analysis shows haploid spermatids covered a range from the early round stage (low expression of protamines) to the late elongating phase (very low expression of *SYCP3*)



Supplemental figure 1. Single cell RNAseq of haploid spermatids identifies chromosome-scale correlations in allelic bias. **A**) t-Distributed Stochastic Neighbor Embedding (tSNE) dimensional reduction for single testis cells enriched for haploid cells. Expression levels in transcripts per million (TPM) are visualized for markers of haploid spermatids (*Prm3*), spermatocytes (*Sycp3*), and spermatogonia (*Zbtb16*). **B**) Cell type annotations based on the above marker genes. **C**) Principal component analysis confirming the tSNE result, showing that all haploid spermatids were strongly distinct from diploid cells. **D**) Left: first two dimensions of diffusion map of haploid spermatids showing the first dimension captured the developmental stage well. Right: Number of genes detected per cell against the first diffusion map dimension (diffusion map pseudotime), showing a decline in those at the latest developmental stage. **E**) Illustration of chromosome-length allelic expression correlation. For one gene on chromosome 1, *Dnah7a* (located at the red line), pairwise correlation of allelic expression ratio was calculated for every gene. Plotted is a loess-smoothed average across each chromosome. Only on chromosome 1 near the *Dnah7a* locus is there a substantial average correlation. **F**) Summary of chromosome-length allelic expression correlations. For each gene, pairwise correlations of allelic expression ratios with all genes on the same chromosome were calculated. The mean correlation in haploid cells or diploid cells across all genes is plotted as a loess-smoothed average. A substantial mean correlation exists for nearby genes in haploid but not diploid cells, and decreases gradually across tens of megabases.

148 S1E-F). We reasoned that this effect could be explained by a combination of correla-
149 tion caused by widespread genoinformative expression and degradation of this correlation
150 with distance by recombination. Therefore, we designed a Bayesian probability framework
151 based on an extension of a Hidden Markov Model to infer the haploid genotype of each
152 cell including recombination breakpoints jointly with genoinformativity. Genoinformative
153 expression was modeled as emissions based on the underlying genotype and propensity of
154 an RNA to be shared across cytoplasmic bridges. Intuitively, this model shares informa-
155 tion between genes across an entire chromosome for each cell, which means that even weak
156 and noisy genoinformative expression signals in individual genes can aggregate to yield
157 robust signals across large stretches of a chromosome. The model output a probability of
158 genotypes for each cell, and a genoinformativity score for each gene representing the esti-
159 mated fraction of transcripts retained from its haploid gene expression. Visual inspection
160 confirmed that our inferred genotypes matched the observed expression biases well (Fig.
161 2A, Fig. S2A). If the inferred genotypes are accurate, the distribution of recombination
162 breakpoints should follow the known recombination density in the mouse genome. Indeed,
163 we saw a significant correlation of inferred recombination density to the published map
164 (Cox et al. 2009) with good agreement at a resolution of 10 to 20 megabases (Fig 2B,
165 S2B-C).

166 Examining for individual genes the concordance between allelic expression and haploid
167 genotype across cells, we observed a wide range of genoinformativity (Fig. 2C): Many
168 genes, like *Sycp3*, had no association between their allelic expression ratio and the inferred
169 genotype, consistent with our null hypothesis of complete sharing across cytoplasmic
170 bridges erasing allelic expression differences; some, such as *Fer1l5*, had virtually complete
171 concordance with their inferred genotype, suggesting minimal sharing across cytoplasmic
172 bridges; a larger set of genes had clear but intermediate genoinformativity, exemplified by
173 *Ccdc28a*, suggesting partial sharing through cytoplasmic bridges. To determine thresholds
174 for confident genoinformativity, we ran our Bayesian algorithm on shuffled data to create

175 an empirical background expectation under the null hypothesis of no genoinformative
176 expression (Fig. S2D-E). Thresholds of parameters for both the posterior distribution
177 of the genoinformativity score and the strength of haplotype inference were selected to
178 achieve an empirical False Discovery Rate of 10%. For convenience, genes that met
179 the criteria for confident genoinformative expression were called genoinformative markers
180 (GIMs), regardless of their effect size. Of the 10,991 genes for which we could estimate
181 genoinformativity, 4,354 (39.6%) were confident GIMs and 3,317 (30.2%) were confidently
182 not GIMs (see Methods; Fig. 2D, inset). We were unable to make a confident call for the
183 remaining 3,320 (30.2%) due to marginal signal for genoinformativity. Of the confident
184 genoinformative set, a wide range of effect sizes was seen, but 3,159 (28.8%) had at least a
185 2-fold average allelic expression ratio in favor of the allele matching the haploid genotype
186 (Fig. 2D).

187 We were surprised that as many as a third of genes were classified as strong GIMs, so
188 we sought to confirm our assumption that this corresponded to incomplete sharing across
189 cytoplasmic bridges. The chromatoid body is a membraneless organelle (a phase-separated
190 condensate) in germ cells that has been shown to shuttle RNA across cytoplasmic bridges
191 to facilitate sharing (Fig. 2E inset; Ventelä, Toppari, and Parvinen 2003). We found
192 that a published set of genes enriched in the chromatoid body (Meikar et al. 2014) had
193 far lower genoinformativity scores than other genes (Fig. 2E), and that there were fewer
194 GIMs enriched in the chromatoid body than expression-matched controls (Fig. S5C). This
195 confirms that GIMs have different subcellular localization of their RNAs from non-GIMs.

196 **4.2 GIMs have specific subcellular localization resulting in in-** 197 **complete sharing across cytoplasmic bridges**

198 To identify what mechanisms might be responsible for the differential localization of GIMs,
199 we compared GIMs to non-GIM controls that were matched for expression across spermiogenesis
200 as closely as possible (Fig. S5A, Table S3-4, Methods). Most eukaryotic mRNA

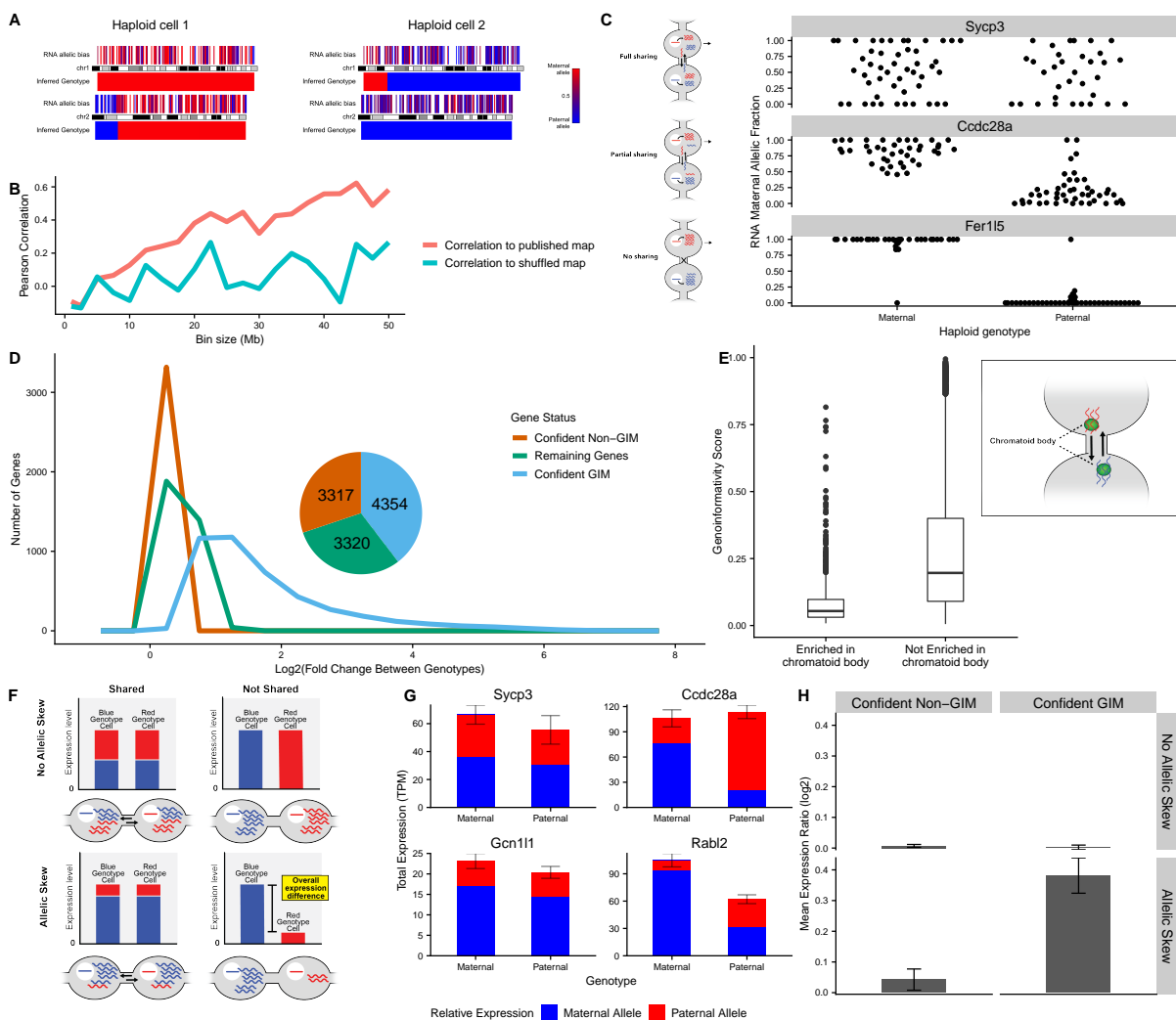
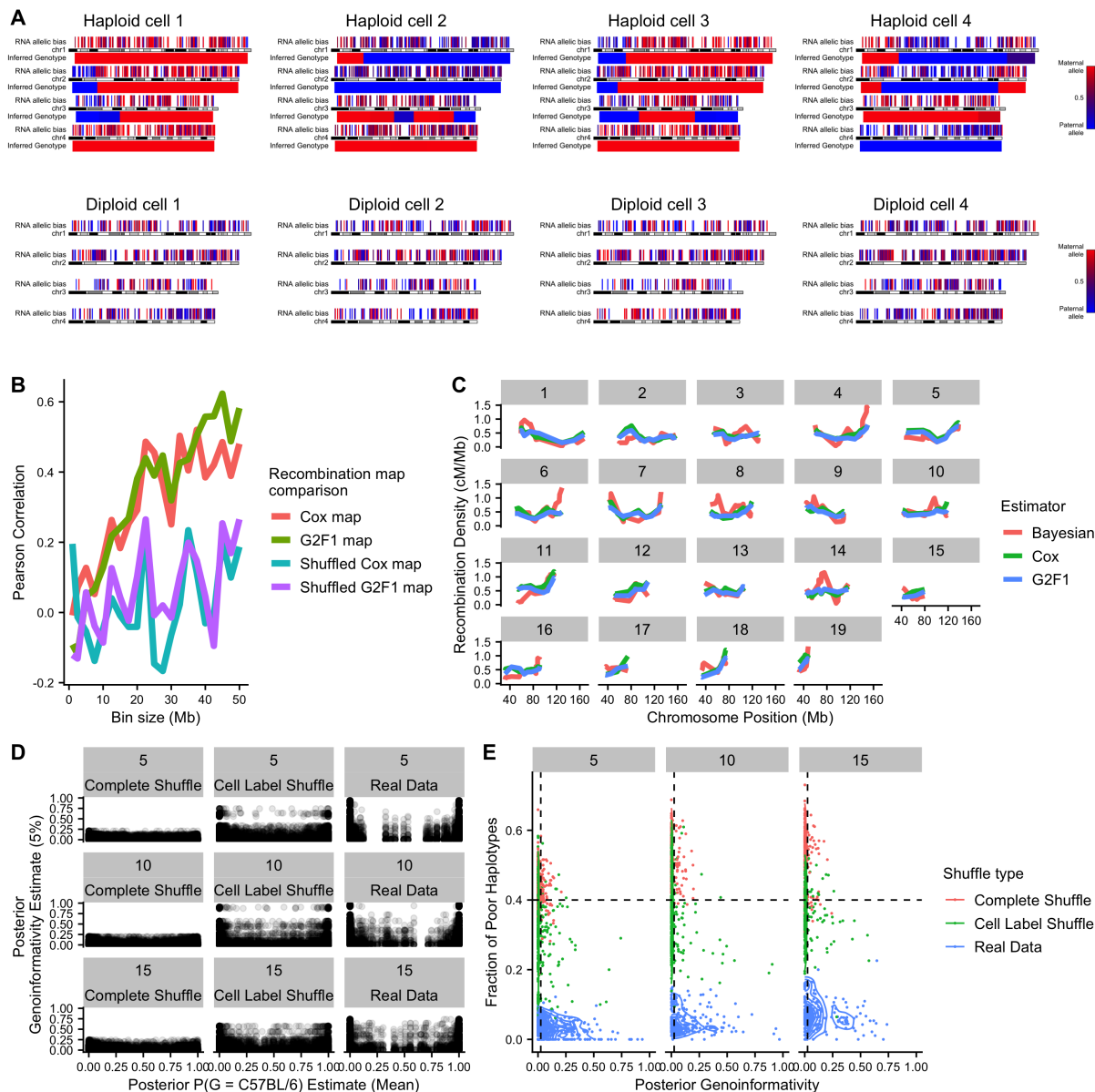


Figure 2: A large fraction of mouse genes exhibit genoinformative expression. **A**) Visualization of allelic bias in the first two chromosomes of two representative haploid cells. Each expressed gene is represented as a vertical line with color representing its allelic ratio (red for more maternal allele, blue for more paternal). Below each chromosome is the genotype inferred by our Bayesian method. **B**) Correlations between inferred recombination densities and a published mouse recombination map (Cox et al. 2009) or a control with recombination densities shuffled between all bins. As bin sizes decrease below about 20 megabases, the variance in our inferred rates increases, causing a degradation of our signal to noise ratio. **C**) Example genes illustrating differing levels of genoinformative expression (right) with their models of sharing (left). *Sycp3* exhibits no association with the haploid genotype, *Ccdc28a* exhibits a strong but incomplete association between the inferred genotype and the expressed allele, and *Fer15* exhibits a near-perfect correlation with the inferred genotype. **D**) GIM classification of all genes. Histogram shows the log₂ of the expression ratio between the concordant allele (i.e. matching the genotype) over the discordant allele on average across cells. Inset: the total number of genes classified in each category of genoinformative expression. **E**) Genes with mRNAs enriched in the chromatoid body have significantly lower genoinformativity scores. Genoinformativity scores range from zero to one and represent the estimated fraction of transcripts originating from a cell's haploid transcription. Inset: depiction of the chromatoid body's role in shuttling mRNAs across cytoplasmic bridges in haploid spermatids. **F**) A model for how allelic skew (e.g. due to eQTLs) interacts with genoinformative expression. Only genes with both allelic skew and genoinformative expression (not shared) have their mean expression level correlated to the haploid genotype. **G**) Example genes matching the categories in (F). Only *Rabl2* has a significant mean expression difference ($p = 1.5 \times 10^{-5}$, Wilcoxon test). **H**) Summary of expression differences (log₂ ratio of genotype concordant with skew to discordant) in all genes in each of the four combinations listed. Only with both allelic skew and GIMs is there an expression difference between cells of differing genotypes.



Supplemental figure 2. Joint inference of genotype and genoinformativity. **A**) Visualization of allelic bias in the first four chromosomes of randomly selected haploid cells and randomly selected diploid cells. Each expressed gene is represented as a vertical line with color representing its allelic ratio (red for more maternal allele, blue for more paternal). Below each chromosome is the genotype automatically inferred by our Bayesian method. **B**) Correlations between inferred recombination densities and two published mouse recombination maps (Cox et al. 2009; Liu et al. 2014) or corresponding controls with recombination densities shuffled between all bins. **C**) Recombination densities across each chromosome (calculated over a 20Mb window) implied by the Bayesian recombination frequencies or for each of the two published recombination maps. **D**) Inferred genotype and genoinformativity for real haploid data and two shuffle types: one permuting both gene and cell labels (complete shuffle) and one permuting only cell labels. Each point is a gene/cell pair, with genotype estimate (x-axis) being a property of the specific gene in a specific cell, and 5% lower bound of genoinformativity (y-axis) being a property of the gene (constant across cells). Three representative chromosomes are plotted (5, 10, and 15). Real data more often have confident genotype estimates and high genoinformativity (upper left and upper right of graph). The cell label shuffle is quite conservative because the genotype structure is maintained, and only the genoinformative expression is randomized. **E**) Summary of the data from (D) illustrating thresholds for calling confident GIMs (dashed lines). Each point is a gene, with poor haplotypes defined as those with less than 95% probability of a genotype. 5% lower bound of posterior genoinformativity probability is plotted on x-axis.

201 localization is dictated by RNA-binding proteins via sequence motifs in 3' UTRs (An-
202 dreassi and Riccio 2009), so we performed an enrichment analysis for known motifs of
203 RNA-binding proteins that are expressed in spermatids. We identified 26 motifs signif-
204 icantly enriched in GIMs relative to controls, and zero significantly depleted in GIMs
205 (Table S5).

206 Similarly, a gene ontology enrichment analysis identified strong enrichment for GIMs
207 for specific protein localizations, especially membrane associations and axoneme or other
208 tail localizations (Table S6). To further refine this result, we performed an enrichment
209 analysis with a comprehensive localization database (Binder et al. 2014). This revealed a
210 strong enrichment for genes with annotated localization in neurons, including both den-
211 drites and axons (Table S7), probably reflecting the fact that subcellular RNA localization
212 has been best studied in neurons but is governed by principles applicable across cell types
213 (Ryder and Lerit 2018). Together, these data suggest a mechanism for genoinformativity
214 whereby RNA-binding proteins bring some mRNAs to specific subcellular locations distal
215 from chromatoid bodies, thus partially avoiding sharing across cytoplasmic bridges.

216 As independent confirmation of our incomplete sharing model for GIMs, we sought
217 to use the much larger set of RNAseq reads that did not overlap a heterozygous site but
218 could be used for estimating overall expression levels. GIMs have allelic expression biases
219 based on the haploid genotype, but because 50% of cells have each genotype, GIMs do
220 not necessarily have a mean allelic expression bias when averaging across many cells (here
221 called allelic skew). However, many genes have a mean allelic skew for other reasons, for
222 example due to expression quantitative trait loci (eQTLs) wherein a genetic variant has
223 differential effects on the expression of a gene. The incomplete sharing model predicts that
224 genes may have different expression levels in spermatids with the paternal versus maternal
225 genotype, but only when they have both an allelic skew and genoinformative expression
226 (Fig. 2F). To illustrate this point, *Sycp3* (Non-GIM, no allelic skew), *Ccdc28a* (GIM,
227 no allelic skew), and *Gcn11l* (Non-GIM, 2.7-fold allelic skew) all have no difference in

228 mean total expression from the maternal and paternal genotype cells (Fig. 2G). However,
229 *Rabl2*, which has a 3.0-fold allelic skew and genoinformativity score of 0.45 has a significant
230 difference in expression between the two spermatid genotypes ($p = 1.2 \times 10^{-5}$, t test).
231 Across all genes, we observe that the expression level of GIMs with allelic skew is linked
232 to the haploid genotype in the expected direction, but not for non-GIMs and not for
233 genes without overall allelic skew (Fig. 2H). Therefore both allele-informative and non-
234 allele-informative RNAseq reads support the identity of GIMs and the incomplete sharing
235 model.

236 **4.3 Sex chromosome genes also exhibit genoinformative expres-** 237 **sion**

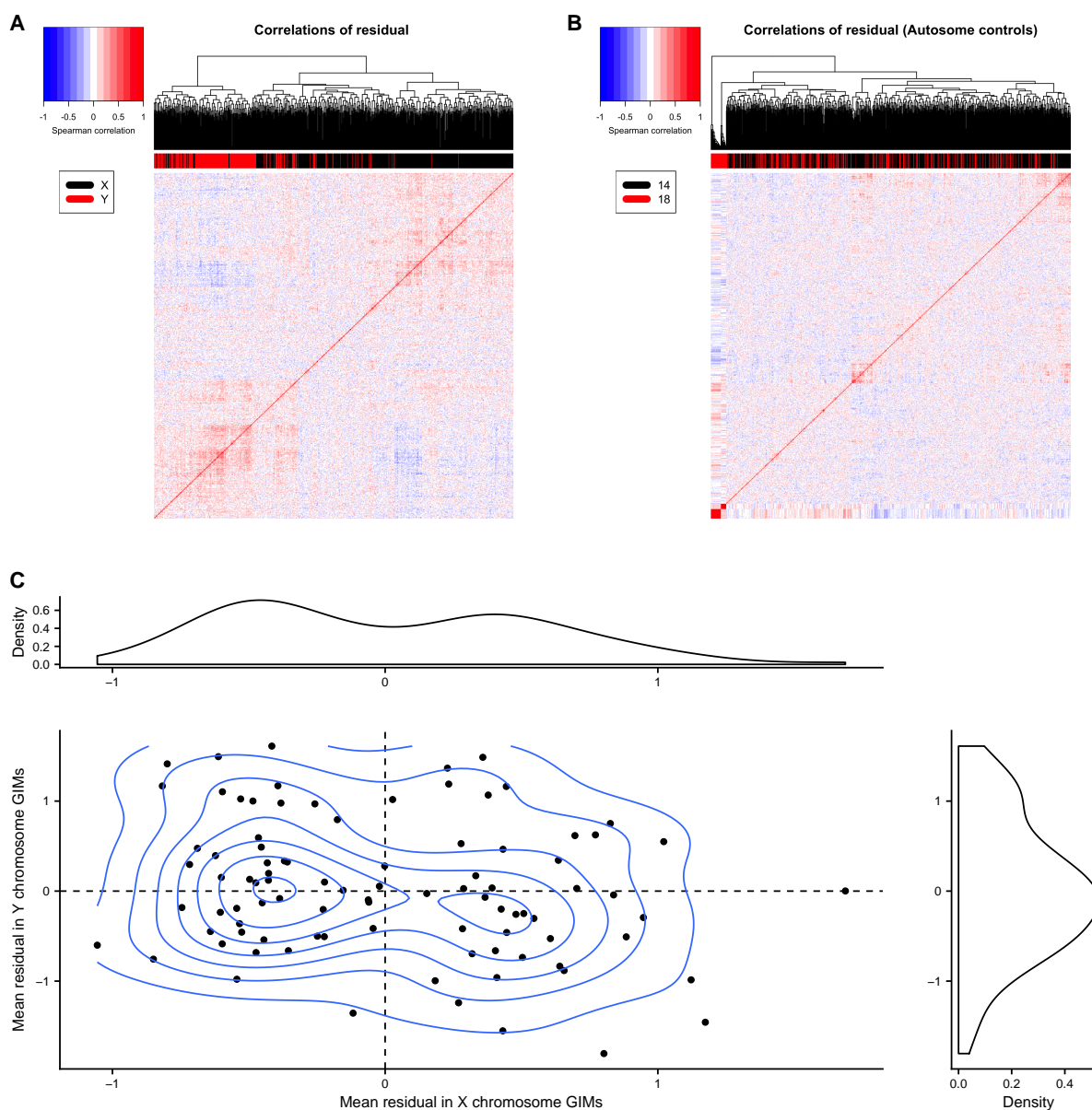
238 Although our Bayesian method for inferring genotype and genoinformativity cannot be
239 applied to sex chromosomes due to the lack of allelic expression data, genoinformative
240 expression of sex chromosome genes would provide an elegant explanation for models
241 of sex ratio distortion in mice (Cocquet et al. 2012; Eep, Pji, and Ellis Email n.d.).
242 We therefore developed a separate method to identify sex chromosome GIMs based on
243 variation in expression levels rather than in allelic ratios. We started by reasoning that
244 X chromosome GIMs should have correlated expression and be anticorrelated with Y
245 GIMs. Because expression levels in any given spermatid can be strongly influenced by
246 developmental stage, we first corrected for the position in the diffusion map pseudotime.
247 Clustering genes by pairwise correlation after correction, we identified two distinct clusters
248 that corresponded overwhelmingly to the X and Y chromosome, respectively (Fig. S3A).
249 In contrast, performing the same analysis on autosomal controls yielded no similar clusters
250 (Fig. S3B). We selected putative GIMs from these distinct clusters that displayed strong
251 correlation signals (see methods), resulting in 63 X GIMs and 84 Y GIMs (Table S2).
252 Spermatids tend to have high or low mean levels of X GIMs, but not intermediate levels
253 (Fig. S3C). Therefore, sex chromosomes appear to be no exception to the prevalence of

254 genoinformative expression, at least on a quantitative level.

255 **4.4 Genoinformativity is conserved between individuals and across** 256 **species**

257 So far, we have only considered mice with one genetic background, so we next asked
258 whether the phenomenon of widespread genoinformative expression extends to other
259 mammals. We dissociated testes from two outbred cynomolgus primates (*Macaca fas-*
260 *cicularis*), isolated haploid spermatids and performed single cell RNAseq. Cynomolgus
261 monkeys have the advantage of being highly heterozygous, with ~ 13 million heterozy-
262 gous SNPs per individual, compared to ~ 3 million for humans. Because our method for
263 inferring genotypes relies on sharing information across entire chromosomes, we required
264 fully phased chromosomes to quantify genoinformative expression. We therefore combined
265 two phasing methods: a dense, short-range phasing using linked read sequencing, and a
266 sparse, long-range phasing using whole genome sequencing of single haploid spermatids
267 (Fig. 3A). Combining the two sources of information led to densely phased chromosomes
268 for each individual, resulting in 11,654,918 and 10,131,178 phased sites in Cynomolgus 1
269 and 2, respectively (Fig. S4A).

270 We were able to quantify allelic expression of a smaller number of genes for cynomolgus
271 spermatids than for mice (7,590 and 4,557 for the two cynomolgus compared to 10,991
272 in mice), mostly due a smaller number of heterozygous sites. Nevertheless, we observed
273 comparable quality of our genotype inference, including significant correlation of inferred
274 recombination rates between individuals, an expression skew in GIMs with allelic expres-
275 sion skew, and substantial differences between real and shuffled data (Fig. S4B-E). Again
276 using an empirical false discovery rate of 10% in each individual, we classified 50.3% and
277 52.3% of spermatid-expressed genes as confident GIMs, respectively (Fig. 3B). The effect
278 sizes were comparable to those seen in mice, with 44.6% and 43.3% of spermatid-expressed
279 genes having at least a 2-fold average expression difference between alleles in favor of the



Supplemental figure 3. Sex chromosome GIMs. **A**) Heatmap of pairwise correlations of sex chromosome genes. Correcting for developmental stage (fitting the expression to the diffusion pseudotime position), the residuals of the log expression levels are correlated between all pairs of sex chromosome genes. Two anticorrelated clusters appear, one principally on the X chromosome (black lines above the heatmap), one principally on the Y chromosome (red lines above the heatmap). **B**) Heatmap of pairwise correlations as in (A), but for autosomal control chromosomes with similar numbers of spermatid-expressed genes (chromosomes 14 and 18). No similar broad clusters appear. **C**) Cells have bimodal expression of putative X chromosome GIMs. For each cell, the mean residual log expression across putative X GIMs and Y GIMs is plotted, with density contours. Density plots on the margins show the kernel density of the mean residual for X GIMs (top) and for Y GIMs (right). Most cells have either a high or a low average expression of X chromosome GIMs, but not intermediate. Cells that have high X GIM expression tend to have lower expression of Y GIMs, and vice versa.

280 haploid genotype. In total, 47.3% of genes that could be quantified met this threshold in
281 either of the two individuals.

282 Because the two individuals have different heterozygous sites, only 2,366 genes had
283 quantified genoinformativity in both. Among these genes, those that were classified as
284 a confident GIM in one individual had far higher genoinformativity scores in the other
285 individual, and those classified as a confident non-GIM had far lower genoinformativity
286 scores in the other individual ($p < 2.2 \times 10^{-16}$; Fig. 3C). This suggests that within a
287 species, the property of genoinformativity is highly consistent. To look across far larger
288 evolutionary timescales, we compared cynomolgus genes to their orthologs in mouse with
289 a genoinformativity score in each ($n = 2,838$). Confident GIMs in cynomolgus had higher
290 genoinformativity in mouse than confident non-GIMs ($p < 2.2 \times 10^{-16}$; Fig. 3C), al-
291 though the relationship was weaker than within a single species. This suggests that the
292 features that confer incomplete sharing across cytoplasmic bridges evolve slowly, so that
293 the identities of GIMs tend to be maintained across evolutionary timescales.

294 **4.5 GIMs show signs of sperm-level natural selection and evo-** 295 **lutionary conflict**

296 The substantial fraction of genes having genoinformative expression at the RNA level is
297 surprising, but it does not necessarily imply functional differences in sperm. For example,
298 proteins could be shared across cytoplasmic bridges, nullifying any allelic differences at
299 the RNA level. In contrast, if GIMs lead to functional differences in sperm linked to their
300 genotype, sperm-level natural selection could result in increased evolutionary forces (both
301 purifying and positive selection) acting on GIMs compared to other genes. Given that the
302 identities of GIMs have been maintained across an appreciable evolutionary distance, we
303 reasoned that functional differences in GIMs would lead to detectable signatures in the
304 genome even if they rarely arise. Selective sweeps entail a beneficial allele experiencing
305 positive selection and rapidly reaching fixation in a population, which leaves a signal that

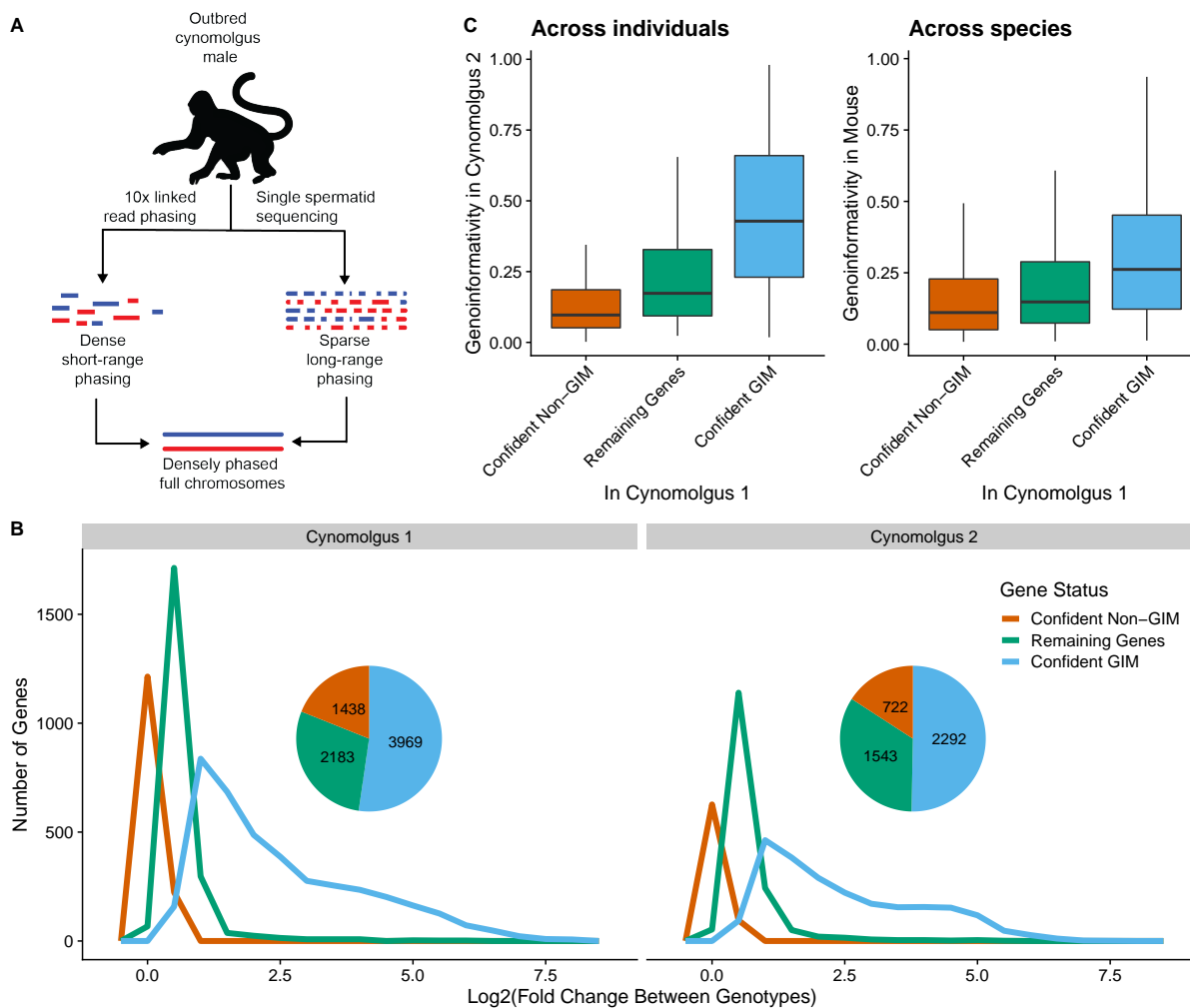
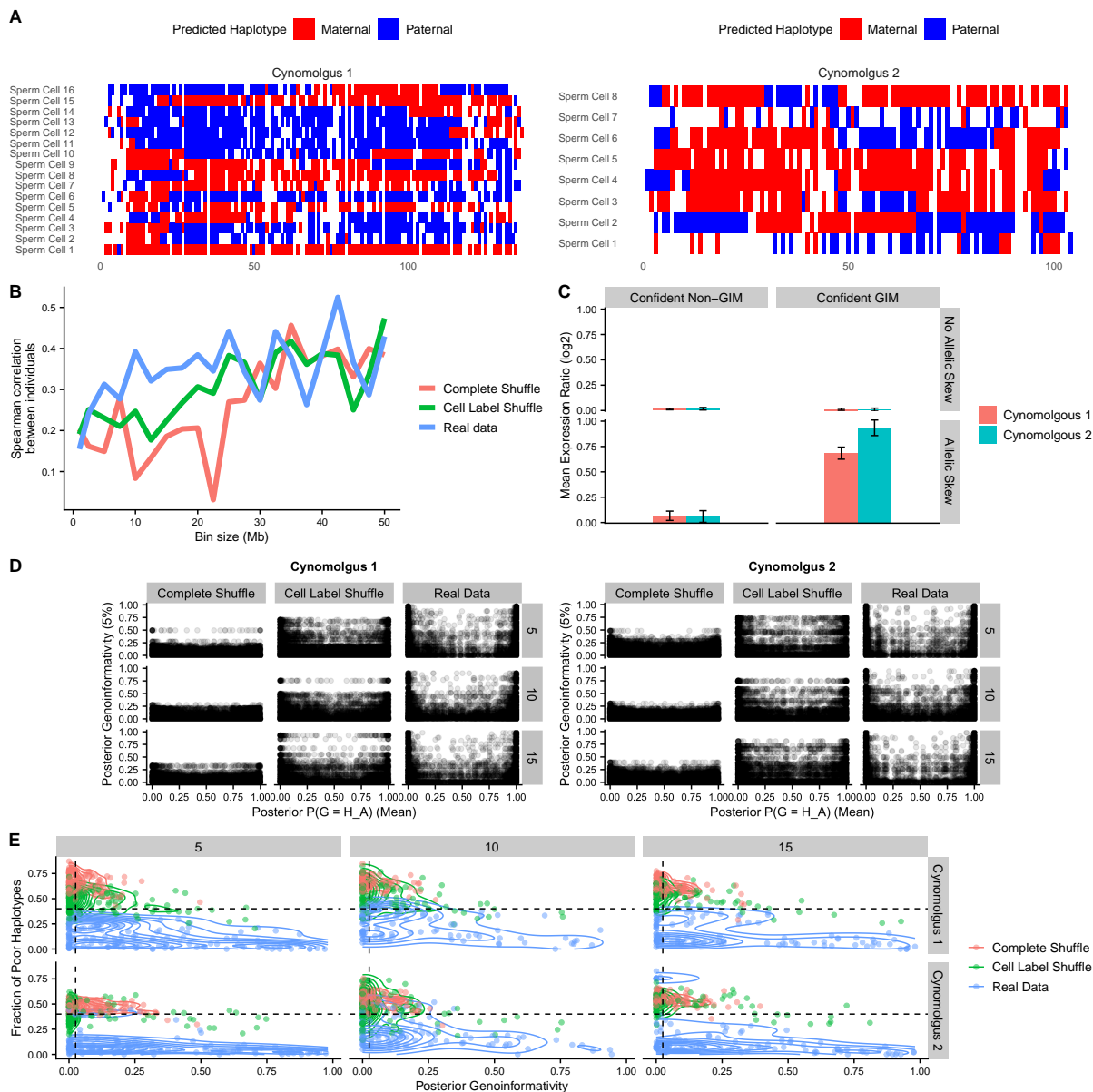


Figure 3: GIMs are conserved between individuals and across species. **A**) Fully phased chromosomes were generated directly from outbred cynomolgus individuals by computationally merging phasing maps from two experimental techniques: short-range phasing from 10x Genomics linked read sequencing, and long-range phasing from whole genome sequencing of several single haploid spermatids. **B**) Genoinformative expression classification of all genes as in Fig. 2D, for each of two cynomolgus individuals. Histogram shows the log₂ of the expression ratio between the concordant allele and the discordant allele on average, where the concordant allele matches the inferred genotype. Inset: the total number of genes classified in each category of genoinformative expression. **C**) Conservation of genoinformativity. Genes are categorized based on their genoinformativity classification in Cynomolgus 1 (x axis), and genoinformativity is plotted for these genes in Cynomolgus 2 (left) or orthologs in mouse (right). Genoinformativity scores range from zero to one and reflect the degree of shared information with genotype.



Supplemental figure 4. Cynomolgus primate genotype and genoinformativity inference. **A)** Single cell DNA sequencing data is displayed as phasing blocks called by the 10x Chromium pipeline for chromosome 1. Blocks are assigned to parental chromosomes based on the single cell sequencing data using the algorithm described in the methods section. The resulting patterns show 1-2 recombinations per cell with very few discordant (incorrectly assigned) blocks. **B)** Spearman correlation between recombination densities inferred for the two individuals. Shuffled data showed lower correlations at low to moderate bin sizes. **C)** Summary of expression differences (\log_2 ratio of genotype concordant with skew to discordant) in all genes in each of the four combinations listed. Only with both allelic skew and GIMs is there an expression difference between cells of differing genotypes, matching the results in mouse. **D)** Inferred genotype and genoinformativity for real haploid data and two shuffle types: one permuting both gene and cell labels (complete shuffle) and one permuting only cell labels. Each point is a gene/cell pair, with genotype estimate (x -axis) being a property of the specific gene in a specific cell, and 5% lower bound of genoinformativity (y -axis) being a property of the gene (constant across cells). Three representative chromosomes are plotted (5, 10, and 15). Real data more often have confident genotype estimates and high genoinformativity (upper left and upper right of graph). The cell label shuffle is quite conservative because the genotype structure is maintained, and only the genoinformative expression is randomized. **E)** Summary of the data from (D) illustrating thresholds for calling confident GIMs (dashed lines). Each point is a gene, with poor haplotypes defined as those with less than 95% probability of a genotype. 5% lower bound of posterior genoinformativity probability is plotted on x -axis.

306 can be detected by a variety of statistical tests over patterns of variation in the genome.
307 We cross-referenced a set of selective sweeps in wild mouse populations (Staubach et
308 al. 2012) with GIMs and non-GIM controls, either randomly selected from spermatid-
309 expressed genes or matched for expression patterns across spermiogenesis. The GIMs
310 were found in significantly more selective sweep regions than expected by chance ($p =$
311 3×10^{-25}) corresponding to an excess of 47 ± 4.6 selective sweeps putatively attributable
312 to genoinformativity (Fig 4A, left). Although we do not know of studies of selective
313 sweeps in cynomolgus, we took advantage of abundant predictions of selective sweeps in
314 humans by examining orthologs of cynomolgus GIMs and non-GIMs. Using a set of human
315 selective sweeps (Refoyo-Martínez et al. 2019), we find a significant enrichment of GIMs
316 ($p \leq .013$) corresponding to 9.4 ± 4.2 sweeps putatively attributable to genoinformativity
317 (Fig 4A, right). We corroborated this enrichment for GIMs in a wide variety of tests for
318 selective sweeps in humans and primates on multiple timescales (Fig. S5B). Examining
319 an even larger set of tests for natural selection using 1000 genomes project data (Pybus
320 et al. 2014), we found significant enrichments in a majority of tests (Fig. S5D). Together,
321 this indicates that GIMs are associated with an increased rate of positive selection over
322 evolutionary time.

323 Sperm-level natural selection poses an evolutionary conundrum: due to its highly
324 specialized function, what is good for the sperm is not necessarily good for the organism.
325 In other words, selection for a beneficial allele in sperm may decrease overall fitness if
326 the allele is deleterious in a somatic cell context (Fig. 4B). Over evolutionary time,
327 this conflict might make genoinformative expression deleterious for genes with somatic
328 functions, but not for genes uniquely expressed in male reproductive tissue. Supporting
329 this hypothesis, we see that GIMs are more likely to be testis-specific in both mouse
330 ($p < 10^{-22}$) and human ($p = 0.006$; Fig. 4C). When it arises, the evolutionary conflict
331 caused by sperm-level selection will cause evolutionary pressure for separating functions
332 for the gene in germ and somatic cells. Examples of this evolutionary pattern include gene

333 duplication followed by subfunctionalization (Fig. 4B), and testis-specific gene isoforms.
334 As predicted, GIMs are significantly enriched in paralog families that are predominantly
335 testis-expressed in both mouse ($p \leq 6.7 \times 10^{-12}$; Fig. 4D, left) and human ($p \leq 0.0007$; Fig.
336 4D, right). Human GIMs are also enriched testis-specific isoforms ($p \leq 1.9 \times 10^{-14}$; Fig.
337 4E, right), and although we are not aware of similar quality isoform-level mouse datasets,
338 mouse GIMs are significantly more likely to have testis-specific exons ($p \leq 3.7 \times 10^{-9}$;
339 Fig. 4E, left).

340 Each of these lines of evidence implies that GIMs with these properties are enriched for
341 causing functional differences in sperm, which would require incomplete sharing of proteins
342 across cytoplasmic bridges. In the mouse *t haplotype*, this occurs in part by translating
343 a protein late in spermiogenesis, as cytoplasmic bridges start to break down (Véron et
344 al. 2009). We therefore predicted that GIMs enriched for causing functional differences
345 in sperm would also be enriched in late translation of their proteins compared to other
346 GIMs. Examining a polysome profiling dataset across mouse spermatogenesis (Iguchi,
347 Tobias, and Hecht 2006), mouse GIMs that were functional candidates based on selective
348 sweeps, testis-specific expression, or testis-specific paralogs, were indeed enriched for late
349 translation ($p = 0.045$, 1.4×10^{-12} , 0.00045 , Fisher's exact test; Fig. 4F). However, we
350 did not see enrichment in late translation for GIMs that had testis-specific exons. These
351 results suggest that late translation of GIMs is one mechanism by which they may lead to
352 sperm-level functional differences, causing a higher rate of selective sweeps and avoidance
353 of evolutionary conflict.

354 5 Discussion

355 Here we have shown that a large fraction of spermatid-expressed genes are not completely
356 shared between haploid spermatids, resulting in allelic expression that is linked to the
357 haploid genotype, which we call genoinformative expression. Our model for the mechanism
358 for this genoinformative expression is subcellular localization of RNAs, occurring through

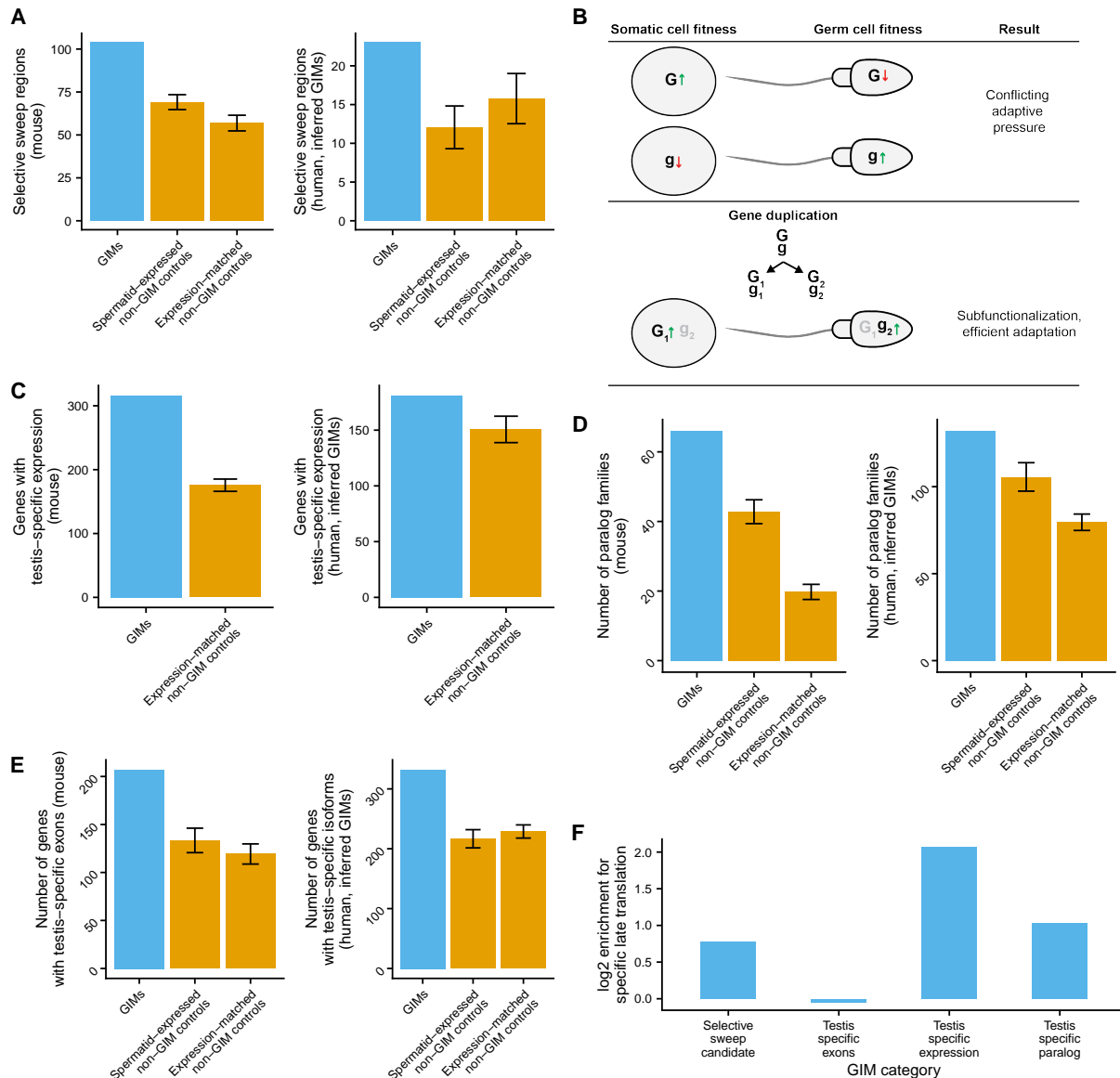
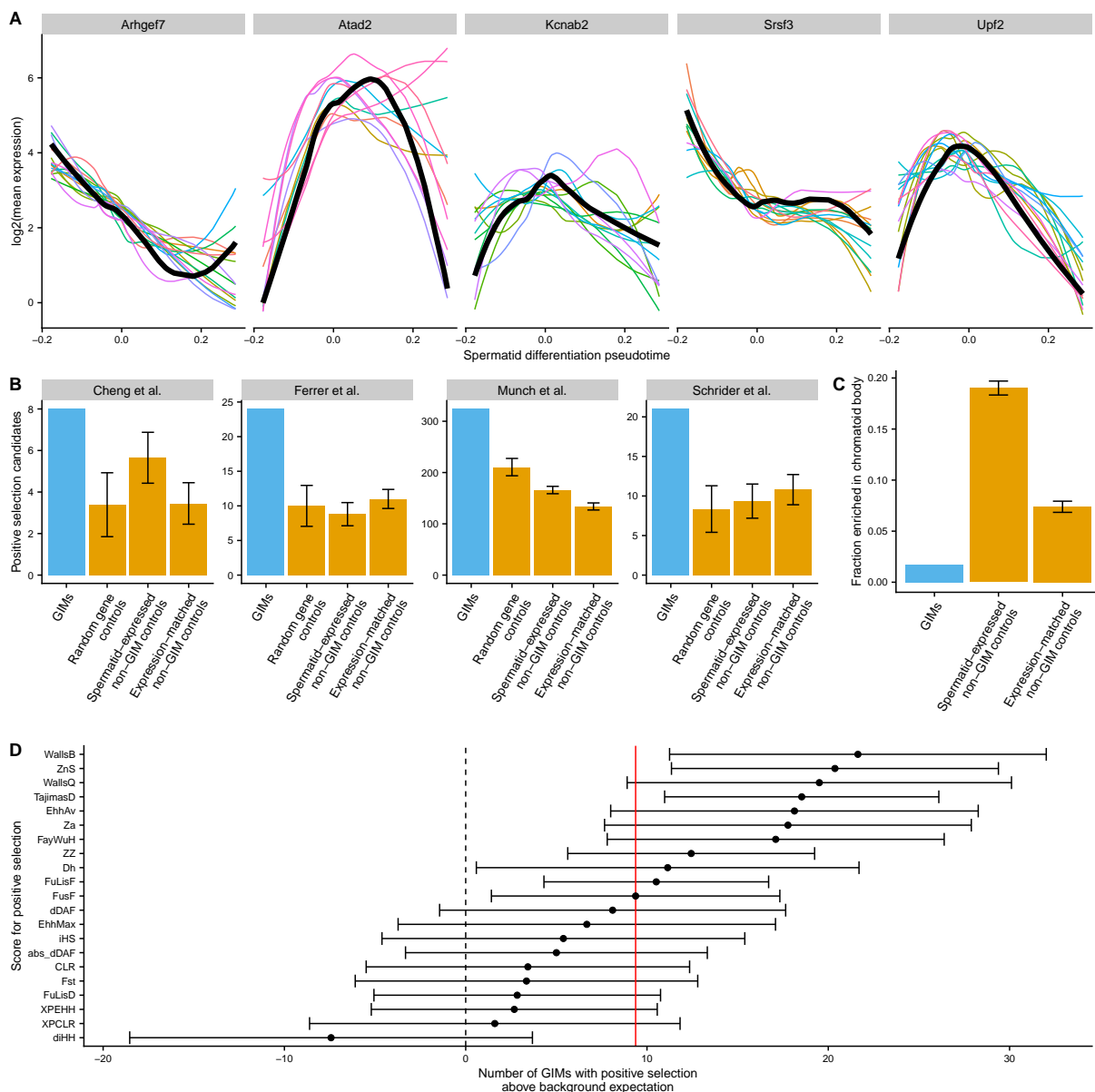


Figure 4: GIMs are associated with sperm-level natural selection and evolutionary conflict. **A**) GIMs are enriched in selective sweep regions in mouse (Staubach et al. 2012) and human (Refoyo-Martinez et al. 2019). Human GIMs were inferred from cynomolgus orthologs. GIMs were compared to control sets (orange bars), either selected from all spermatid-expressed confident non-GIMs, or confident non-GIMs matched to GIMs by their spermatid expression trajectory. **B**) Model for evolutionary conflict between sperm-level and organism-level natural selection. The gene has one allele with beneficial effect in somatic cells but detrimental effect in sperm (G) and one allele with the reverse pattern (g), resulting in positive selection for g at the sperm level, but negative selection at the organism level. A resolution to conflict can be achieved by duplication into two genes, G_1/g_1 expressed in somatic cells and G_2/g_2 expressed in sperm. Selection will then favor the G_1 and g_2 alleles, with no detrimental effects at either level. **C**) GIMs are enriched for testis-specific expression in mice and human, defined as 10-fold higher expression than any other tissue. GIMs were only compared to non-GIMs matched for spermatid expression trajectory, because testis-specific expression is by definition dependent on spermatid expression level. **D**) GIMs represent a higher number of paralog families than non-GIMs in mice and humans. Controls as in (A). **E**) GIMs are enriched in testis-specific isoforms in humans and testis-specific exons in mice. Controls as in (A). **F**) GIMs that are functional candidates are enriched for specific late translation. The GIMs are taken from the blue bars in panel A, C, D, and E, respectively. GIMs in each functional category are compared with GIMs not in that category, and the proportion with specific late translation was calculated. The log₂ of the ratio of these proportions is plotted.



Supplemental figure 5. GIM functional characterization. **A**) Illustration of expression-matched control selection for representative GIMs. Thick black lines represent log₂ of the loess fit of the expression (in TPM) of GIMs across the spermatid differentiation diffusion pseudotime. Colored lines represent the same loess fit for the 20 genes selected as controls for this gene based on their expression pattern and dropout rate. **B**) The number of positive selection (selective sweep) candidates from several publications (Schrider and Kern 2016; Ferrer-Admetlla et al. 2014; Cheng, Racimo, and Nielsen 2019; Munch et al. 2016) overlapping GIMs or several types of controls. Error bars represent the mean \pm standard deviation over the 20 control sets of mock GIMs. GIMs are enriched for selective sweeps in all cases ($p < 0.0276$, $p < 1.01 \times 10^{-6}$, $p < 9.65 \times 10^{-6}$, $p < 8.60 \times 10^{-6}$, respectively). **C**) The fraction of genes overlapping the genes annotated as enriched in the chromatoid body (Meikar et al. 2014) overlapping with each gene category. Bars represent mean \pm standard deviation over the 20 control sets of mock GIMs. **D**) Enrichment for GIMs in positive selection candidates based on raw scores for positive selection calculated based on 1000 genomes project data. The background expectation was calculated using the expression-matched non-GIM control set, and error bars represent the mean \pm twice the standard deviation of these controls.

359 RNA-binding protein motifs in the 3' UTRs or other mechanisms, resulting in depletion
360 of GIMs from the chromatoid body (which facilitates sharing across cytoplasmic bridges).
361 GIMs are substantially conserved across populations and evolutionary timescales, so we
362 predict these mechanisms are conserved as well.

363 In light of this finding, a number of cases of sperm-level functional differences in
364 the literature can be putatively attributed to GIMs (Conway et al. 1994; P. A. Martin-
365 DeLeon et al. 2005; Butler et al. 2007; Véron et al. 2009; Cocquet et al. 2012; Alavioon
366 et al. 2017; Nadeau 2017; Umehara, Tsujita, and Shimada 2019). Despite the growing
367 number of examples of sperm-mediated transmission ratio distortion, it has been widely
368 assumed these are isolated cases and that mammalian sperm are functionally diploid as
369 a rule. The fact that GIMs were so common (over a third of spermatid-expressed genes)
370 surprised us, and suggests that many more cases of sperm-level functional heterogeneity
371 based on genotype will be found.

372 Mendel's first law dictates that alleles of genes are inherited with equal probability,
373 requiring sperm to be functionally equivalent regardless of their haploid genotype. We
374 believe that remains the case for the majority of genes in mammals at any given time,
375 since transmission ratio distortion has not been commonly observed. However, we show
376 that over evolutionary timescales, GIMs are associated with an increased rate of selective
377 sweeps, suggesting selection at the level of sperm based on functional differences linked
378 to alleles. At first glance, reconciling the sperm-level selection with the predominance of
379 Mendel's first law seems difficult, but there are several reasons to believe they are com-
380 patible: 1) We find evidence for only tens to hundreds of selective sweeps across deep
381 timescales and across thousands of GIMs, suggesting that they are relatively rare; 2) Se-
382 lective sweeps happen quickly on an evolutionary timescale, erasing standing variation
383 and making transmission ratio distortion a rare phenomenon at any one time; 3) Because
384 most GIMs lead to only modest allelic differences (2-4 fold), sperm with these differences
385 may be functionally equivalent or will lead to modest transmission ratio distortion, as is

386 observed for example in mouse Yq deletions or Slx knockdowns (Conway et al. 1994; Coc-
387 quet et al. 2012), which is challenging to quantify in most mammals; and 4) Avoidance of
388 evolutionary conflict by evolving sperm-specific expression removes cases of balancing se-
389 lection, which might have resulted in observable transmission ratio distortion on standing
390 variation.

391 While genoinformative expression is widespread at the RNA level, we do not have
392 direct evidence for how common it is at the protein level. One reason to believe there
393 are substantially fewer protein-level GIMs than RNA-level GIMs is that proteins can
394 be shared across cytoplasmic bridges. This is consistent with the fact that GIMs that
395 are preferentially translated late in spermiogenesis, when there is little to no time to
396 be shared across cytoplasmic bridges, are enriched in evidence for selection or avoidance
397 of evolutionary conflict. Even extremely late-expressed GIMs may not always lead to
398 functional differences in sperm, because epididymal exosomes deliver proteins from diploid
399 cells to sperm after they cease transcription and translation, potentially masking allelic
400 differences in mature sperm. Another mechanism for masking the functional consequences
401 of GIMs may be the abundant post-translational regulation of mature sperm, for example
402 during capacitation, which might create larger cell-to-cell variation among sperm than
403 GIMs.

404 Nevertheless, the ability of GIMs to lead to sperm-level natural selection may have
405 profound evolutionary consequences. We have shown strong enrichments of GIMs for
406 testis-specific expression, testis-specific paralogs, and testis-specific isoforms or exon us-
407 age. There are two forces that could give rise to these results: first, evolutionary con-
408 flict arises repeatedly in GIMs, which provides an evolutionary advantage to evolve dis-
409 tinct sperm-level function; second, that evolutionary conflict provides pressure to decrease
410 genoinformativity (i.e. increase sharing across cytoplasmic bridges), so that the remaining
411 GIMs preferentially have more sperm-specific expression. It is impossible to distinguish
412 between these models with the data here, but it is likely that both forces contribute.

413 More comprehensive catalogs of GIMs across species may be necessary to identify which
414 is predominant.

415 This provokes a profound question: why, from an evolutionary perspective, do GIMs
416 exist at all? For sex chromosome genes, such as *Akap4* (an X-chromosome gene required
417 for sperm motility), it is clear that some degree of sharing is required for sperm function
418 and specific mechanisms have evolved to facilitate sharing (Morales et al. 2002). However,
419 it is not clear that genes need to be shared equally or that absolute functional equivalence
420 is achieved; in some cases, a 2-fold or 4-fold difference in allelic expression may not have
421 strong enough functional effects to exert evolutionary pressure to fully share transcripts.
422 Also, sex chromosome genes are a special case that are hemizygous in males, so there may
423 be even less pressure to share equally for autosomes. For genes or isoforms that are sperm
424 specific, there could in fact be a benefit to sperm-level selection: an intensification of
425 both purifying and positive selection by adding a selective layer on top of organism-level
426 selection. In these cases, there would be no evolutionary conflict between the two selective
427 layers, so some GIMs could become "selfish elements" whose interests are aligned with
428 the organism: improving sperm function, which in turn increases the number of offspring.

429 The testis-expressed genome has long been a puzzling outlier, including by far the most
430 tissue-specific gene expression, the most tissue-specific paralogs, and the most rapidly
431 evolving genes. The widespread presence of GIMs raises the possibility that sperm-level
432 selection and resulting evolutionary conflicts are common enough to provide an elegant
433 explanation for these phenomena. If functional and molecular heterogeneity of sperm
434 can be understood in enough detail, it is even possible that it could be exploited to
435 isolate and eliminate sperm carrying severe Mendelian disease genes, reducing the risk
436 of disease transmission across generations, as has been previously suggested (e.g. Butler
437 et al. 2007). Given the rarity of GIM-related selective sweeps, it may be technically
438 challenging to identify and leverage this expanded source of sperm heterogeneity. However,
439 the surprisingly widespread existence of GIMs raises the possibility that a wide variety of

440 severe diseases could be prevented by means of sperm selection.

441 **6 Methods**

442 **6.1 Spermatid isolation and cell sorting**

443 Testes were reduced to a single-cell suspension (breaking apart the intracellular bridges
444 between germ cells in the process), using the two-step digestion protocol of Gaysinskaya
445 et al. 2014. Digestions were performed in 6 ml for mouse, with one whole testis as starting
446 material (tunica albuginea removed); and in 30 ml for non-human primate, with 600mg
447 of diced testis tissue as starting material. First, to disperse the seminiferous tubules,
448 testis tissue was incubated in digestion solution 1: Hanks' Balanced Salt solution (HBSS,
449 Sigma Aldrich), 1 mg/ml collagenase Type I (Worthington Biochemical), and 6 U/ml
450 DNase I (Sigma Aldrich). Incubation was at 37°C for 10 min with horizontal agitation.
451 Tubules were then allowed to settle and the supernatant (containing somatic cells) was
452 discarded. Digestion solution 2 was then added to reduce the tubules to a single-cell
453 suspension: HBSS, 1 mg/ml collagenase, 6 U/ml DNase, and 0.05% trypsin (Gibco, 2.5%
454 stock solution). Incubation was for 25 min at 37°C with horizontal agitation; tubules were
455 pipetted every 5 minutes, and an additional 0.025% trypsin was added halfway through
456 the incubation. Successful digestion was confirmed by examining the cell suspension under
457 a light microscope. Digestion was quenched with Newborn Calf Serum (Gibco).

458 After digestion, the single-cell suspension was filtered through a 100 µm cell strainer
459 and centrifuged for 10 minutes at 500g. The supernatant was discarded, and the cell pellet
460 was gently resuspended at $1-2 \times 10^6$ cells/ml in PBS + 5 mg/ml BSA. Hoechst 33342 was
461 added at 10 µg/ml and cells were incubated for 30 minutes at 37°C. Propidium iodide
462 (PI) was added at 1 µg/ml during the last 5 minutes of incubation. Samples were filtered
463 through a 40 µm mesh immediately before sorting.

464 Single live spermatids were then sorted into 96-well plates as described below, using a

465 BD FACS Aria, a Beckman Coulter MoFlo Astrios, or a SONY Synergy SY3200 instru-
466 ment. Our gating strategy was as follows: Selected for 1n cells (spermatids and sperm)
467 based on Hoechst 33342 fluorescence intensity (with 355 nm excitation and a 448/59 nm
468 bandpass emission filter) (Gaysinskaya et al. 2014); Selected for PI-negative cells to get a
469 live population (PI was measured with 561 nm excitation and a 614/20 nm bandpass emis-
470 sion filter)); Enriched for round spermatids by selecting cells with high forward scatter
471 (Bastos et al. 2005)

472 **6.2 Cynomolgus primates**

473 Adult male cynomolgus monkeys (*Macaca fascicularis*) were used for the non-human pri-
474 mate studies conducted at the University of Kentucky. Monkeys were singly housed in
475 climate-controlled conditions with 12-hour light/dark cycles. Monkeys were provided wa-
476 ter ad libitum and fed Teklad Global 20% Protein Primate Diet. Spermatid isolation and
477 sorting was performed at the University of Kentucky with two male monkeys. Monkeys
478 were euthanized, testes were promptly removed and placed in Hanks' Balanced Salt Solu-
479 tion (HBSS) on ice, prior to proceeding to tissue digestion and subsequent preparation of
480 a single cell suspension for cell sorting. All animal care, procedures, and experiments were
481 based on approved institutional protocols from the University of Kentucky Institutional
482 Animal Care and Use Committee IACUC (protocol #2015-2294).

483 **6.3 Single-cell RNA sequencing**

484 Single cells meant for RNA processing were sorted into 96-well full-skirted Eppendorf
485 plates that were pre-chilled at 4°C and were prefilled with 10µL of lysis buffer consisting
486 of TCL buffer (Qiagen) supplemented with 1% beta-mercaptoethanol. Sorted plates with
487 single-cell lysates were subsequently sealed, vortexed, spun down at 300g at 4°C for 1
488 minute, immediately placed on dry ice to flash-freeze the lysates, and then moved to -80°C
489 for storage. The Smart-Seq2 protocol was performed on single sorted cells as previously

490 described (1-3), with some modifications described below.

491 **Reverse transcription** Single-cells lysates were thawed on ice for 2 minutes, then
492 centrifuged at 3,000rpm at 4°C for 1 minute. 20µL of Agencourt RNAClean XP SPRI
493 beads (Beckman-Coulter) was added to lysates, mixed slowly, to not introduce bub-
494 bles and subsequently incubated at room temperature for 10 minutes. The 96-well
495 plate was then placed onto a magnet (DynaMag-96 Side Skirted Magnet, Life Tech-
496 nologies) for 5 minutes while covered. The supernatant was removed, and the SPRI
497 beads were washed three times with 100µL of freshly prepared 80% ethanol, careful to
498 avoid loss of beads during the washes. Upon completely removing ethanol after the last
499 wash, SPRI beads were left to dry at room temperature for up to 10 minutes. Beads
500 were resuspended in using 4µL of the following Elution Mix: 0.1µL 10µM RT primer
501 (5'AAGCAGTGGTATCAACGCAGAGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTVN-
502 3', IDT), 1µL 10 mM dNTP (Life Technologies), 0.1µL Recombinant RNase-Inhibitor (40
503 U/µL, Clontech), and 2.8µL nuclease-free water. The plates were sealed and then spun
504 down briefly, 5 seconds max to get up to 150rpm. The samples were denatured at 72°C
505 for 3 minutes and placed immediately on ice afterwards. 7µL of the Reverse Transcription
506 Mix was subsequently added in every well, consisting of: 2µL 5x RT buffer (Thermo Fisher
507 Scientific), 2µL 5 M Betaine (Sigma-Aldrich), 0.09µL 1M MgCl₂ (Sigma-Aldrich), 0.1µL
508 100µM TSO (5'- AAGCAGTGGTATCAACGCAGAGTACATrGrG+G-3', Exiqon), 0.25
509 µL Recombinant RNase-Inhibitor (40 U/µL, Clontech), 0.1µL Maxima H Minus Reverse
510 Transcriptase (200U/µL, Thermo Fisher Scientific), and 2.46µL nuclease-free water. Ev-
511 ery well was mixed with the resuspended beads. Reverse transcription was carried out by
512 incubating the plate at 50°C for 90 minutes, followed by heat inactivation at 85°C for 5
513 minutes.

514 **PCR amplification and cDNA purification** 14µL of PCR Mix was added in each
515 well: 0.05µL 100µM PCR primer (5'- AAGCAGTGGTATCAACGCAGAGT-3', IDT),

516 12.5 μ L 2x KAPA HiFi HotStart ReadyMix (KAPA Biosystems), 1.45 μ L nuclease-free
517 water. The reaction was carried out with an initial incubation at 98°C for 3 minutes,
518 followed by 22 cycles at (98°C for 15 seconds, 67°C for 20 seconds, and 72°C for 6 minutes)
519 and a final extension at 72°C for 5 minutes. PCR products were purified by mixing them
520 with 20 μ L (0.8X) of Agencourt AMPureXP SPRI beads (Beckman-Coulter), followed by
521 a 5 minutes incubation period at room temperature. The plate was then placed onto a
522 magnet for 6 minutes prior to removing the supernatant. SPRI beads were washed twice
523 with 100 μ L of freshly prepared 70% ethanol, carefully to avoid loss of beads during the
524 washes. Upon removing all residual ethanol traces, SPRI beads were left to dry at room
525 temperature for up to 10 minutes. The beads were then resuspended in 20 μ L of TE buffer
526 (Teknova) and incubated at room temperature for 5 minutes. The plate was placed on
527 the magnet for 5 minutes prior to transferring the supernatant containing the amplified
528 cDNA to a new 96-well plate. This cDNA SPRI clean-up procedure was repeated a second
529 time to remove all residual primer dimers and resuspended in a final volume of 15 μ L of
530 TE buffer. The concentration of amplified cDNA was measured using the Qubit dsDNA
531 High Sensitivity Assay Kit (Life 7 Technologies/Thermo Fisher Scientific). The cDNA
532 size distribution of few selected wells was assessed on a High-Sensitivity Bioanalyzer Chip
533 (Agilent). Expected single cell cDNA quantification was around 0.5-2 ng/ μ L with size
534 distribution sharply peaking around 2kb.

535 **Library preparation** Library preparation was carried out using the Nextera XT DNA
536 Sample Kit (Illumina) with indexing adapters that allow 96 single cell libraries to be
537 simultaneously sequenced. For each library, the amplified cDNA was normalized to a
538 0.12-0.20ng/ μ L concentration range. The tagmentation reaction consisted of mixing 1.25
539 μ L of normalized cDNA with 2.5 μ L of Tagmentation DNA Buffer and 1.25 μ L of Amplicon
540 Tagment enzyme Mix. The 5 μ L reaction was mixed well, spun at 3,000 rpm for 3 minutes,
541 incubated at 55°C for 10 minutes and then immediately placed on ice upon completing this
542 incubation step. The reaction was quenched with 1.25 μ L of Neutralize Tagment Buffer

543 and incubated at room temperature for 10 minutes. The libraries were amplified by adding
544 3.75 μ L of Nextera PCR Master Mix, 2.5 μ L of mixed indices (Nextera XT Index Kit). The
545 PCR was carried out at an initial incubation at 72°C for 3 minutes, 95°C for 30 seconds,
546 followed by 12 cycles of (95°C for 10 seconds, 55°C for 30 seconds, 72°C for 1 minute),
547 and a final extension at 72°C for 5 minutes. Following PCR amplification, 2.5 μ L of each
548 library were pooled together in a 2.0 mL Eppendorf tube. The pool was mixed with 216
549 μ L (0.9X ratio for 2.5 μ L of 96 cells pooled together) of Agencourt AMPureXP SPRI beads
550 (Beckman-Coulter) and incubated at room temperature for 5 minutes. The pool was then
551 placed on a magnet (DynaMag-2, Life Technologies) and incubated for 5 minutes. The
552 supernatant was removed and the SPRI beads were washed twice with 1 mL of freshly
553 prepared 70% ethanol. Upon removing all residual ethanol traces, the SPRI beads were
554 left to dry at room temperature for 10 minutes. The beads were resuspended in 100 μ L of
555 TE buffer and incubated at room temperature for 5 minutes. The tube was then placed
556 back on the magnet for 3 minutes prior to transferring the supernatant to a new 1.5 mL
557 Eppendorf tube. This SPRI clean-up procedure of the library was repeated a second time
558 to remove all residual primer dimers, using the same approach and the final resuspension
559 was done in 30 μ L of TE buffer. The concentration of the pooled libraries was measured
560 using the Qubit dsDNA High Sensitivity Assay Kit (Life Technologies/Thermo Fisher
561 Scientific), and the library size distribution measured on a High-Sensitivity Bioanalyzer
562 Chip (Agilent). Expected concentration of the pooled libraries was 30-50 ng/ μ L with size
563 distribution of 300-700 bp.

564 **6.4 Single-cell DNA sequencing**

565 Single haploid cells meant for DNA processing were sorted into 96-well full-skirted Ep-
566 pendorf plates either in (1) 5 μ L of Cell Extraction Buffer (4.8 μ L of Extraction Enzyme
567 Dilution Buffer, 0.2 μ L Cell Extraction enzyme, New England BioLabs) and processed
568 using the PicoPlex kit, or (2) 5 μ L of Cell Lysis Reaction Mix (4.9 μ L of Cell Lysis Buffer,

569 0.1 μ L Cell Lysis enzyme, Yikon Genomics). Sorted plates with single-cell lysates were
570 subsequently sealed, vortexed, spun down at 300g at 4°C for 1 minute, immediately placed
571 on dry ice to flash-freeze the lysates, and then moved to -80°C for storage. All single-cells
572 plates were thawed on ice for 2 minutes, then centrifuged at 3,000 rpm at 4°C for 1 minute
573 prior to processing.

574 **PicoPlex Amplification** For PicoPlex amplification, the plates were incubated at 75°C
575 for 10 minutes, followed by 95°C for 4 minutes and held at 22°C until ready for the next
576 steps. The pre-amplification Master Mix, consisting of 4.8 μ L of Pre-Amp Buffer and 0.2
577 μ L of Pre-Amp Enzyme Mix was added to each cell, the reaction was mixed well, spun
578 at 3,000 rpm for 1 minute. The PCR was carried out at an initial incubation at 95°C
579 for 2 minutes, followed by 12 cycles of (95°C for 15 seconds, 15°C for 50 seconds, 25°C
580 for 40 seconds, 35°C for 30 seconds, 65°C for 40 seconds, 75°C for 40 seconds), and a
581 hold at 4°C. Following the pre-amplification reaction, each well was mixed well with the
582 Amplification Master mix, consisting of 25 μ L of Amplification Buffer, 34.2 μ L of nuclease-
583 free water and 0.8 μ L of Amplification Enzyme Mix. The reactions were mixed well, spun
584 at 3,000 rpm for 1 minute and incubated at 95°C for 2 minutes, followed by 16 cycles
585 of (95°C for 15 seconds, 65°C for 1 minute, 75°C for 1 minute) and a hold at 4°C. The
586 concentration of each cell was measured using the Qubit dsDNA High Sensitivity Assay
587 Kit (Life Technologies/Thermo Fisher Scientific). Expected concentration of the single
588 cell lysates was 20-50 ng/ μ L with size distribution of 300-1000 bp.

589 **MALBAC Amplification** For MALBAC amplification, the plates were incubated at
590 50°C for 50 minutes, followed by 80°C for 10 minutes and held at 4°C until ready for the
591 next steps. The pre-amplification Reaction Mix, consisting of 29 μ L of Pre-Amp Buffer
592 and 1 μ L of Pre-Amp Enzyme Mix was added to each cell, the reaction was mixed well,
593 spun at 3,000 rpm for 1 minute. The PCR was carried out at an initial incubation at
594 94°C for 3 minutes, followed by 8 cycles of (20°C for 40 seconds, 30°C for 40 seconds,

595 40°C for 30 seconds, 50°C for 30 seconds, 60°C for 30 seconds, 70°C for 4 minutes, 95°C
596 for 20 seconds, 58°C for 10 seconds), and a hold at 4°C. Following the pre-amplification
597 reaction, each well was mixed well with the Amplification Reaction mix, consisting of 29.2
598 µL of Amp Buffer and 0.8 µLamp Enzyme Mix. The reactions were mixed well, spun at
599 3,000 rpm for 1 minute and incubated at 94°C for 30 seconds, followed by 21 cycles of
600 (94°C for 20 seconds, 58°C for 30 seconds, 72°C for 3 minutes) and a hold at 4°C. The
601 concentration of each cell was measured using the Qubit dsDNA High Sensitivity Assay
602 Kit (Life Technologies/Thermo Fisher Scientific). Expected concentration of the single
603 cell lysates was 20-60 ng/µL with size distribution of 300-2000 bp.

604 **6.5 Haplotype Phasing**

605 **6.5.1 Mouse**

606 We downloaded the combined VCF of laboratory mouse strains from The Mouse Genome
607 project (Keane et al. 2011) and defined maternal and paternal haplotypes utilizing SNPs
608 unique to either C57BL/6J or PWK/PHJ, respectively. For all analyses, we disregarded
609 indels and only considered SNPs. This resulted in a total of 20,986,995 heterozygous
610 SNPs, which overlapped 28,497 expressed genes in mouse round spermatids.

611 **10X Chromium Alignment and Haplotype Calling** We created maternal and pa-
612 ternal haplotypes of the two non-human primates using a combination of 10X Chromium
613 linked read sequencing on diploid cells and sparse single cell DNA sequencing on haploid
614 spermatid cells. We aligned the 10X Chromium reads to the *Macaca fascicularis* genome
615 *Macaca_fascicularis_MacFac_5.0* from Ensembl, herein referred to as Ensembl-MF5-G,
616 by first creating a custom reference using the longranger mkref command , and then run-
617 ning longranger using this reference and default parameters to generate 10X Chromium
618 alignment data.

619 For *Cynomolgus* 1, the instrument generated 1,850,208 Gel Beads in Emulsion (GEMs)

620 and the software mapped 819,440,960 reads for 37.2x average coverage across the genome.
621 This resulted in 361,465 haplotype blocks with N50 length of 1.6 MB. Each block contained
622 an average of 36 SNPs for a total of 12,758,999 heterozygous SNPs. Cynomolgus 2 10x
623 Chromium data featured 1,889,596 GEMs that led to 812,899,614 reads mapping at an
624 average coverage of 37.4X. It had 318,516 blocks with N50 length of 1.8 MB and an
625 average of 40 heterozygous SNPs per block, for a total of 12,744,826 heterozygous SNPs.

626 **Mature Sperm scWGS Alignment and Processing** We genotyped the haploid
627 sperm scWGS samples using a custom pipeline. First, the paired-end reads were aligned
628 to Ensembl-MF5 using BWA v0.7.5 (Li and Durbin 2009) using the mem option with
629 default parameters. The resulting bam files were sorted using samtools v1.4.1 sort (Li,
630 Handsaker, et al. 2009) and duplicates were removed using sambamba v0.6.6 (Tarasov
631 et al. 2015). `samtools mpileup` with a bed file of the 10X Chromium identified variant
632 positions calculated the allelic depths per heterozygous site. We then filtered the file to
633 only include allelic depths of variant alleles. For Cynomolgus 1, this resulted in an average
634 of 1.2M heterozygous sites per spermatid sample, for a total overlap of 3.3M sites across
635 the 17 spermatid samples. With the 8 spermatid samples for Cynomolgus 2, we covered
636 3.5M total sites with an average 1.2M sites per sample at roughly 1X coverage.

637 **Creating Chromosome-Length Haplotype Blocks** The final step involved stitching
638 the the haplotype blocks generated by 10X Chromium sequencing into chromosome-length
639 haplotypes using the haploid cell haplotypes as a guide. In the case of no recombination,
640 the stitching is trivial and requires only a single sperm sample. Due to recombination, we
641 used multiple sperm single cell WGS samples, and utilized a dynamic programming frame-
642 work tuned to minimizing the number of recombination events to assign the chromium
643 blocks to maternal and paternal haplotypes.

644 **6.6 Allele-specific Expression Quantification**

645 **StringTie Transcriptome Assembly** Due to unavailability of a publicly available
646 testes transcriptome of *Cynomolgus*, we created a custom cynomolgus transcriptome us-
647 ing single cell RNA-seq samples of round spermatid and elongating spermatid cells from
648 the two individuals. First, we aligned the samples to Ensembl-MF5-G using STAR v2.5.3
649 with default parameters, and merged and sorted the bams using samtools. This resulted
650 in 3.7 billion total reads aligned across the corpus of 480 samples. We fed the merged
651 bam into StringTie v1.3.3 with default options except for `-p 39` to indicate a large number
652 of available threads. We compared the StringTie generated transcriptome to MacFas_5.0,
653 a Ensembl-generated transcriptome of *Macaca fascicularis* using Cufflinks v.2.2.1 `gffcom-`
654 `pare`, and created a dictionary to map the StringTie annotation ids back to known gene
655 symbols.

656 **RNA-Seq Processing and Alignment** To reduce allelic bias in read mapping, we
657 used bcftools consensus to generate masked genomes, in which all bases in heterozygous
658 positions were modified to the IUPAC character N in the reference genomes. We used
659 STAR v2.5.3 to align the round and elongating spermatid single cell RNAseq reads, but
660 created custom STAR genomes with either Ensembl GRCm38 or the previously described
661 StringTie-generated transcriptomes. We utilized STAR options `-outFilterMultimapNmax`
662 `1` to eliminate multi-mapping reads, `-alignSJOverhangMin 4` to force large overlap be-
663 tween RNA-seq reads and the genome, and `-outSAMattributes NH HI NM nM MD XS`
664 attributes, and removed duplicated reads using sambamba. `featureCounts` (Liao, Smyth,
665 and Shi 2014) was used to generate gene transcripts per million (TPM) values with op-
666 tions `-s 0` for unstranded reads, `-p` for paired end reads, and `-B` to require both ends of
667 the read to be mapped.

668 **Generating Allele-specific Counts** To quantify allele specific expression of genes,
669 we first assigned each heterozygous SNP to a gene using the `snpEff` Cingolani et al. 2012

670 annotate tool using custom snpEff databases. Then, after splitting the aligned RNA-
671 seq bams into chromosome-specific BAMs, we generated the allele counts for each gene
672 and spermatid sample combination. To avoid double-counting of reads that overlapped
673 multiple sites, each read was only counted once in favor of either allele, and if a read
674 matched variants on both alleles, we tagged it as a discordant read and did not utilize it
675 for further analysis. For mouse, we average 145 allele-specific reads per gene per sample
676 across 11,542 phaseable genes in 95 spermatid samples.

677 We performed an additional step to quantify allele specific counts in the non-human
678 primate samples. Due to limited coverage across the length of an entire gene, StringTie
679 often splits a single Ensembl gene annotation into multiple gene annotations. As such,
680 we summed reads from separate StringTie genes overlapping known annotations. The
681 resulting allele counts files for the monkeys are a combination of known genes annotated
682 by Ensembl and novel genes identified by StringTie only. For *Cynomolgus* 1, in 187
683 spermatid samples, we average 122 reads per gene per sample for 8956 phased genes. For
684 *Cynomolgus* 2, in 185 spermatid samples, we average 131 reads per gene per sample in
685 8216 genes.

686 **6.7 Haplotype and Genoinformativity Inference**

687 To study haploid-biased gene expression, we require knowledge of the underlying haplo-
688 type. We reasoned that if there was true haploid biased expression, it would be possible
689 to infer the haplotype from the allele specific expression data. As such, we derived a
690 model to perform both haplotype and genoinformativity inference simulataneously. Here,
691 we first describe a model for transcript sharing across a syncytium and then extend it to
692 a probability model for observing allele specific reads from round spermatids in a single
693 cell RNA-seq assay.

694 **Model of Genoinformative Transcripts** We begin by describing a simple model
695 for the number of transcripts of a single gene g in a single cell c . The total number of
696 transcripts T in the haploid cell is the combination of external transcripts E and retained
697 transcripts R .

$$T = E + R \quad (1)$$

698 Here, external transcripts indicates transcripts that were not transcribed by the hap-
699 loid cell, but rather were transported into the cell through the cytoplasmic bridge. Re-
700 tained transcripts are the transcripts that were transcribed by the cell and not shared
701 through the cytoplasmic bridge.

702 We can also write down the total transcripts T as the combination of transcripts from
703 the maternal allele of the gene M or the paternal allele of the gene P .

$$T = M + P \quad (2)$$

704 Note that we can marginalize the maternal and paternal transcripts in terms of external
705 and retained transcripts.

$$M = E_M + R_M$$

$$P = E_P + R_P$$

$$T = E_M + E_P + R_M + R_P$$

706 Before deriving a model for genoinformativity, we introduce two last definitions in the
707 form of ratios. The ratio of E_M to E , or the skew of transcripts towards the maternal
708 allele S , and the ratio of R to T , or the genoinformativity of the transcript.

$$S = \frac{E_M}{E} \quad (3)$$

$$G = \frac{R}{T} \quad (4)$$

709 Assuming no eQTL effects, genome imprinting, technical bias, or other mechanisms
710 for differential allelic expression, this allelic skew S is 0.5, i.e. the number of haploid cells
711 that contain maternal and paternal genotype are equal and the number of transcripts
712 transferring into the cell is equal from either allele.

713 **Haploid Cell with Maternal Allele** Given the previous system and definitions, we
714 now derive the empirical genoinformativity for a single haploid cell. Consider the case
715 of a cell c having the maternal allele for the gene or haplotype H_M . Then, we further
716 deconvolve the total transcripts by the transcripts from the maternal allele M and the
717 transcripts from the paternal allele P . Note that this classification is only relevant for
718 autosomes, where it is possible to have transcripts from either chromosome in the haploid
719 cell.

$$\begin{aligned} M|H_M &= E_M|H_M + R_M|H_M \\ &= E_M + R_M|H_M \end{aligned} \quad (5)$$

$$\begin{aligned} P|H_M &= E_P|H_M + R_P|H_M \\ &= E_P \end{aligned} \quad (6)$$

720 Since the cell has a maternal allele only for the gene of interest, there are no retained
721 reads from the paternal allele. Finally, let's express the total transcripts T in terms of
722 the maternal and paternal transcripts.

$$\begin{aligned} T|H_M &= E_M + E_P + R_M \\ &= E + R_M \end{aligned} \tag{7}$$

723 Given equation 3, 4, and 7, we can restate equation 5 as

$$\begin{aligned} M|H_M &= ES + R_M = (T - R_M)S + R_M \\ &= TS + (1 - S)R_M \\ &= TS + (1 - S)TG \\ &= (S + (1 - S)G)T \end{aligned} \tag{8}$$

724 We can derive similar equations for $P|H_M$, $M|H_P$ and $P|H_P$.

$$\begin{aligned} P|H_M &= ((1 - S) - (1 - S)G)T \\ M|H_P &= (S - SG)T \\ P|H_P &= ((1 - S) + SG)T \end{aligned} \tag{9}$$

725 **Probability Model for Allele-Specific Reads** We now focus our attention on devel-
 726 oping a model for observing allele-specific reads using single cell RNA-Seq from haploid
 727 round spermatids. We derive a probability model for observing counts of alleles from the
 728 maternal allele C^M and paternal allele C^P for N individuals and G genes. Given param-
 729 eters θ , each cell i and gene j is independent of each other and the collective probability
 730 can be written as:

$$P(D|\theta) = \prod_i^N \prod_j^G P(C_{ij}^M, C_{ij}^P|\theta) \tag{10}$$

731 For simplicity, we will write the set of counts C_{ij}^M and C_{ij}^P as D_{ij} where applicable.
 732 The main reason we are able to treat each set of counts independently is because we
 733 marginalize the probability over the haplotype H_{ij} of cell i at gene j .

$$\begin{aligned}
 P(D_{ij}|H_{ij}, \theta) &= P(D_{ij}|H_{ij} = H_M, \theta)P(H_{ij} = H_M|\theta) \\
 &+ P(D_{ij}|H_{ij} = H_P, \theta)P(H_{ij} = H_P|\theta)
 \end{aligned}
 \tag{11}$$

734 Using the above formulation, it is possible to split the inference goal into two separate
 735 sub-tasks: haplotype inference and genoinformativity inference.

736 **Haplotype Inference** We use a Markov chain across a single chromosome to estimate
 737 the haplotype given a recombination rate r .

$$P(H_{ij}|\theta) = P(H_{ij}|H_{ij-1})P(H_{ij-1}|\theta)
 \tag{12}$$

738 where

$$P(H_{ij}|H_{ij-1}) = \begin{cases} (1-r) & H_{ij} = H_{ij-1} \\ r & H_{ij} \neq H_{ij-1} \end{cases}
 \tag{13}$$

739 We set the initial probability of each cell's haplotype to be equal at 0.5.

740 **Genoinformativity Inference** Given the haplotype H_{ij} of cell i at gene j , the counts
 741 of the maternal and paternal allele follows from the generative model described above.
 742 Due to overdispersion in RNA-seq data, we model the counts using a beta-binomial dis-
 743 tribution, which is specified by shape parameters α and β . In fitting the model, we only
 744 fit the shape parameter β and reparameterize α in terms of skew S and genoinformativity
 745 G . More explicitly, we can model the system as

$$\frac{\alpha_j}{\alpha_j + \beta_j} = \begin{cases} S_j + G_j(1 - S_j) & H_{ij} = H_M \\ S_j - G_j(S_j) & H_{ij} = H_P \end{cases}
 \tag{14}$$

$$P(D_{ij}|H_{ij}, \theta)P(\theta) = P(D_{ij}|H_{ij}, S_j, G_j, \beta_j)P(\beta_j)P(G_j)P(S_j) \quad (15)$$

$$P(D_{ij}|H_{ij}, S_j, G_j, \beta_j) = \text{Beta-Binomial}(\alpha_j, \beta_j, D_{ij}) \quad (16)$$

746 In addition to overdispersion, single cell RNA-seq data also contains high amount of
 747 allelic dropout and amplification of a single molecule. To alleviate the impact of allelic
 748 dropout on estimates of genoinformativity, we introduce a Zero-and-N-inflated Beta Bino-
 749 mial distribution parameterized by an additional variable ζ_j which defines the probability
 750 of allelic dropout for the gene.

$$P(C_{ij}^M, C_{ij}^P|H_{ij}, S_j, G_j, \beta_j, \zeta_j) = \begin{cases} \frac{\zeta_j \alpha_j}{\alpha_j + \beta_j} + (1 - \zeta_j) \text{Beta-Binomial}(\alpha_j, \beta_j, D_{ij}) & C_{ij}^M = 0 \\ \frac{\zeta_j \beta_j}{\alpha_j + \beta_j} + (1 - \zeta_j) \text{Beta-Binomial}(\alpha_j, \beta_j, D_{ij}) & C_{ij}^P = 0 \\ \text{Beta-Binomial}(\alpha_j, \beta_j, D_{ij}) & \text{otherwise} \end{cases} \quad (17)$$

751 6.7.1 Implementation

752 **Haplotype Inference** Unfortunately due to inherent noise in the system and the cost
 753 of sampling the aforementioned Markov chain, we do not compute the Markov chain for
 754 each gene independently. Instead, we bin the genes into buckets B and perform a similar
 755 inference task with each bucket k . Each bucket on average contained 10 genes in our fits.

$$P(H_{ik}|\theta) = P(H_{ik}|H_{ik-1})P(H_{ik-1}|\theta) \quad (18)$$

756 We also used a fixed recombination rate r for each cell and each chromosome with the
 757 assumption that a cell would have on average 0.5 recombination events per chromosome.

758 **Genoinformativity Inference** Instead of learning the parameter S_j for each gene, we
759 use an empirical estimate of S_j derived from dividing the number of H_M reads for a gene j
760 by the total number of reads for that gene across all cells. We also tested using the mean
761 of the empirical S_{ij} derived from each cell separately, and did not notice large differences
762 in the model fits.

Priors

$$P(H_{ij}) = \begin{cases} 0.5 & H_{ij} = H_M \\ 0.5 & H_{ij} = H_P \end{cases} \quad (19)$$

$$P(G_j) = \text{Uniform}(0, 1) \quad (20)$$

$$P(\beta_j) = \text{Uniform}(3, 30) \quad (21)$$

$$P(\zeta_j) = \text{Uniform}(0.005, 1) \quad (22)$$

763 **Two-stage Fitting** For computational efficiency, we split the inference task into two
764 stages. In the first stage, we fit both the haplotype and genoinformativity inference steps
765 for highly expressed genes ($\text{TPM} > 20$). Then, in the second stage, we only performed
766 genoinformativity inference using fixed haplotype probabilities. We used the mean pos-
767 terior of the haplotypes from the first stage, and interpolated the probability for genes
768 that were unique to the second stage. There was 99% correlation between the posterior
769 mean of the genoinformativity values, indicating low variance in the posterior haplotype
770 distributions and high confidence in haplotype inference.

771 **Samplers** We used PyMC3 (Salvatier, Wiecki, and Fonnesbeck 2016) as the frame-
772 work for sampling the model. For the haplotype sampling, we used a Categorical Gibbs

773 Metropolis sampler. All the other parameters were sampled using the No U-turn Sampler
774 (NUTS) with a target accept probability of 0.8. We sampled the model for 5000 steps with
775 two separate chains and used the last 500 steps for estimating the posterior distribution
776 across the 2 chains.

777 **6.7.2 Sex chromosome GIMs**

778 Mouse gene-level transcripts per million (TPM) values were collected for all genes in all
779 spermatids using all RNAseq reads, not only allele-informative reads. For each gene, a
780 loess regression was used to fit its log₂ expression across the diffusion pseudotime with a
781 pseudocount of 1 TPM, using the R loess function with a gaussian function family and
782 0.75 span. The residuals from this fit were then used to calculate pairwise Spearman
783 correlations between all sex chromosome genes. Pairwise correlations were hierarchically
784 clustered using the complete linkage method, with the results visualized in heatmaps. A
785 cutoff height of 6 was empirically found to split the data into three clusters: a distinct X
786 cluster, an anti-correlated distinct Y cluster, and a mixed X and Y cluster with no strong
787 correlation patterns. Genes in the first two clusters were considered potential GIMs. We
788 calculated the mean pairwise Spearman correlation between pairs of potential GIMs, with
789 the sign reversed for genes in opposite clusters. Genes with a mean pairwise correlation
790 of greater than 0.05 (roughly the median value over potential GIMs) were selected as
791 putative sex chromosome GIMs.

792 **6.8 GIM classification**

793 To classify each gene as a “Confident GIM”, “Confident Non-GIM”, or “Remaining Gene”,
794 we fit the Bayesian model to shuffled data, and compared the posterior distributions for
795 H_{ik} , G_j , and β_j between real and shuffled data. We utilized two main shuffling methods:
796 complete shuffle and cell-label shuffle for each chromosome independently. The complete
797 shuffle shuffled the allele counts randomly across the population of cells and genes. For the

798 cell-label shuffle, the allele counts were randomized across the cells, but the distribution
799 of counts in a gene remained the same. We trained our Bayesian HMM using the same
800 default parameters and priors as the real data, and compared the model fits. Since β_j can
801 capture both the variance of single cell rna-seq as well as the variance in genoinformativity,
802 we created a new measure γ_j as an alternative measure of genoinformativity that combines
803 both posterior mean estimates of G_j and β_j .

$$\alpha_j = \frac{G_j(\beta_j)}{1 - G_j}$$
$$\gamma_j = \text{CDF of Beta}(\alpha_j, \beta_j) == 0.05$$

804 To reflect the confidence of the haplotype fits H_{ik} across all n samples, we also created
805 an aggregated measure, fraction of poor haplotypes $\zeta_j = \frac{\sum_{i=1}^N \sum_{k=1}^K I[(H_{ik} < 0.95) | H_{ik} > 0.05]}{NK}$,
806 which reflected the proportion of haplotypes that a posterior mean haplotype probability
807 less than 0.95 for either the maternal or paternal haplotypes.

808 We performed a grid search across thresholds for highest posterior density (hpd) eval-
809 uated at 5% and 95% for genoinformativity G_j and γ_j and ζ_j , which controlled the eFDR
810 at 10% for confident and non-confident gims compared to the shuffled control. For a
811 particular gene, the thresholds for a "Confident GIM" are: hpd 5% of genoinformativity
812 > 0.025 , hpd 95% of genoinformativity > 0.2 , $\gamma_j > 0.025$, fraction of poor haplotypes
813 < 0.4 For "Confident Non-GIMs" are restricted to hpd 95% of genoinformativity < 0.2 .
814 Genes that fall outside these bounds were considered "Remaining Genes".

815 6.9 GIM characterization

816 **Expression-matched control selection** The expression trajectory across spermiogen-
817 esis was first tabulated for each gene by cross-referencing the log2 of the TPM expression
818 level (with a pseudocount of 1 and complete dropout considered a zero) against the diffu-

819 sion map pseudotime value for each cell (i.e. the first dimension of the map). To reduce
820 noise at the single cell level, a smoothed loess fit was used as the expression trajectory
821 (fit using default parameters for the R `loess` function).

822 Next, all confident non-GIMs expressed in spermatids were considered as controls for
823 all GIMs. Pools of controls were first reduced for each GIM based on two hard filters:
824 first, all genes were equally distributed into 5 bins based on their dropout rates; second,
825 the slope of a linear fit to the expression trajectory was required to differ by no more than
826 0.2. This helped to control for any confounders resulting in oversampling, as well as large
827 expression changes in a small number of cells, generally in the extreme early or late part
828 of the trajectory.

829 For each GIM, all non-GIMs remaining in its pool were ranked by their mean squared
830 difference in log2 expression level, and the top 20 were selected as mock GIM controls,
831 whose ranks were then scrambled. This resulted in 20 control sets of mock GIMs having
832 similar dropout rates, slope of expression trajectory, and low difference in expression
833 trajectory. For analyses limited to protein coding genes, control selection was performed
834 again with both the GIMs and the control pools limited to protein coding genes.

835 For the cynomolgus samples, the expression trajectories were averaged across the two
836 individuals. Where stringtie genes overlapped with Ensembl annotations, the aggregated
837 expression for the Ensembl annotation was used for both GIMs and controls. A gene
838 was considered a GIM if it was called as a confident GIM in either individual, and was
839 considered as a non-GIM if it was called as a confident non-GIM in either individual. The
840 rare genes having conflicting calls in each individual were excluded from these analyses.
841 Human GIMs and non-GIMs were inferred from homologous cynomolgus annotations with
842 homology defined as having the same Ensembl gene symbol (i.e. standard gene name).

843 For spermatid-expressed non-GIM controls, control sets were selected from all confi-
844 dent non-GIMs randomly, without filtering for dropout bin or expression trajectory fit.

845 **Gene Ontology** Mouse gene ontology annotations were downloaded from Ensembl
846 Biomart with the Ensembl Genes 93 / GRCh38.p6 annotation dataset. The mean and
847 standard deviation of number of GIMs expected with each annotation was calculated
848 based on the 20 control sets. Nominal probabilities were then calculated using the normal
849 distribution, and multiple testing was corrected using the Benjamini-Hochberg method to
850 result in false discovery rates. GO terms were considered significant if they had at least
851 20 GIMs, an $FDR \leq 0.001$ and a moderated \log_2 enrichment (using a pseudocount of 5)
852 of at least 0.5.

853 For COMPARTMENTS comparisons, fewer controls had at least one annotation than
854 GIMs, which could artificially inflate significance for individual categories. Therefore,
855 we performed an additional normalization for the expected number of GIMs with an
856 annotation. The number of controls in a set having a GO annotation was converted
857 to a fraction out of those have any annotation, and then multiplied by the number of
858 GIMs to yield the total number expected with each annotation specifically. Otherwise
859 the enrichment analysis was the same as for the GO analysis above.

860 **3' UTR motifs** Only protein-coding genes were considered. The 3' UTR annotations
861 of GIMs and their controls were taken from the highest expressed Ensembl transcript in
862 spermatids. UTRs annotated as less than 7 nucleotides in length were discarded. All 20
863 sets of control UTRs were combined into a single background set, allowing duplicates.
864 AME, a tool from the MEME suite, was run with default parameters using a motif
865 database comprised of the CISBP-RNA and Ray2013 mouse and human sets provided
866 by MEME. The enrichment search was performed using GIMs as foreground and the
867 combined control set as background, with foreground and background switched for the
868 depletion analysis.

869 For candidate RNA-binding proteins, only those with a maximum TPM of 10 at any
870 point in the loess-smoothed expression trajectory were considered. Enrichments were con-
871 sidered significant at an E-value cutoff of 0.01. Motifs having the same IUPAC consensus

872 were merged into a single result.

873 **Selective sweeps** Candidates for mouse selective sweeps were taken from Staubach et
874 al. 2012. Sweep regions in any population were considered. All candidate genes within
875 600kb of each other were collapsed into a single region. For GIMs or each control set,
876 the number of regions having at least one overlapping gene was counted. For a p-value of
877 this difference, the mean and standard deviation of the control sets was used to generate
878 a one-sample t-test.

879 Candidates for human selective sweeps were taken from Refoyo-Martínez et al. 2019;
880 Schrider and Kern 2016; Ferrer-Admetlla et al. 2014; Cheng, Racimo, and Nielsen 2019;
881 Munch et al. 2016, with selective sweep regions as in each paper. In cases in which
882 the paper predicted selective sweep regions but did not annotate associated genes, all
883 genes overlapping the regions were considered selective sweep candidates. Otherwise this
884 analysis was as in the mouse.

885 Direct testing for human selective sweeps was performed using statistics from Pybus et
886 al. 2014 based on analysis of the 1000 genomes project data. To help control for differences
887 in gene length, the median score overlapping the 3' UTR was used to represent the gene.
888 The “best” score for each gene was taken across each population, where “best” signifies
889 the raw score most in favor of a selective sweep for that score. Selective sweep candidates
890 were defined as any where the score was at least 3 standard deviations beyond the mean
891 in this direction. The number of GIM sweep candidates was compared to background
892 expectation of the mean and standard deviation among the 20 control sets.

893 **Testis-specific paralogs** Paralog and tissue-specificity data were taken from Guschan-
894 ski, Warnefors, and Kaessmann 2017. Testis-specific paralogs were defined as those with
895 a “Tissue specificity” (as defined by the paper) of at least 0.90.

896 **Alternative splicing** Mouse alternative splicing was taken from events in VastDB with
897 quality greater than zero and testis specific was defined as a difference in PSI of at least
898 50 between testis and the median PSI across all other tissues.

899 Human isoform expression was taken from the GTEX consortium (file `GTEX_Analysis_2016-01-15_v7.L`).
900 Since individual isoform estimates can be unstable, we considered subsets of isoforms that
901 are expressed higher in testis. Each transcript was ranked by the difference between testis
902 isoform usage (i.e. ratio of transcript TPM to gene TPM in that tissue) to the median
903 tissue isoform usage across other tissues. The maximum of the cumulative sum of excess
904 isoform usage in testis was counted as the testis specificity (testis isoform usage minus
905 other tissue isoform usage). A cutoff of 0.5 was considered testis-specific (equivalent to
906 50 PSI).

907 **Late translation** Translation data was taken from Iguchi, Tobias, and Hecht 2006
908 (GSE4711 on GEO). Translation efficiencies were calculated as the median across repli-
909 cates of the fold change from polysome to RNP samples. Genes were defined as having
910 specific late translation if they were in the bottom quartile of this score at day 22 (which
911 is depleted for late spermiogenesis), and the top quartile with respect to fold-change in-
912 crease in translation efficiency between day 22 and adult mice. For each of the functional
913 readouts of GIMs (e.g. selective sweeps), we compared the fraction of GIMs in that cat-
914 egory that were specifically late translated to those that were not in that category (i.e.
915 not functional candidates by that measure).

916 References

917 **Alavioon2017** Ghazal Alavioon et al. “Haploid selection within a single ejaculate in-
918 creases offspring fitness.” In: *Proceedings of the National Academy of Sciences of*
919 *the United States of America* 114.30 (July 2017), pp. 8053–8058. ISSN: 1091-6490.
920 DOI: 10.1073/pnas.1705601114. URL: <http://www.ncbi.nlm.nih.gov/>

921 [pubmed/28698378%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5544320](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5544320).

923 **Andreassi2009** Catia Andreassi and Antonella Riccio. “To localize or not to localize:
924 mRNA fate is in 3UTR ends”. In: *Trends in Cell Biology* 19.9 (Sept. 2009), pp. 465–
925 474. ISSN: 09628924. DOI: 10.1016/j.tcb.2009.06.001. URL: [http://www.
926 ncbi.nlm.nih.gov/pubmed/19716303%20https://linkinghub.elsevier.com/
927 retrieve/pii/S096289240900141X](http://www.ncbi.nlm.nih.gov/pubmed/19716303%20https://linkinghub.elsevier.com/retrieve/pii/S096289240900141X).

928 **Angerer2016** Philipp Angerer et al. “destiny : diffusion maps for large-scale single-cell
929 data in R”. In: *Bioinformatics* 32.8 (Apr. 2016), pp. 1241–1243. ISSN: 1367-4803.
930 DOI: 10.1093/bioinformatics/btv715. URL: [http://www.ncbi.nlm.nih.gov/
931 pubmed/26668002%20https://academic.oup.com/bioinformatics/article-
932 lookup/doi/10.1093/bioinformatics/btv715](http://www.ncbi.nlm.nih.gov/pubmed/26668002%20https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btv715).

933 **Bastos2005** Henri Bastos et al. “Flow cytometric characterization of viable meiotic and
934 postmeiotic cells by Hoechst 33342 in mouse spermatogenesis”. In: *Cytometry Part*
935 *A* 65A.1 (May 2005), pp. 40–49. ISSN: 1552-4922. DOI: 10.1002/cyto.a.20129.
936 URL: <http://doi.wiley.com/10.1002/cyto.a.20129>.

937 **Binder2014** J. X. Binder et al. “COMPARTMENTS: unification and visualization of
938 protein subcellular localization evidence”. In: *Database* 2014.0 (Feb. 2014), bau012–
939 bau012. ISSN: 1758-0463. DOI: 10.1093/database/bau012. URL: [https://academic.
940 oup.com/database/article-lookup/doi/10.1093/database/bau012](https://academic.oup.com/database/article-lookup/doi/10.1093/database/bau012).

941 **Braun1989** Robert E. Braun et al. “Genetically haploid spermatids are phenotypically
942 diploid”. In: *Nature* 337.6205 (Jan. 1989), pp. 373–376. ISSN: 00280836. DOI: 10.
943 1038/337373a0. URL: <http://www.nature.com/articles/337373a0>.

944 **Butler2007** Avigdor Butler et al. “Sperm abnormalities in heterozygous acid sphin-
945 gomyelinase knockout mice reveal a novel approach for the prevention of genetic
946 diseases.” In: *The American journal of pathology* 170.6 (June 2007), pp. 2077–88.

947 ISSN: 0002-9440. DOI: 10.2353/ajpath.2007.061002. URL: <http://www.ncbi.nlm.nih.gov/pubmed/17525274>
948 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1899442>.
949

950 **Buxbaum2015** Adina R Buxbaum, Gal Haimovich, and Robert H Singer. “In the right
951 place at the right time: visualizing and understanding mRNA localization.” In:
952 *Nature reviews. Molecular cell biology* 16.2 (Feb. 2015), pp. 95–109. ISSN: 1471-
953 0080. DOI: 10.1038/nrm3918. URL: <http://www.ncbi.nlm.nih.gov/pubmed/25549890>
954 [http://www.pubmedcentral.nih.gov/articlerender.fcgi?](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4484810)
955 [artid=PMC4484810](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4484810).

956 **Cheng2019** Jade Yu Cheng, Fernando Racimo, and Rasmus Nielsen. “Ohana: detecting
957 selection in multiple populations by modelling ancestral admixture components”.
958 In: *bioRxiv* (Feb. 2019), p. 546408. DOI: 10.1101/546408. URL: <https://www.biorxiv.org/content/10.1101/546408v1>.
959

960 **Cingolani2012** Pablo Cingolani et al. “A program for annotating and predicting the
961 effects of single nucleotide polymorphisms, SnpEff”. In: *Fly* 6.2 (Apr. 2012), pp. 80–
962 92. ISSN: 1933-6934. DOI: 10.4161/fly.19695. URL: <http://www.tandfonline.com/doi/abs/10.4161/fly.19695>.
963

964 **Cocquet2012** Julie Cocquet et al. “A Genetic Basis for a Postmeiotic X Versus Y Chro-
965 mosome Intragenomic Conflict in the Mouse”. In: *PLoS Genetics* 8.9 (Sept. 2012).
966 Ed. by Michael W. Nachman, e1002900. ISSN: 1553-7404. DOI: 10.1371/journal.pgen.1002900.
967 <https://dx.plos.org/10.1371/journal.pgen.1002900>.

968 **Conway1994** S. J. Conway et al. “Y353/B: a candidate multiple-copy spermiogene-
969 sis gene on the mouse Y chromosome”. In: *Mammalian Genome* 5.4 (Apr. 1994),
970 pp. 203–210. ISSN: 0938-8990. DOI: 10.1007/BF00360546. URL: <http://link.springer.com/10.1007/BF00360546>.
971

- 972 **Cox2009** Allison Cox et al. “A new standard genetic map for the laboratory mouse”.
973 In: *Genetics* 182.4 (Aug. 2009), pp. 1335–1344. ISSN: 00166731. DOI: 10.1534/
974 genetics.109.105486. URL: [http://www.ncbi.nlm.nih.gov/pubmed/19535546%](http://www.ncbi.nlm.nih.gov/pubmed/19535546%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2728870)
975 [20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2728870](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2728870).
- 976 **Crow1979** J F Crow. “Genes That Violate Mendel Rules”. In: *Scientific American* 240.2
977 (1979), pp. 134–&. DOI: 10.2307/24965130. URL: [https://www.jstor.org/
978 stable/24965130](https://www.jstor.org/stable/24965130).
- 979 **Eep2019** Johnson Eep, Ellis Pji, and Peter Ellis Email. “Differential sperm motility
980 mediates the sex ratio drive shaping mouse sex chromosome evolution”. In: (). DOI:
981 10.1101/649707. URL: <http://dx.doi.org/10.1101/649707>.
- 982 **Ferrer2014** Anna Ferrer-Admetlla et al. “On detecting incomplete soft or hard selective
983 sweeps using haplotype structure.” In: *Molecular biology and evolution* 31.5 (May
984 2014), pp. 1275–91. ISSN: 1537-1719. DOI: 10.1093/molbev/msu077. URL: [http:
985 //www.ncbi.nlm.nih.gov/pubmed/24554778%20http://www.pubmedcentral.
986 nih.gov/articlerender.fcgi?artid=PMC3995338](http://www.ncbi.nlm.nih.gov/pubmed/24554778%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3995338).
- 987 **Gaysinskaya2014** Valeriya Gaysinskaya et al. “Optimized flow cytometry isolation of
988 murine spermatocytes”. In: *Cytometry Part A* 85.6 (June 2014), pp. 556–565. ISSN:
989 15524922. DOI: 10.1002/cyto.a.22463. URL: [http://doi.wiley.com/10.1002/
990 cyto.a.22463](http://doi.wiley.com/10.1002/cyto.a.22463).
- 991 **Guschanski2017** Katerina Guschanski, Maria Warnefors, and Henrik Kaessmann. “The
992 evolution of duplicate gene expression in mammalian organs”. In: *Genome Research*
993 27.9 (Sept. 2017), pp. 1461–1474. ISSN: 1088-9051. DOI: 10.1101/gr.215566.116.
994 URL: [http://www.ncbi.nlm.nih.gov/pubmed/28743766%20http://www.
995 pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5580707%20http:
996 //genome.cshlp.org/lookup/doi/10.1101/gr.215566.116](http://www.ncbi.nlm.nih.gov/pubmed/28743766%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5580707%20http://genome.cshlp.org/lookup/doi/10.1101/gr.215566.116).

- 997 **Iguchi2006** Naoko Iguchi, John W Tobias, and Norman B Hecht. “Expression profil-
998 ing reveals meiotic male germ cell mRNAs that are translationally up- and down-
999 regulated.” In: *Proceedings of the National Academy of Sciences of the United States*
1000 *of America* 103.20 (May 2006), pp. 7712–7. ISSN: 0027-8424. DOI: 10.1073/pnas.
1001 0510999103. URL: <http://www.ncbi.nlm.nih.gov/pubmed/16682651>
1002 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1472510>.
- 1003 **Immler2008** Simone Immler. “Sperm competition and sperm cooperation: the potential
1004 role of diploid and haploid expression.” In: *Reproduction (Cambridge, England)*
1005 135.3 (Mar. 2008), pp. 275–83. ISSN: 1741-7899. DOI: 10.1530/REP-07-0482.
1006 URL: <http://www.ncbi.nlm.nih.gov/pubmed/18299420>.
- 1007 **Joseph2004** Sarah B. Joseph and Mark Kirkpatrick. *Haploid selection in animals*. Nov.
1008 2004. DOI: 10.1016/j.tree.2004.08.004. URL: [https://www.sciencedirect.](https://www.sciencedirect.com/science/article/pii/S0169534704002381)
1009 [com/science/article/pii/S0169534704002381](https://www.sciencedirect.com/science/article/pii/S0169534704002381).
- 1010 **Keane2011** Thomas M. Keane et al. “Mouse genomic variation and its effect on phe-
1011 notypes and gene regulation”. In: *Nature* 477.7364 (Sept. 2011), pp. 289–294. ISSN:
1012 0028-0836. DOI: 10.1038/nature10413. URL: [http://www.nature.com/articles/](http://www.nature.com/articles/nature10413)
1013 [nature10413](http://www.nature.com/articles/nature10413).
- 1014 **Kleene2005** Kenneth C. Kleene. “Sexual selection, genetic conflict, selfish genes, and
1015 the atypical patterns of gene expression in spermatogenic cells”. In: *Developmental*
1016 *Biology* 277.1 (Jan. 2005), pp. 16–26. ISSN: 0012-1606. DOI: 10.1016/J.YDBIO.
1017 2004.09.031. URL: [https://www.sciencedirect.com/science/article/pii/](https://www.sciencedirect.com/science/article/pii/S0012160604006682)
1018 [S0012160604006682](https://www.sciencedirect.com/science/article/pii/S0012160604006682).
- 1019 **Lecuyer2007** Eric Lécuyer et al. “Global Analysis of mRNA Localization Reveals a
1020 Prominent Role in Organizing Cellular Architecture and Function”. In: *Cell* 131.1
1021 (Oct. 2007), pp. 174–187. ISSN: 00928674. DOI: 10.1016/j.cell.2007.08.003.
1022 URL: <http://www.ncbi.nlm.nih.gov/pubmed/17923096>
1023 <http://linkinghub.elsevier.com/retrieve/pii/S0092867407010227>.

- 1024 **Li2009** H. Li and R. Durbin. “Fast and accurate short read alignment with Burrows-
1025 Wheeler transform”. In: *Bioinformatics* 25.14 (July 2009), pp. 1754–1760. DOI: 10.
1026 1093/bioinformatics/btp324. URL: [https://academic.oup.com/bioinformatics/
1027 article-lookup/doi/10.1093/bioinformatics/btp324](https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp324).
- 1028 **Li2009a** H. Li, B. Handsaker, et al. “The Sequence Alignment/Map format and SAM-
1029 tools”. In: *Bioinformatics* 25.16 (Aug. 2009), pp. 2078–2079. ISSN: 1367-4803. DOI:
1030 10.1093/bioinformatics/btp352. URL: [http://www.ncbi.nlm.nih.gov/
1031 pubmed/19505943](http://www.ncbi.nlm.nih.gov/pubmed/19505943)[http://www.pubmedcentral.nih.gov/articlerender.
1032 fcgi?artid=PMC2723002](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2723002)[https://academic.oup.com/bioinformatics/
1033 article-lookup/doi/10.1093/bioinformatics/btp352](https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp352).
- 1034 **Liao2014** Y. Liao, G. K. Smyth, and W. Shi. “featureCounts: an efficient general purpose
1035 program for assigning sequence reads to genomic features”. In: *Bioinformatics* 30.7
1036 (Apr. 2014), pp. 923–930. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btt656.
1037 URL: [https://academic.oup.com/bioinformatics/article-lookup/doi/10.
1038 1093/bioinformatics/btt656](https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btt656).
- 1039 **Liu2014** Eric Yi Liu et al. “High-Resolution Sex-Specific Linkage Maps of the Mouse
1040 Reveal Polarized Distribution of Crossovers in Male Germline”. In: *Genetics* 197.1
1041 (May 2014), pp. 91–106. ISSN: 0016-6731. DOI: 10.1534/GENETICS.114.161653.
1042 URL: <https://www.genetics.org/content/197/1/91.short>.
- 1043 **Martin-DeLeon2005** Patricia A. Martin-DeLeon et al. “Spam I-associated transmission
1044 ratio distortion in mice: Elucidating the mechanism”. In: *Reproductive Biology and
1045 Endocrinology* 3.1 (2005), p. 32. ISSN: 14777827. DOI: 10.1186/1477-7827-3-32.
1046 URL: <http://rbej.biomedcentral.com/articles/10.1186/1477-7827-3-32>.
- 1047 **Meikar2014** Oliver Meikar et al. “An atlas of chromatoid body components.” In: *RNA*
1048 (*New York, N.Y.*) 20.4 (2014), pp. 483–95. ISSN: 1469-9001. DOI: 10.1261/rna.
1049 043729.113. URL: <http://www.ncbi.nlm.nih.gov/pubmed/24554440>.

- 1050 **Morales2002** Carlos R. Morales et al. “A TB-RBP and Ter ATPase Complex Accom-
1051 panies Specific mRNAs from Nuclei through the Nuclear Pores and into Intercel-
1052 lular Bridges in Mouse Male Germ Cells”. In: *Developmental Biology* 246.2 (June
1053 2002), pp. 480–494. ISSN: 0012-1606. DOI: 10.1006/DBIO.2002.0679. URL: <https://www.sciencedirect.com/science/article/pii/S0012160602906792>.
1054
- 1055 **Munch2016** Kasper Munch et al. “Selective Sweeps across Twenty Millions Years of Pri-
1056 mate Evolution”. In: *Molecular Biology and Evolution* 33.12 (Dec. 2016), pp. 3065–
1057 3074. ISSN: 0737-4038. DOI: 10.1093/molbev/msw199. URL: <https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msw199>.
1058
- 1059 **Nadeau2017** Joseph H Nadeau. “Do Gametes Woo? Evidence for Their Nonrandom
1060 Union at Fertilization.” In: *Genetics* 207.2 (Oct. 2017), pp. 369–387. ISSN: 1943-
1061 2631. DOI: 10.1534/genetics.117.300109. URL: <http://www.ncbi.nlm.nih.gov/pubmed/28978771>.
1062
- 1063 **Pybus2014** Marc Pybus et al. “1000 Genomes Selection Browser 1.0: a genome browser
1064 dedicated to signatures of natural selection in modern humans”. In: *Nucleic Acids*
1065 *Research* 42.D1 (Jan. 2014), pp. D903–D909. ISSN: 0305-1048. DOI: 10.1093/nar/
1066 gkt1188. URL: <http://www.ncbi.nlm.nih.gov/pubmed/24275494>
1067 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3965045>
1068 <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkt1188>.
1069
- 1070 **Refoyo-Martinez2019** Alba Refoyo-Martínez et al. “Identifying loci under positive se-
1071 lection in complex population histories.” In: *Genome research* (July 2019), gr.246777.118.
1072 ISSN: 1549-5469. DOI: 10.1101/gr.246777.118. URL: <http://www.ncbi.nlm.nih.gov/pubmed/31362936>.
1073
- 1074 **Ryder2018** Pearl V. Ryder and Dorothy A. Lerit. “RNA localization regulates diverse
1075 and dynamic cellular processes”. In: *Traffic* 19.7 (July 2018), pp. 496–502. ISSN:

1076 13989219. DOI: 10.1111/tra.12571. URL: [http://doi.wiley.com/10.1111/tra.](http://doi.wiley.com/10.1111/tra.12571)
1077 12571.

1078 **Salvatier2016** John Salvatier, Thomas V. Wiecki, and Christopher Fonnesbeck. “Prob-
1079 abilistic programming in Python using PyMC3”. In: *PeerJ Computer Science* 2
1080 (Apr. 2016), e55. ISSN: 2376-5992. DOI: 10.7717/peerj-cs.55. URL: [https :](https://peerj.com/articles/cs-55)
1081 [//peerj.com/articles/cs-55](https://peerj.com/articles/cs-55).

1082 **Schrider2016** Daniel R. Schrider and Andrew D. Kern. “S/HIC: Robust Identification
1083 of Soft and Hard Sweeps Using Machine Learning”. In: *PLOS Genetics* 12.3 (Mar.
1084 2016). Ed. by Bret A. Payseur, e1005928. ISSN: 1553-7404. DOI: 10.1371/journal.
1085 [pgen.1005928](https://dx.plos.org/10.1371/journal.pgen.1005928). URL: <https://dx.plos.org/10.1371/journal.pgen.1005928>.

1086 **Staubach2012** Fabian Staubach et al. “Genome Patterns of Selection and Introgression
1087 of Haplotypes in Natural Populations of the House Mouse (*Mus musculus*)”. In:
1088 *PLoS Genetics* 8.8 (Aug. 2012). Ed. by Michael H. Kohn, e1002891. ISSN: 1553-
1089 7404. DOI: 10.1371/journal.pgen.1002891. URL: [http://dx.plos.org/10.](http://dx.plos.org/10.1371/journal.pgen.1002891)
1090 [1371/journal.pgen.1002891](http://dx.plos.org/10.1371/journal.pgen.1002891).

1091 **Tarasov2015** Artem Tarasov et al. “Sambamba: fast processing of NGS alignment for-
1092 mats”. In: *Bioinformatics* 31.12 (June 2015), pp. 2032–2034. ISSN: 1367-4803. DOI:
1093 10.1093/bioinformatics/btv098. URL: [https://academic.oup.com/bioinformatics/](https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btv098)
1094 [article-lookup/doi/10.1093/bioinformatics/btv098](https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btv098).

1095 **Umehara2019** Takashi Umehara, Natsumi Tsujita, and Masayuki Shimada. “Activation
1096 of Toll-like receptor 7/8 encoded by the X chromosome alters sperm motility and
1097 provides a novel simple technology for sexing sperm”. In: *PLOS Biology* 17.8 (Aug.
1098 2019). Ed. by Yukiko M Yamashita, e3000398. ISSN: 1545-7885. DOI: 10.1371/
1099 [journal.pbio.3000398](http://dx.plos.org/10.1371/journal.pbio.3000398). URL: [http://dx.plos.org/10.1371/journal.pbio.](http://dx.plos.org/10.1371/journal.pbio.3000398)
1100 [3000398](http://dx.plos.org/10.1371/journal.pbio.3000398).

1101 **Ventela2003** Sami Ventelä, Jorma Toppari, and Martti Parvinen. “Intercellular organelle
1102 traffic through cytoplasmic bridges in early spermatids of the rat: mechanisms of
1103 haploid gene product sharing.” In: *Molecular biology of the cell* 14.7 (July 2003),
1104 pp. 2768–80. ISSN: 1059-1524. DOI: 10.1091/mbc.E02-10-0647. URL: <http://www.ncbi.nlm.nih.gov/pubmed/12857863>
1105 [http://www.pubmedcentral.
1106 nih.gov/articlerender.fcgi?artid=PMC165675](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC165675).

1107 **Veron2009** Nathalie Véron et al. “Retention of gene products in syncytial spermatids
1108 promotes non-Mendelian inheritance as revealed by the t complex responder.” In:
1109 *Genes & development* 23.23 (Dec. 2009), pp. 2705–10. ISSN: 1549-5477. DOI: 10.
1110 1101/gad.553009. URL: <http://www.ncbi.nlm.nih.gov/pubmed/19952105>
1111 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2788329>.

1112 **Zheng2001** Y. Zheng, X. Deng, and P.A. Martin-DeLeon. “Lack of Sharing of Spam1
1113 (Ph-20) among Mouse Spermatids and Transmission Ratio Distortion1”. In: *Biology
1114 of Reproduction* 64.6 (June 2001), pp. 1730–1738. ISSN: 0006-3363. DOI: 10.1095/
1115 biolreprod64.6.1730. URL: [https://academic.oup.com/biolreprod/article-
1116 lookup/doi/10.1095/biolreprod64.6.1730](https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod64.6.1730).

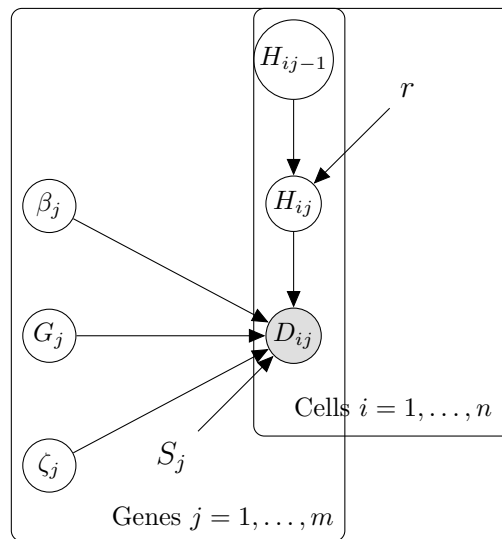


Figure 5: Graphical model for Bayesian method