Heterochrony explains convergent testis evolution in primates

Ekin Saglican[1], Ezgi Ozkurt[1], Haiyang Hu[2], Babur Erdem[1], Philipp Khaitovich[2], Mehmet Somel[1]

[1] Department of Biological Sciences, METU, Ankara, Turkey.

[2] CAS-MPG Partner Institute for Computational Biology, Shanghai, PRC.

*Corresponding author: msomel@metu.edu.tr

Abstract

The testis displays striking anatomical divergence among primates. Multi-male species, such as chimpanzees, have recurrently evolved large testicles relative to single-male species, such as humans. However, the developmental mechanisms behind testis divergence and whether they involve convergent molecular changes, have remained unknown. Through comparative analysis of transcriptomic data, we show that a species' testis expression profile, like testis size, can be a reliable indicator of mating type among primates, and possibly murids. Differential expression, in turn, largely reflects changes in the relative proportions of somatic/pre-meiotic versus meiotic/post-meiotic cell types. By studying mouse and macaque testis development, we find that single-male species' testis expression profiles are paedomorphic relative to multi-male species' profiles. For instance, human and gorilla testis profiles resemble those of adolescent mice. Our results suggest that heterochronic shifts involving conserved transcription regulators have been repeatedly employed in primate evolution, leading to rapid, convergent changes in testis size and histology.

Introduction

Studies on hominid evolution frequently focus on the human brain, which has tripled in size since our common last ancestor with chimpanzees [1-3]. But compared with change in testis size, hominid brain size evolution appears only minor. Chimpanzee testicles (150-170 g) can be 3-10 times larger than human testicles (16-50 g) [4-6]. This change is anatomically even more striking given that human body size is 30% larger than chimpanzee body size.

It has been hypothesized that multi-male mating systems, where a female mates with multiple males during estrous, leads to inter-male sperm competition, and would strongly select for high copulation frequency, higher ejaculate volume, and larger testis size [7]. This, indeed, could explain large testicles in the promiscuous chimpanzees, compared to humans who in many societies adopt monogamy and/or polygyny [5]. Testing this notion with primate data, Harcourt and colleagues found an obvious relationship between relative testis size (testis-body weight ratios) and mating systems, such that multi-male (polyandrous) species had larger testes than single-male (monandrous) species [6]. Later work likewise found relative testis size correlated with mating type, sperm competition, as well as multiple paternity, in different mammalian taxa [5, 8-10].

The distribution of mating systems and relative testis sizes within the primate phylogeny indicates that large and small testicles evolved multiple times, independently. For example, surveying only 33 primate species published by Harcourt and colleagues [6], we can predict that the ancestors of these taxa must have switched between the single-male, small testis mode and the multi-male, large testis mode at least 6 times. We can even observe changes within a single genus (*Papio*). Notably, testis histology and spermatogenesis rates also display convergent evolution. Multi-male species' testes contain 1.5-3 times higher proportion of seminiferous tubules to interstitial (connective) tissue, compared to ratios close to 1:1 in single-male primates [4, 6]. Species with large relative testis size also have higher rates of spermatogenesis than other species [11].

These observations suggest that the evolution of testis anatomy and histology can be rapid, possibly more so than in other tissues. This could be due to strong positive selection on testis and sperm phenotypes [5, 7, 12]; it could also be due to relatively modular organization of testis development [12, 13]. Unfortunately, molecular and developmental mechanisms of primate testis evolution have yet received meager attention. One exception has been the studies by Khaitovich and colleagues, who found significantly higher gene expression divergence, and low within-species diversity in the

testis compared to other tissues, including the brain and liver [14, 15]. These authors suggested that expression divergence in the testis could have been driven by positive selection on gene expression levels [14, 16]. Meanwhile, a recent study noted higher similarity between human and gorilla testis transcriptomes relative to chimpanzee (although this human-gorilla affinity was not quantified) [17]. The authors speculated on the possible role of chimpanzee promiscuity as a source of this expression pattern.

In this study, we first ask whether testis transcriptome profiles evolved in convergent fashion among species with similar mating types – as reported for size and histology. Finding a positive indication, we then link transcriptome divergence with cell type proportion differences within the testis, and further, with developmental timing differences between species, or heterochrony.

Results and Discussion

We combined two published adult testis transcriptome datasets [14, 17] comprising a total of 8 humans, 7 chimpanzees (including one bonobo), one gorilla and two rhesus macaques (Table S1) (Materials and Methods). Hierarchical clustering analysis showed that species' transcriptome profiles in this combined dataset group according to their known phylogeny (Figure S1). As reported previously, we found pervasive testis expression divergence between humans and chimpanzees [14]: among 9,017 common genes, 4,845 (54%) showed significant differential expression (t-test, Benjamini-Hochberg-corrected q-value < 0.10).

Transcriptome profile comparisons inform on mating type

Could testis gene expression differences between human and chimpanzee represent universal differences between single- and multi-male species? In order to address this question, we additionally collected published adult testis transcriptome data from the common marmoset, and from four non-primate mammalian species: house mouse, Norway rat, gray short-tailed opossum, and platypus. In this dataset, monadry is represented by human, gorilla and marmoset, whereas polyandry is represented by chimpanzee, rhesus macaque, mouse, rat, opossum, and platypus (Table 1 and references therein). We used a simple scheme, gauging whether each species' mean testis transcriptome is better correlated with that of humans or of chimpanzees, for which we have the best data. Among the seven taxa tested, the two single-male primates showed higher affinity to human, one multi-male primate and two multi-male rodents to

chimpanzee (Kolmogorov-Smirnov test, $p<0.02$ in each test) (Figure 1, Table S2). The gray short-tailed opossum showed no trend, and the platypus showed higher affinity to human. Thus, primates and rodents appear to support our hypothesis, whereas more distant, non-eutherian mammals, do not.

Notably, the three primate comparisons are phylogenetically independent: First, gorilla, macaque, and marmoset are equally distant to either human or chimpanzee. Second, gorilla and macaque are phylogenetically closer to each other than to the marmoset, but display contrasting affinities. On the other hand, the mouse and rat are sister taxa and therefore their results are not independent. Overall, the hypothesis that transcriptome differences reflect mating type is supported in four independent lineages among primates and murids. This is suggestive but not conclusive evidence.

Cell type proportion shifts reflect mating type

We next hypothesized that the monandrous and polyandrous transcriptome profiles, at least among primates and murids, could reflect parallel shifts in cell type composition. As mentioned before, chimpanzee testicles contain a higher concentration of seminiferous tubules than interstitial tissue, and produce more sperm compared to human testicles [5, 6, 12].

Chalmel and colleagues had purified testis cell types from mice, rats and humans, and reported evolutionarily conserved differences between transcriptomes of pre-meiotic/somatic cells (PRE; including spermatogonia and Sertoli cells) and meiotic/post-meiotic cells (POST; including spermatocytes and spermatids) [18]. Using this and another mouse cell type dataset [namekawa], we confirmed that independent of species origin, PRE cell type transcriptomes cluster to the exclusion POST cell types (data not shown). We then used the combined mouse PRE and POST profiles to predict the relative proportions of PRE and POST cells within testis tissue of each species. We did not use human cell types, so as not to bias the analysis.

Employing a deconvolution scheme based on linear regression on 2792 common genes (Materials and Methods), we found that, as hypothesized, POST profiles are represented at least twice more common than PRE profiles among all four multi-male primates and murids, and vice versa among all three single-male primates (Figure 2). This observation fits well with histological data on tubule:interstitia proportions [5, 6], and that the tubule transcriptome is dominated by POST cell type profiles [18]. Thus, at least among eutherian mammals, convergent cell type composition changes appear to have evolved

in distinct taxa. If we consider human-gorilla and mouse-rat as sister taxa, our hypothesis appears to have been confirmed in 5 independent cases (one-sided binomal test p=0.03).

Single-male species' testes paedomorphic relative to multi-male species'

Cell type composition shifts could arise from hetero

chronic shifts in conserved developmental pathways [19, 20]. For example, by extending or accelerating a developmental pathway involving progenitor germ cell production during sexual maturation, males of a lineage that recently switched to polyandry could evolve larger testes, enabling the production of copious amounts of sperm. To test this idea, we used a published testis development transcriptome series, where testis tissue had been sampled from newborn to adult mice [21]. We first constructed loess-based interpolated curves for each gene, and then compared the interpolated mice transcriptome profiles with average PRE and POST cell type transcriptome profiles (Materials and Methods). As might be expected, PRE cells showed highest correlation to pre-adolescent mice, and POST cells, to adult mice (Figure 3A). We next added mean adult human, chimpanzee, gorilla and macaque expression profiles to the analysis. This revealed that, while adult chimpanzee and macaque testes show highest affinity to those of adult mice, human and gorilla adult testes show highest affinity to adolescent mice, ~20 days of age. By comparing individual human and chimpanzee profiles (Figure 3B), we further confirmed that the species difference was significant (one-sided Mann-Whitney U test, p=0.0007). Thus, single-male species' testis transcriptome profiles appear paedomorphic relative to those of multi-male species.

To confirm this result, we next generated testis development microarray dataset from 12 macaques ranging from newborns to old adults (Materials and Methods). Performing the same analysis as above, we found a trend in the similar direction, although more modest (Figure 3C). Nevertheless, relative to chimpanzees, humans showed higher affinity to younger macaques at a marginally significant level (Figure 3D; one-sided Mann-Whitney U test, p=0.05).

We then asked whether such relative timing difference might be found in other organs, such as the brain. We had previously reported that postnatal human brain development trajectories can be neotenic relative to those of chimpanzees and other primates, but that this trend involves only specific gene groups [22, 23]. Because we wished to apply the above-described test to the brain, we generated a mouse prefrontal cortex

development transcriptome dataset, including 8 mice with ages ranging from newborns to 122 days (Materials and Methods). We then compared mouse brain development profiles with those of adult humans, chimpanzee, gorilla and macaque (Figure 3E). All primates showed highest affinity to young adult mice (~40 days; Figure 3F). The level of heterochrony we observe in testis development, with human and gorilla adults showing affinity toward *adolescent* stages of mice, might be a unique phenomenon.

A role of RFX1 in testis transcriptome divergence

To determine whether the same regulatory factors might play role in transcriptome divergence in different primates, we clustered genes based on their similarity in adult primate, macaque and mouse development, and PRE and POST expression (Materials and Methods). We defined 4 groups, although choosing other numbers yields similar results (Figure S2). One gene group, Cluster 4, was enriched in multiple spermatogenesis- and cell division-related functional processes (Bonferroni adjusted Fisher's exact test $p<0.05$, Table S3). This was also the Cluster showing high expression in POST cells and multi-male species, leading us to speculate that it represents germ stem cell development-related genes (Figure 4A).

We then sought common transcription factors that target cluster 4 genes using the Transfac database [24]. We found one such common regulator, RFX1, which showed enrichment in Cluster 4 genes (Bonferroni adjusted Fisher's exact test $p=0.002$). RFX1 had predicted binding sites among 14% of Cluster 4 genes' promoters. Notably, this transcription factor was also previously implicated in testis development [25, 26]. Checking the *RFX1* expression profile among the four taxa, we noticed higher expression in both chimpanzee and macaque, relative to human and gorilla (Kolmogorov-Smirnov test $p=0.02$ using only hominids, $p=0.006$ using all four species). Thus, assuming the primate common ancestor was monandrous, the observed transcriptome divergence patterns could be partly explained by convergent *RFX1* up-regulation in chimpanzee and macaque. Assuming the primate common ancestor was polyandrous, *RFX1* down-regulation in the ape common ancestor would be required, followed by its up-regulated in chimpanzee.

Conclusion

Relative testis size has been established as a useful method to predict a species' mating type, with respect to monandry and polyandry (e.g. [6, 8-10]). This implies rapid and

convergent evolution of testis size. Here we have shown that the testis transcriptome also evolves in convergent fashion and in parallel with mating type and relative size, at least among primates and murids. In the more distant opossum and platypus, the relationship between mating type and testis transcriptome (as detected in primates) disappears. Given these results, future work on comparative transcriptomics could inform on mating systems of ecologically understudied species, and supplement relative testis size information. However, our results also suggest that this approach will function within a limited phylogenetic range. Here we found consistent patterns only across primates and murids, which are phylogenetically closer to each other than other to most other mammals.

Our results also provide insight into heterochronic mechanisms of rapid testis evolution. We hypothesize as follows: In primates and murids, postnatal testis development follows a conserved pathway, whereby there is increased accumulation of spermatocytes and spermatids with age. In species that evolve under a multi-male regime, positive selection shifts this developmental pathway towards prolonged activity, or hyper-activation. This leads to an increased proportion of seminiferous tubules and POST cell types in the testis, and enables faster spermatogenesis [11]. In turn, in species that switch to a monandrous regime, mutations causing slower activation or earlier developmental cessation eventually fix in the population. This leads to a paedomorphic testis profile, in transcriptome, histology, and size; it also leads to slower spermatogenesis.

We have detected one common transcription factor that could be involved in driving convergent transcriptome changes. RFX1 is member of a family of transcription factors that play role in spermatogenesis [25] and shows particularly high expression in spermatids [26] (although its exact role in spermatogenesis is less known than other RFX members, such as RFX2). In addition to binding site enrichment, we find that *RFX1* is expressed at higher levels in adult chimpanzee and macaque testes relative to those of human and gorilla.

Hence, some of the observed testis expression divergence patterns could be driven by convergent *cis*- or *trans*- changes in *RFX1* regulation. These particular genetic changes, as well as the role of positive selection and/or relaxation of constraints in driving them, remain to be identified.

Materials and Methods

Ethics statement: Samples used for microarray analysis in this study were derived from animals that died of reasons independent of this work. Biomedical Research Ethics Committee of Shanghai Institutes for Biological Sciences completed the review of the use and care of the animals within the research project (approval ID: ER-SIBS-260802 P).

Published datasets – primate dataset 1: This RNA-sequencing dataset was published in [17]. The testis dataset includes 2 adult humans (*Homo sapiens*), 2 adult chimpanzees (1 *Pan paniscus* and 1 *Pan troglodytes*, which we do not distinguish as explained below), 1 adult gorilla (*Gorilla gorilla*), 2 adult rhesus macaques (*Macaca mulatta*). The prefrontal cortex dataset includes 2 male humans, 6 male chimpanzees, and the same number of individuals for the other species as the testis dataset (we did not include 3 humans that showed unusual profiles and did not group with other humans; data not shown). This RNA sequencing dataset was generated on the Illumina Genome Analyser IIx platform. RPKM expression levels were called using constitutive, orthologous and alignable primate exons, and summarized as Ensembl genes by the authors. We redefined these expression levels as log2(x+1), where x stands for the RPKM level per gene.

Published datasets – primate dataset 2: This microarray dataset was published in [14]. It comprises 6 adult human and 5 adult chimpanzee cerebral prefrontal cortex and testis tissue samples. The data was generated using Affymetrix Human HGU133Plus2 microarrays. We downloaded the raw CEL files from EBI ArrayExpress (www.ebi.ac.uk/arrayexpress/), under accession E-AFMX-11. We masked probes that do not match either human or chimpanzee genomes, and used a probeset definition based on Ensembl genes (http://www.ensembl.org/) [27, 28], as described in [22]. Affymetrix CEL files were processed using the Bioconductor “affy” package “rma” function [29], which includes background subtraction, log transformation, normalization, and summary across probesets.

Published datasets – marmoset: This RNA-sequencing dataset was published in [30], generated on the Illumina Genome Analyser IIx platform. The processed data was obtained from the file titled “GSE50747_marmoset_normalized_all_tissues_FPKM.txt”, downloaded from NCBI GEO (www.ncbi.nlm.nih.gov/geo/) under the accession number GSE50747, where the authors had calculated average FPKM values for marmoset

(*Callithrix jacchus*) testis. We mapped these values to orthologous human genes, only including "1-to-1 orthologs" as defined by Ensembl (v.77) [27]. We redefined expression levels as log2(x+1), where x stands for the FPKM level per gene.

Published datasets – amniote dataset: This RNA-sequencing dataset was published in [17]. This testis dataset includes the same individuals described above in "primate dataset 1", as well as additional non-primate amniotes. Among these, we used testis data from 2 adult mice (*Mus musculus*) and 2 adult opossum (*Monodelphis domestica*) individuals. The dataset also differs from "primate dataset 1" in that, it was constructed using constitutive, orthologous and alignable amniote exons [17]. The data was generated as described above, and we followed the same steps used in preprocessing "primate dataset 1".

Published datasets – rat dataset: This dataset was published in [18]. It was generated using Affymetrix Rat 230.2 microarrays. Rat Affymetrix CEL files were downloaded from EBI ArrayExpress under accession number E-TABM-130. These were processed using the Bioconductor "affy" package "rma" function [29]. Affymetrix probeset IDs were mapped to rat genes and further to orthologous human Ensembl genes following the above-described procedure. From this dataset, only the two adult rat testis profiles were chosen.

Published datasets – mouse testis development: This microarray dataset was published in [21]. This testis dataset includes 15 mice testis samples from individuals, with ages ranging from 0 days to adult (assumed 42 days of age; www.genomics.senescence.info/species, [31]). Affymetrix CEL files were downloaded from NCBI GEO [ref] with accession number GSE640, processed using the Bioconductor "affy" package "rma" function [29]. Affymetrix probeset IDs were converted to Ensembl gene IDs (Ensembl v.77) downloaded from the Ensembl Biomart (www.ensembl.org/biomart/). For each gene that matched multiple probesets, we chose the probeset with the highest mean expression level. We then mapped mouse genes to orthologous human genes, only including "1-to-1 orthologs" as defined by Ensembl.

Published datasets – mouse testis cell types: This data was published in [18] and [32]. Both datasets include expression profiles from adult mouse spermatogonia,

spermatocytes, and spermatids. In addition, the first dataset includes purified adult mouse Sertoli cells. Both had used Affymetrix Mouse 430.2 microarrays. Affymetrix CEL files were downloaded from EBI ArrayExpress under accession number E-TABM-130, and processed using the Bioconductor "affy" package "rma" function [29]. Affymetrix probeset IDs were mapped to mouse genes and further to orthologous human Ensembl genes following the above-described procedure.

Novel datasets – mouse brain development: Post-mortem prefrontal cortex samples were obtained from 8 C57BL/6 mice aged between 2 days and 122 days, housed and fed under standard conditions in the MPI-EVA Animal Facility (Leipzig, Germany). The mice were sacrificed for reasons independent of this study, their tissues were harvested and frozen immediately, and stored at -80°C. We extracted total RNA from supplied frozen prefrontal cortex tissue, using the Trisol reagent. RNA was processed and hybridized to Affymetrix Mouse Gene 1.0ST arrays following standard Affymetrix protocols. The raw data is deposited to NCBI GEO under accession number (pending). Affymetrix CEL files were processed using the Bioconductor "oligo" package "rma" function [33]. Affymetrix probeset IDs were mapped to mouse genes and further to orthologous human Ensembl genes following the above-described procedure.

Novel datasets – macaque testis development: Post-mortem testis samples from 12 rhesus macaques of different ages (16 days to 26 years) were obtained from the SuZhou Experimental Animal Center (SuZhou, China). The individuals were housed under standard conditions, were healthy, and died of causes with no relation to the tissue used, and of reasons unrelated to this study. Whole testis samples were stored at -80°C. Total RNA from ~100 mg tissue was isolated using the Trisol reagent, and hybridized to Affymetrix Human Gene 1.0ST arrays following standard Affymetrix protocols. The raw data is deposited to NCBI GEO under accession number (pending). Affymetrix CEL files were processed using the Bioconductor "oligo" package "rma" function [33]. Affymetrix probeset IDs were mapped to human Ensembl genes following the above-described procedure.

Combining primate datasets: We applied the following strategy to combine primate dataset 1 and 2. We first joined datasets based on 9,088 common Ensembl genes. As both datasets included adult human and chimpanzee testis samples, we normalized

mean and standard deviation (s.d.) of each gene in each dataset based on human and chimpanzee samples. For each gene, we calculated the mean expression value ($\mu_{h,c}$) and standard deviation ($\sigma_{h,c}$) for equal numbers of humans and chimpanzees in each dataset (n=5 each for dataset 1, and n=2 each for dataset 2). In each dataset, from the expression values for each gene, we subtracted the corresponding $\mu_{h,c}$ and divided by $\sigma_{h,c}$. This procedure shifts each genes' expression level (for humans and chimpanzees) to an average of ~0.

Finally, in order not to loose information on expression level differences across genes, for each gene, we calculated the mean expression value for 2 humans and 2 chimpanzees from each dataset (8 in total), and added this value to all individuals' normalized expression values in both datasets.

The final combined dataset was analyzed using hierarchical clustering based on Euclidean distance, using the "hclust" function in R.

The result indicated that species cluster according to the known phylogeny [34], and samples from the two datasets cannot be distinguished (Figure S1). We thus concluded that our approach could effectively remove major sources of between-dataset technical variance.

Note that in the original publication for dataset 1 [17], the authors had reported that gorilla and human appear as sister taxa, to the exclusion of the chimpanzee. We do not observe this pattern in either dataset 1 or the combined dataset, which might be related to the exact genes incorporated in the analysis.

Differential expression. Gene expression differences between humans and chimpanzees was tested for each gene using a two-sided t-test, and correcting for multiple testing by converting p-values to q-values with the Benjamini-Hochberg method [35], as implemented in the "p.adjust" function in R.

Combining cell type datasets. Previous reports indicate that the main shift in transcriptome profiles during spermatogenesis occurs during meiosis – such that pre-meiotic and somatic cell types' profiles cluster together, while meiotic and post-meiotic cell types profiles create another cluster [18]. Given this result, we combined the two mouse testis cell type datasets such that each contained 4 pre-meiotic/somatic and 4 meiotic/post-meiotic samples – thus maintaining balance. To remove technical influence on gene expression, we normalized each genes' mean to 0, and standard deviation to 1,

in each dataset. This yielded a dataset of 14,934 human Ensembl genes. Using hierarchical clustering as described above, we determined that pre-meiotic/somatic and meiotic/post-meiotic samples cluster together (Figure S3), and dataset origins cannot be separated.

Bonobo and common chimpanzee: As mentioned above, throughout the analysis we treated the single bonobo (*Pan paniscus*) individual in "primate dataset 1", as a chimpanzee. In our analysis, this individual's testis profile showed no detectable difference from *Pan troglodytes* individuals (data not shown).

Correlation analysis with hominids: We checked whether across the transcriptome, non-hominid species might show higher affinity to humans or chimpanzees. For this, we calculated Spearman correlation coefficients between each hominid individual and that of gorilla, macaque, marmoset, mouse, and opossum. For mouse and opossum, for which we have data from 2-3 individuals each, we used the mean expression level per gene across all individuals. Testing each individual separately gives qualitatively the same results (data not shown). For each species, we then tested for significant differences between its correlation coefficients with humans, and with chimpanzees, using the non-parametric two-sided Kolmogorov-Smirnov test. The number of genes used in each comparison, and the mean correlation coefficients with human and chimpanzee, and the Kolmogorov-Smirnov test p-values are shown in Table S2 and Figure 1.

Cell type proportion prediction: Here we predicted the relative contribution of pre-meiotic/somatic (PRE) versus meiotic/post-meiotic (POST) cell types to the overall adult testis transcriptomes using an approach related to deconvolution [36]. First, for each of the 14,934 genes in the cell type dataset, we calculated mean expression levels for all PRE (n=8) and POST (n=8) samples, which we call E_{PRE} and E_{POST}. Next, using these average expression levels, and overlapping genes' expression values in each whole testis sample (E_{WT}), we constructed a linear regression model:

$$E_{WT} = a + b_{PRE}*E_{PRE} + b_{POST}*E_{POST} + \text{error} \quad (1)$$

where a represents the intercept, and b_{PRE} and b_{POST} represent regression coefficients.

We then calculated the ratio $\log_2(b_{POST} / b_{PRE})$ for each individual, which we expect to reflect the relative contribution of POST cells to the tissue average (Figure 1B). We

further repeated the analysis using average expression levels for all individuals of a species. We finally tested lower $\log_2$ ratios in monoandrous than polyandrous species, using a one-sided Mann-Whitney U test.

Heterochrony analysis: Our goal is understand whether testis development in closely related monoandrous and polyandrous species, such as human and chimpanzee, differ in timing. Given severe limitations in such sample acquisition, we resorted to a indirect approach. We first used the mouse testis development dataset described above. We first identified genes showing significant change with mouse age using the Spearman correlation test, and correcting for multiple testing using the Benjamini-Hochberg method, as implemented in the "p.adjust" function in R (at q-value < 0.10). For each of these 6010 genes, we interpolated loess regression curves of expression change with age using the "loess" function in R, across 20 points, and with the degree parameter fixed at 1.

We then calculated Spearman correlation coefficients between these interpolated mouse testis expression levels, and each adult species' testis profile, at each interpolated point, across all 1621 common genes. We performed this analysis using the mean expression level for each species and each cell type. The correlation coefficients were plotted against mouse age. To better observe differences among correlation coefficient curves, we further normalized each to mean=0, and s.d.=1.

To test whether the human curve is significantly shifted to earlier ages compared to the chimpanzee curve, we again calculated correlation coefficients between each hominid and interpolated mouse profile. For each hominid, we then determined the mouse age at which the correlation is maximized. Finally, we tested whether these maximum expression similarity ages are lower for humans than for chimpanzees, using a one-sided Mann-Whitney U test.

Regulatory analysis: To identify potential common regulators of testis development that show differential expression across primates, we first clustered gene expression data from the combined primate dataset, the mouse and macaque testis developmental datasets, and mouse cell type datasets. In each dataset, we normalized each genes' expression to mean=0, and s.d.=1, to be able to best use information on variation among samples. We clustered the genes into 4 major expression groups using the k-means algorithm in R. We then tested enrichment in Gene Ontology categories, downloaded

from the Ensembl Biomart database (v.77), using the Fisher's exact test, and correcting for multiple testing using the Bonferroni correction. Second, we tested each cluster for common transcription factor binding sites in their promoters, using the Transfac database [24], as described previously [37]. Again, we used the Fisher's exact test and Bonferroni correction.

Acknowledgements: We thank Duha Alioglu, Aysegul Dede and all members of the Comparative and Evolutionary Biology Group at METU, members of the Comparative Biology group at Shanghai, Tulin Cetin, and Mesut Muyan for helpful discussion and suggestions. We thank Tulin Yanik for support. M.S. was supported by a TUBITAK 2232 fellowship and a Science Academy-Turkey Award (BAGEP-2014).

References:

1. Carroll, S., Genetics and the making of Homo sapiens. *Nature*, 2003. 422(6934): p. 849-857.
2. Sherwood, C.C., et al., Human brain evolution writ large and small. *Prog Brain Res*, 2012. 195: p. 237-54.
3. Herculano-Houzel, S., The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proc Natl Acad Sci USA*, 2012. 109 Suppl 1: p. 10661-8.
4. Schultz, A., The relative weight of the testes in primates. *Anat Rec*, 1938. 72(3): p. 387-394.
5. Dixson, A.F., *Primate Sexuality: Comparative Studies of the Prosimians, Monkeys, Apes, and Humans*. 2012: Oxford University Press.
6. Harcourt, A., et al., Testis weight, body weight and breeding system in primates. *Nature*, 1981. 293(5827): p. 55-57.
7. Short, R.V., Sexual Selection and Its Component Parts, Somatic and Genital Selection, as Illustrated by Man and the Great Apes. *Advances in the Study of Behavior*, 1979. Volume 9: p. 131-158 T2 -.
8. Kenagy, G. and S. Trombulak, Size and Function of Mammalian Testes in Relation to Body Size. *Journal of Mammalogy*, 1986. 67(1): p. 1-22.
9. Ramm, S.A., G.A. Parker, and P. Stockley, Sperm competition and the evolution of male reproductive anatomy in rodents. *Proc Biol Sci*, 2005. 272(1566): p. 949-55.
10. Hosken, D.J., Sperm competition in bats. *Proc Biol Sci*, 1997. 264(1380): p. 385-92.
11. Ramm, S.A. and P. Stockley, Sperm competition and sperm length influence the rate of mammalian spermatogenesis. *Biol Lett*, 2010. 6(2): p. 219-21.
12. Stockley, P., Sperm competition in mammals. *Hum Fertil (Camb)*, 2004. 7(2): p. 91-7.
13. Sekido, R. and R. Lovell-Badge, Genetic control of testis development. *Sex Dev*, 2013. 7(1-3): p. 21-32.

14. Khaitovich, P., et al., Parallel Patterns of Evolution in the Genomes and Transcriptomes of Humans and Chimpanzees. *Science*, 2005. 309(5742): p. 1850-1854.
15. Khaitovich, P., et al., Functionality of Intergenic Transcription: An Evolutionary Comparison. *PLoS Genetics*, 2006. 2(10): p. e171.
16. Khaitovich, P., et al., Evolution of primate gene expression. *Nat Rev Genet*, 2006. 7(9): p. 693-702.
17. Brawand, D., et al., The evolution of gene expression levels in mammalian organs. *Nature*, 2011. 478(7369): p. 343-8.
18. Chalmel, F., et al., The conserved transcriptome in human and rodent male gametogenesis. *Proceedings of the National Academy of Sciences*, 2007. 104(20): p. 8346-8351.
19. Shea, B., Heterochrony in human evolution: The case for neoteny reconsidered. *American Journal of Physical Anthropology*, 1989. 32(S10): p. 69-101.
20. McKinney, M. and K. McNamara, *Heterochrony: The evolution of ontogeny*. 1991.
21. Schultz, N., F. Hamra, and D. Garbers, A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. *PNAS*, 2003. 100(21): p. 12201-12206.
22. Somel, M., et al., Transcriptional neoteny in the human brain. *Proc Natl Acad Sci USA*, 2009. 106(14): p. 5743-5748.
23. Liu, X., et al., Extension of cortical synaptic development distinguishes humans from chimpanzees and macaques. *Genome Res*, 2012.
24. Kel, A.E., et al., MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Research*, 2003. 31(13): p. 3576-9.
25. Grimes, S.R., Testis-specific transcriptional control. *Gene*, 2004. 343(1): p. 11-22.
26. Kistler, W.S., et al., Differential expression of Rfx1-4 during mouse spermatogenesis. *Gene Expr Patterns*, 2009. 9(7): p. 515-9.
27. Hubbard, T., et al., Ensembl 2007. *Nucl Acids Res*, 2007. 35(suppl_1): p. D610-617.
28. Dai, M., et al., Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucl Acids Res*, 2005. 33(20): p. e175-.
29. Gautier, L., et al., affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*, 2004. 20(3): p. 307-315.
30. Cortez, D., et al., Origins and functional evolution of Y chromosomes across mammals. *Nature*, 2014. 508(7497): p. 488-93.
31. de Magalhães, J.P. and J. Costa, A database of vertebrate longevity records and their relation to other life-history traits. *Journal of Evolutionary Biology*, 2009. 22(8): p. 1770-4.
32. Namekawa, S., et al., Postmeiotic Sex Chromatin in the Male Germline of Mice. *Current Biology*, 2006. 16(7): p. 660-667.
33. Carvalho, B.S. and R.A. Irizarry, A framework for oligonucleotide microarray preprocessing. *Bioinformatics*, 2010. 26(19): p. 2363-7.
34. Uddin, M., et al., Sister grouping of chimpanzees and humans as revealed by genome-wide phylogenetic analysis of brain gene expression profiles. *PNAS*, 2004. 101(9): p. 2957-2962.
35. Benjamini, Y. and Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 1995. 57(1): p. 289-300.

36. Gong, T., et al., Optimal deconvolution of transcriptional profiling data using quadratic programming with application to complex clinical blood samples. *PLoS ONE*, 2011. 6(11): p. e27156.
37. Somel, M., et al., MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res*, 2010. 20: p. 1207-1218.

Figure legends:

Figure 1: Correlation between mean adult testis profiles with hominids. The y-axis shows Spearman correlation coefficients of gorilla, macaque, marmoset, rat, mouse, opossum, platypus, with those of humans (n=8) and chimpanzees ("chimp"; n=7). Only genes showing significant differential expression between human and chimpanzee are used (4845 to 2217 genes, depending on comparison). "m.m" and "s.m" next to species' names indicate multi- and single-male mating types, respectively. The full results, as well as results using all detected genes, is listed in Table S2. Asterisks indicate (*): Kolmogorov-Smirnov test $p<0.05$; (**): $p<0.01$; (***): $p<0.001$.

Figure 2: Predicted POST/PRE cell type ratios (log2) for each adult individuals' testis profile among primate and murid species. Blue and orange boxes stand for multi- and single-male species, respectively. The ratios are predicted using linear regression with mean POST and PRE cell type profiles. In this analysis, the same set of 2792 common genes were used.

Figure 3: Heterochrony analysis comparing mouse and macaque developmental profiles with those of adult primates and adult cell type profiles. **(A)** Correlation between mean adult testis profiles of human, chimpanzee, gorilla, and macaque, as well as mean POST and PRE cell types, with those of mice aged 0 to adults, across common genes. The y-axis shows Spearman correlation coefficients normalized to mean=0 and s.d.=1 for each comparison; the x-axis shows mouse age in days. **(B)** The same analysis as in Panel A, but using individual humans (red) and chimpanzees (blue). **(C)** The same analysis as in Panel A, but using macaque testis development data, calculating correlations across 3492 common genes. The y-axis represents macaque ages. **(D)** The same analysis as in Panel B, applied to macaque testis development data. **(E)** The same analysis as in Panel A, but performed by comparing mouse brain development data and adult primate brain data, calculating correlations across 3898 common genes.

Figure 4: Expression profile of cluster 4 genes **(A)** and their predicted regulatory *RFX1* **(B)**. In panel A, the y-axis shows mean expression levels (after normalizing each gene to mean=0 and s.d.=1) across all genes in the cluster. Each point represents an individual or cell type sample, with identities indicated within the figure.

Table Legends:

Table 1: The species used in the analysis, their mating types, and sources of expression data.

Supporting Figure Legends:

Figure S1: A hierarchical clustering plot of the combined primate dataset. ds1 and ds2 refer to datasets 1 and 2, respectively, as described in Materials and Methods.

Figure S2: Expression profile of 4 cluster gene clusters, constructed using kmeans anlaysis. The y-axis shows mean expression levels (after normalizing each gene to mean=0 and s.d.=1) across all genes in the cluster. Each point represents an individual or cell type sample, with identities indicated within the figure.

Figure S3: A hierarchical clustering plot of the combined cell type profiles dataset. ds1 and ds2 refer to datasets 1 and 2, respectively, as described in Materials and Methods.

Supporting Table Legends:

Table S1: Description of datasets used in the study, including microarray or sequencing platforms, the number of individuals, and number of human genes.

Table S2: Comparisons of correlations with humans and chimpanzees for 7 mammalian species. KS test: Kolmogorov-Smirnov test.

Table S3: Results for transcription factor binding site and Gene Ontology Biological Process enrichment for the 4 gene clusters.

Figure 1:

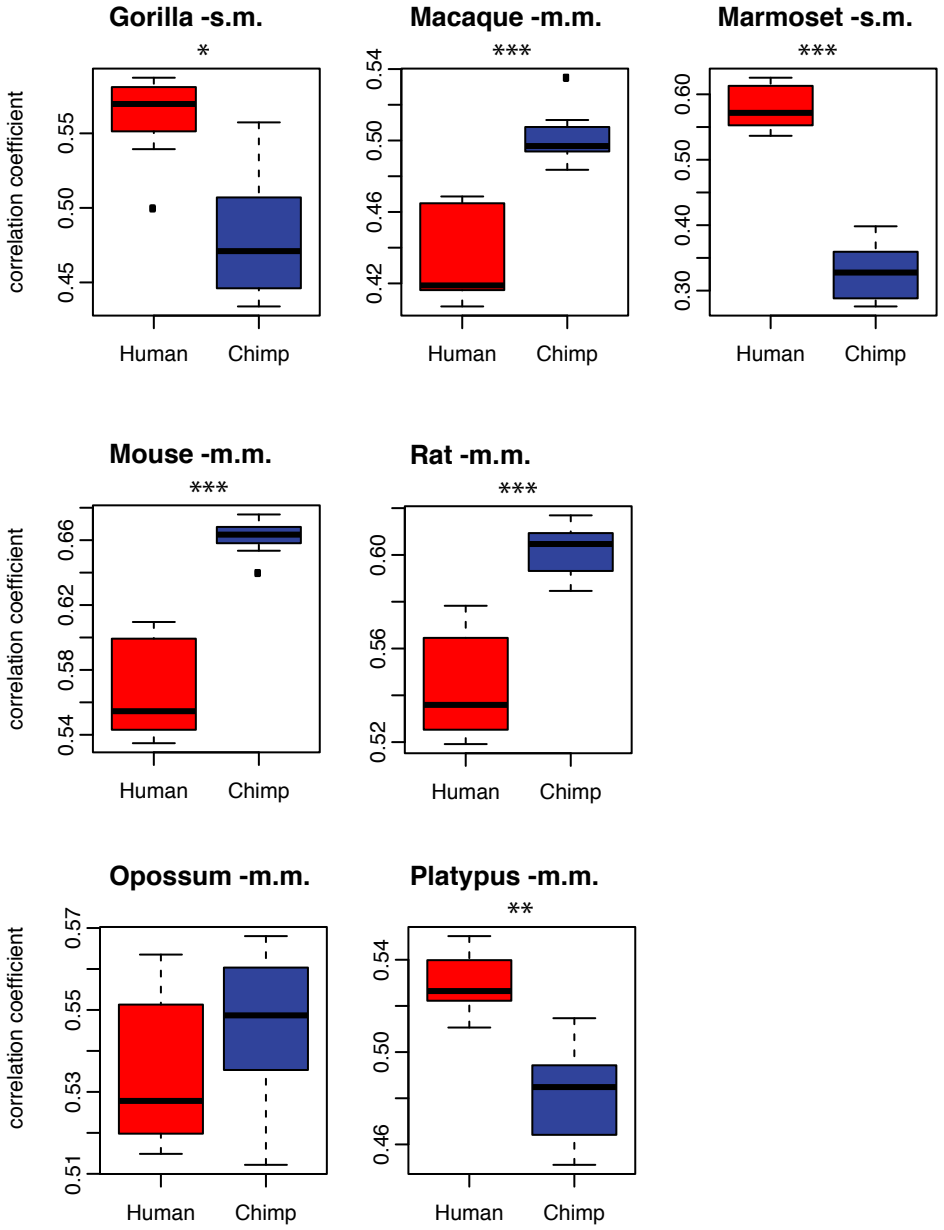

Figure 2:

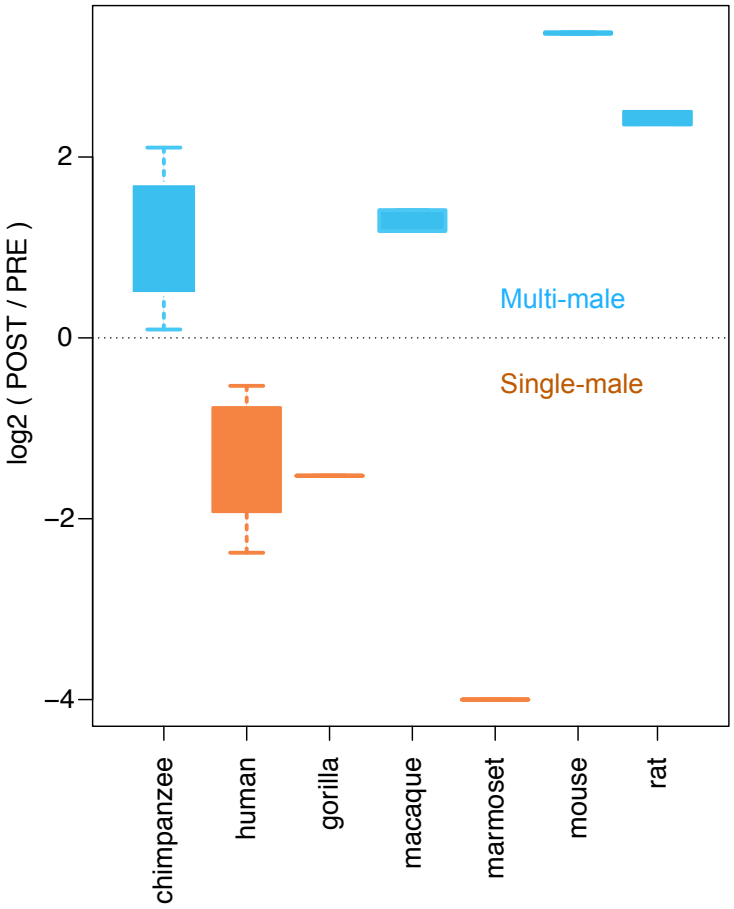

Figure 3:

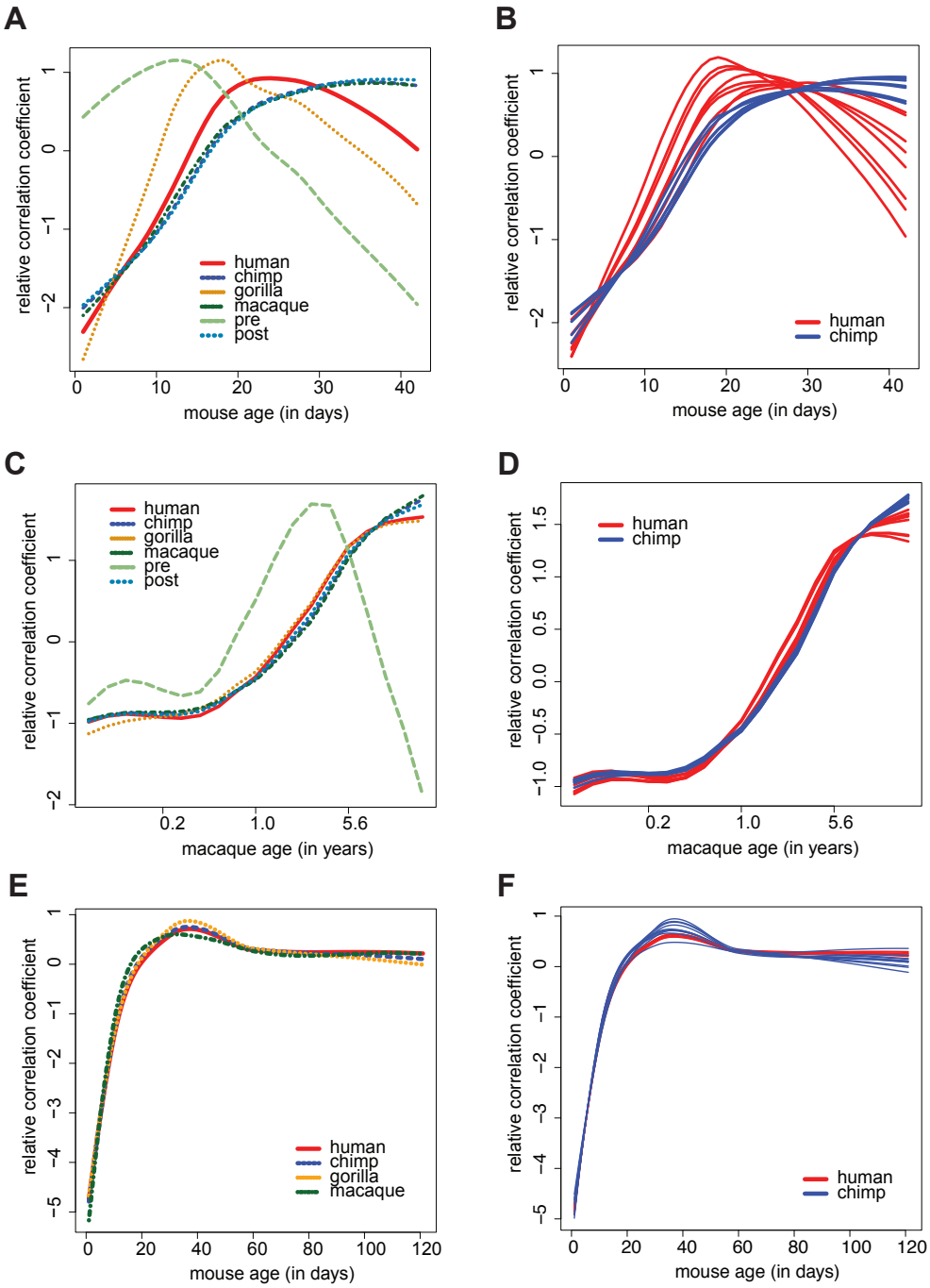

Figure 4:

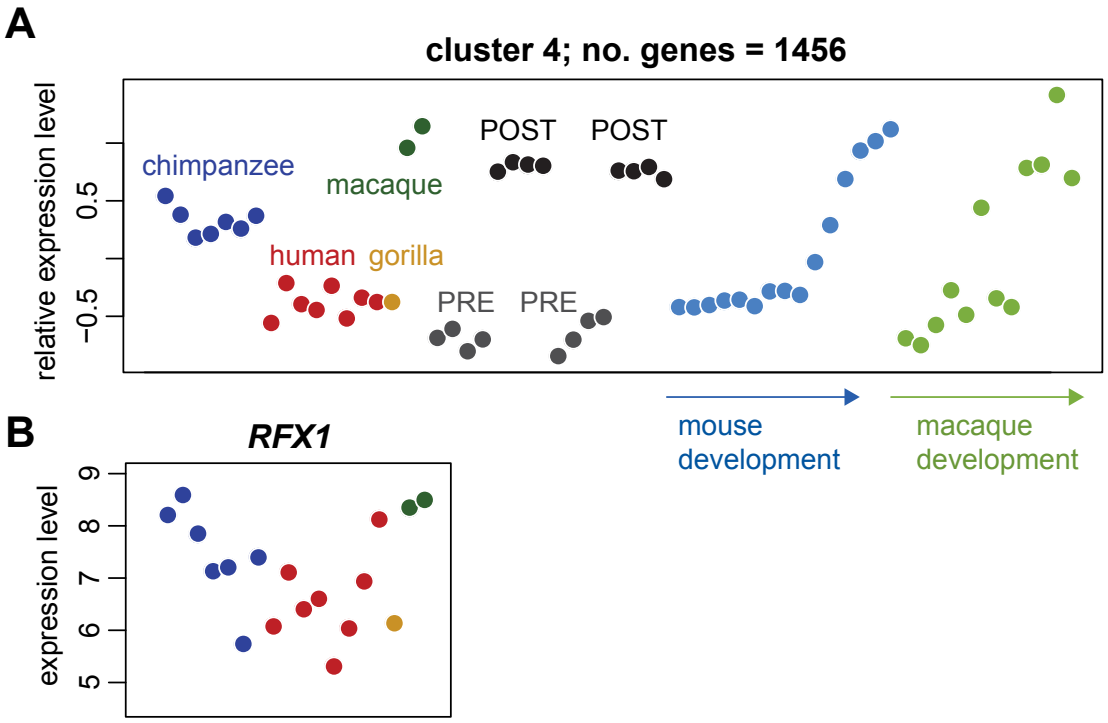

Figure S1:

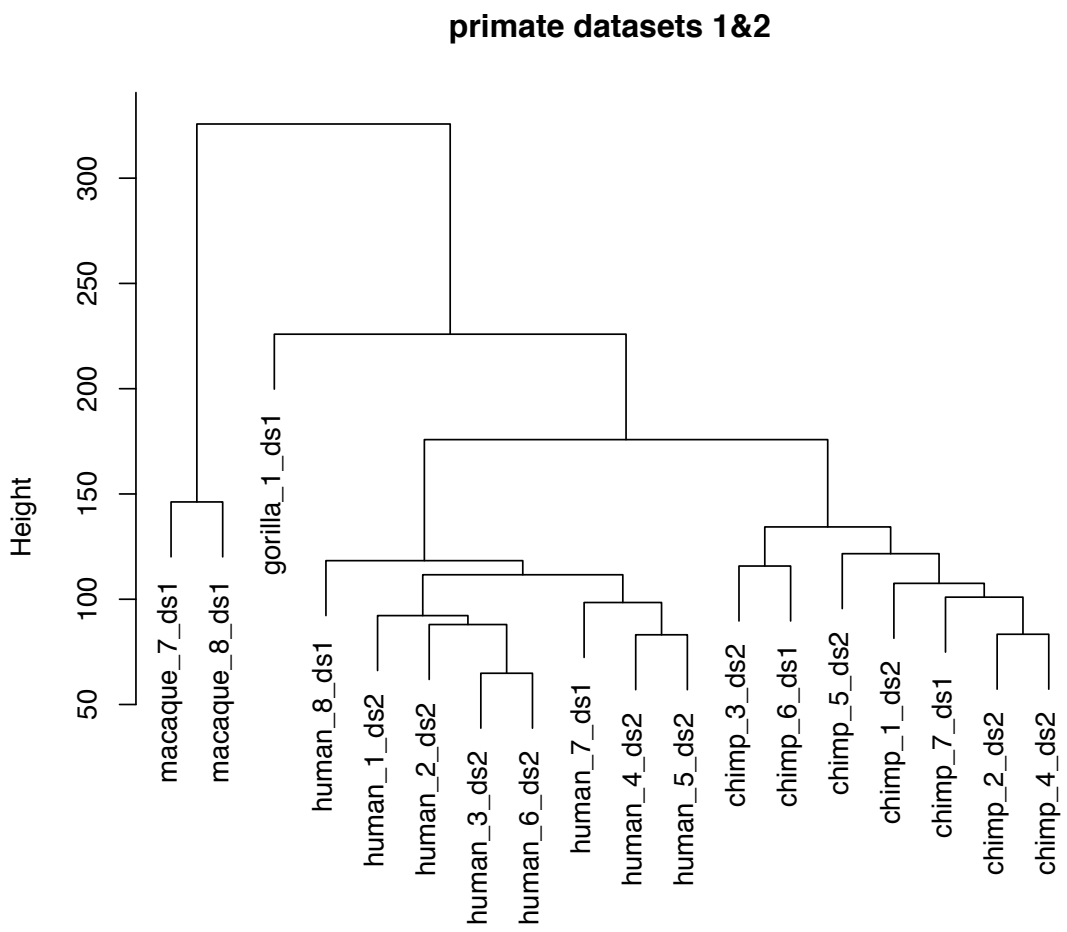

Figure S2:

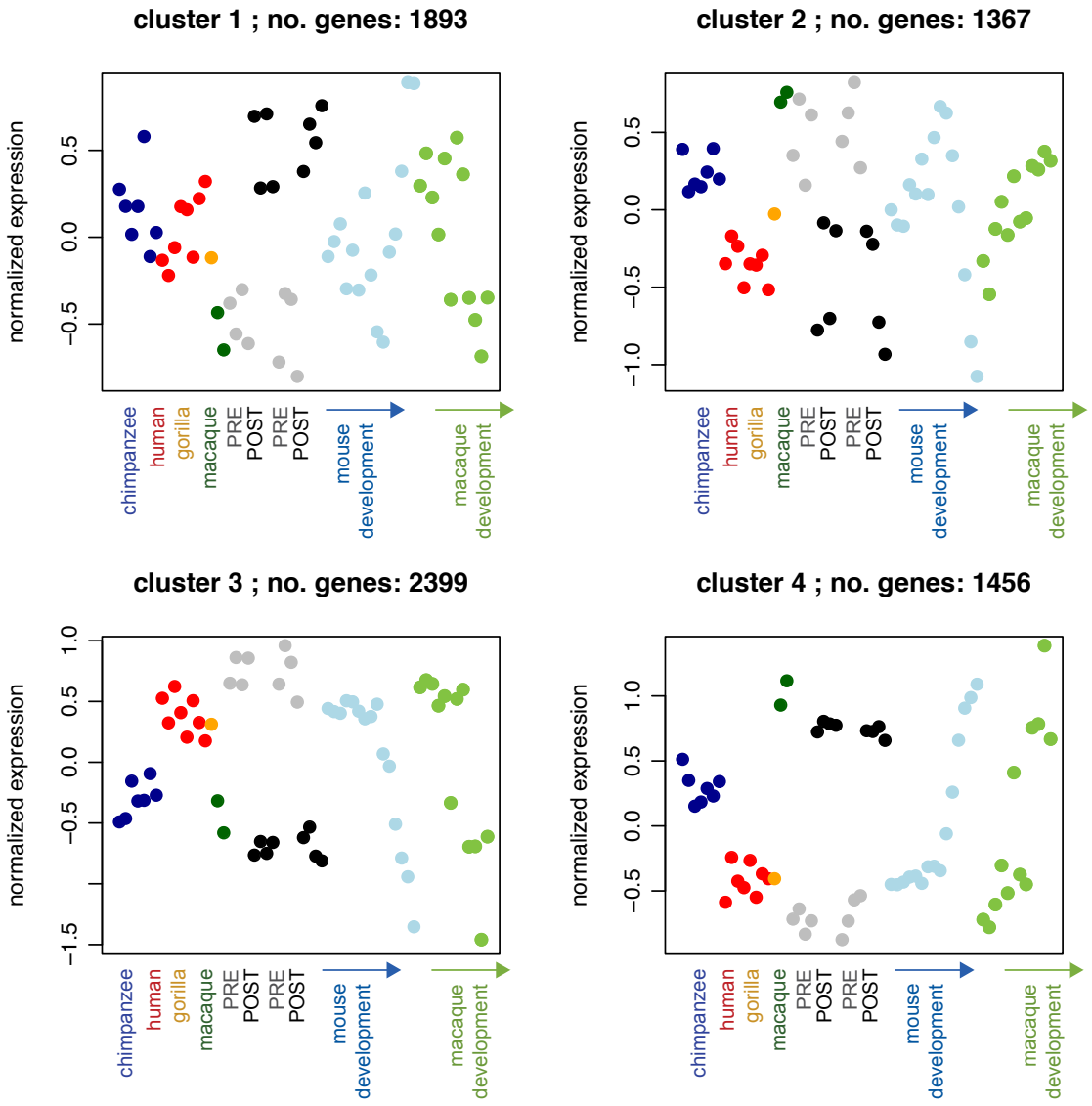

Figure S3: