Sexual selection and pseudogenization in primate fertilization

Jolie A. Carlisle[1], Andrew L. Bickell[2], Michael W. Hart[2,3]

[1]Cornell University, Department of Molecular Biology and Genetics

[2]Simon Fraser University, Department of Biological Sciences

[3]Corresponding author: Department of Biological Sciences, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A1S6, Canada. Email: mwhart@sfu.ca

Short running title: Gene loss under sexual selection

Keywords: Neofunctionalization; sexual conflict; mating system; speciation; infertility

Abstract

The mouse sperm protein ZP3R interacts with proteins in the egg coat and mediates sperm–egg adhesion in a species-specific manner. Understanding the function and evolution of such genes has important implications for studies of speciation, reproductive success, and infertility. A recent analysis showed that (1) the human pseudogene *C4BPAP1* is the ortholog of *Zp3r*, (2) *ZP3R* pseudogenization evolved independently in parallel among several primate lineages, and (3) functional *ZP3R* genes evolve under positive selection among other primate species. The causes of this pseudogenization of *ZP3R* are unknown. To explore one plausible cause (changes in sexual selection on males), we searched for *ZP3R* pseudogenes in recently published genomes, then compared sexually selected male traits among lineages with and without a functional *ZP3R*. We found that traits associated with more intense sexual selection on males (large male body size, larger male canines, larger testes) were associated with functional *ZP3R* expression, and suggest that a relaxation of sexual selection may be associated with selection for *ZP3R* pseudogenization. This proposed causal relationship implies an evolutionary cost to maintaining redundancy in the suite of primate fertilization genes.

Lay summary

In sexual interactions more is often assumed to be better. But the evolution of animal genomes suggests that sometimes less is more: the adaptive loss of genes that function in sex may be favored by selection. How could this happen? One surprising answer comes from analyzing humans and some other primate species that have turned off a key gene called *ZP3R* that helps sperm bind to eggs. The loss of that gene function in some primates is associated with male traits (smaller bodies, smaller canines, smaller testes) that often indicate less vigorous selection on males to compete for matings with females. That correlation implies that the same selection acting on male morphological traits may also act on sperm molecular traits. The correlation also implies that it's expensive to keep some genes turned on, and that when they're no longer helpful it's adaptive to turn them off or allow them to become fallow. This economy of gene expression in sex is an under-explored area of research.

Introduction

Successful fertilization is critical for all sexually reproducing organisms. Interactions between sperm and egg in animals are mediated in part by a suite of proteins that are expressed on the surfaces of gametes and mediate several types of cellular interactions at the time of fertilization, ranging from sperm chemoattraction to gamete fusion (Hirohashi et al., 2008). The genes that encode these gamete surface proteins are often under strong selection (Metz & Palumbi, 1996; Vacquier, 1998; Swanson & Vacquier, 2002). Understanding how selection shapes the evolution of gamete interaction can potentially illuminate the phenotypic and genetic causes of fertility variation and reproductive compatibility between mates, and the evolution of reproductive isolation

between populations and species. Understanding the molecular basis for sperm–egg interactions in mammals is of particular interest for the development of both contraceptives (Meeusen et al., 2007) and infertility treatments (Huang et al., 2014).

Although many important molecules have been described for their role in mammalian fertilization (e.g., sperm–egg plasma membrane fusion; Inoue et al., 2013; Allingham & Floriano, 2021; Brukman et al., 2022), the catalogue of genes and proteins that mediate fertilization can be incomplete (Carlisle & Swanson, 2021). That gene repertoire evolves in part by whole-gene duplication, the addition of new paralogs into related fertilization functions, functional redundancy of some genes (Ikawa et al., 2010; Avella et al., 2013; Lu et al., 2019), and the loss of functional members of such gene families (recently reviewed by Carlisle et al., 2022; Rivera & Swanson, 2022). For example, the mouse egg coat proteins ZP2 and ZP3 interact with mouse *Zp3r* (or *Zona pellucida 3 receptor*, also known as *Sp56*) expressed in the sperm acrosome (Buffone et al., 2008; Wassarman, 2009). However, male *Zp3r*- / *Zp3r*- null homozygotes are fertile (Muro et al., 2012), probably because *Zp3r* is functionally redundant with other male-expressed genes and proteins that bind eggs. Functional redundancies in fertilization may evolve as part of the arms race between males and females driven by sex differences in optimal values for reproductive traits including the specificity or strength of sperm–egg binding interactions (Gavrilets, 2000; Turner et al., 2008).

The human ortholog of mouse *Zp3* has long been well known (Buffone et al., 2008), but until recently the identity of its human ortholog was uncertain. The RCA gene cluster on mouse chromosome 1 (chr1) includes *Zp3r* and several other members of the C4b-binding protein family of genes and pseudogenes (McLure et al., 2004) that encode regulators of complement activation in the innate immune system (Arenzana et al., 1995). Mouse *Zp3r* is descended from mouse *C4bp* by an ancient gene duplication event, and *Zp3r* has acquired a novel pattern of expression and protein localization (in the testis and acrosome) and a novel function (in sperm binding to the egg coat).

The human RCA cluster includes only one functional gene (*C4BPA*) and a non-functional pseudogene (*C4BPAP1*). Several previous population genetic studies assumed that the functional gene *C4BPA* is the human ortholog of *Zp3r* (Rohlfs et al., 2010; Hart et al., 2018; Morgan & Hart, 2019). Carlisle et al. (2024) analyzed synteny between genes in the mouse and human RCA clusters, and phylogenetic relationships among the members of the C4b binding protein gene family, to show that *C4BPAP1* is the human ortholog of mouse *Zp3r*. Expression of that human *ZP3R* pseudogene mRNA in the testis (Carithers & Moore, 2015) is consistent with its former function in fertilization.

Carlisle et al. (2024) also identified *ZP3R* pseudogenization events that occurred independently across some but not all other primate lineages. The parallel loss of *ZP3R* function was unambiguously inferred from differences in the genomic locations and types of frameshift mutations and premature stop codons among *ZP3R* coding sequences from eight other lineages. Given the evidence for positive selection acting on divergence among functional primate *ZP3R* genes (Carlisle et al., 2024), and the

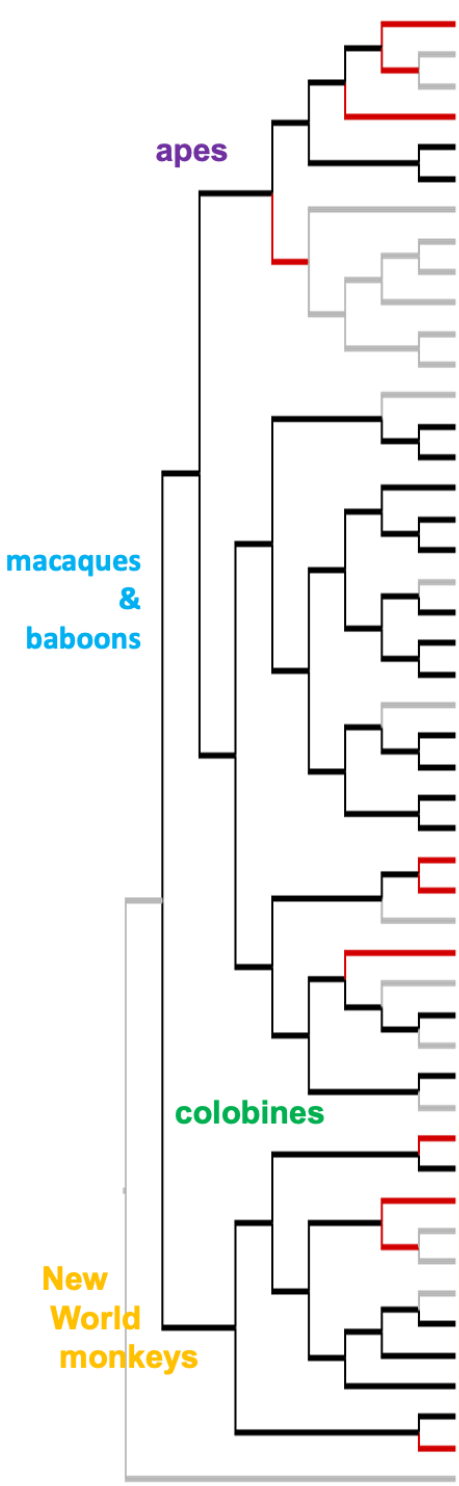
general tendency for sexual selection to cause rapid divergence of genes that mediate reproductive success (Gavrilets, 2000; Uy & Borgia, 2000; Van Doorn et al., 2001; Swanson & Vacquier, 2002; Patiño et al., 2016), Carlisle et al. (2024) concluded that "recurrent, lineage-specific gene loss events [are] suggestive of strong selection for gene loss [but] it is still unclear what is driving the loss of *ZP3r* in primates."

Here we explore evidence for one possible mechanism driving the evolutionary loss of *ZP3R* expression: changes in the strength of sexual selection on males that permits or favours the loss of gene function and reduced redundancy in sperm adhesion to the egg coat. We added to the data from Carlisle et al. (2024) by finding functional *ZP3R* coding sequences or pseudogenes in recently published genomes of 17 additional primate species. We used previously published morphological data to characterize variation among primate lineages in three sexually-selected traits, and we show that smaller males with smaller canines (but not smaller testes) are associated with *ZP3R* pseudogenization events. One interpretation of that pattern across primate lineages is that phenotypic markers of the intensity of precopulatory sexual selection predict the conditions for loss (and perhaps for gain) of gene function in fertilization.

Methods

We used data for *ZP3R* pseudogenization from Carlisle et al. (2024), including their discoveries of functional or non-functional *ZP3R* gene structure in 30 non-human primate species (Figure 1). We used the same methods (Appendix S1) to find functional *ZP3R* genes or pseudogenes in the genomes of 17 additional species from Shao et al. (2023).

We analyzed the coevolution of *ZP3R* with phenotypic traits that are expected to vary with (and reflect past selection arising from) the intensity of sexual selection including male-male competition for mates or for fertilizations. We compiled data from the literature on female body mass (g), male body mass, female canine size, male canine size, and testis mass (g) from 48 species (Appendix S2). We used mesiodistal length (mm) of the maxillary (upper) canine (Thorén et al., 2006). The data include 12 species of apes, 15 species of macaques and baboons (Cercopithecinae), 9 species of colobines, 11 species of New World monkeys, and 1 tarsier (Figure 1). We interpreted the results as evidence for a correlated response to sexual selection on male morphological traits and on male molecular traits.



Species		Body mass (g)		Maxillary canine mesiodistal length (mm)		Testis mass (g)
		Female	Male	Female	Male	
Homo sapiens		40100 [g]	47900	8.24 [q]	9.29	40.50 [z]
Pan paniscus		33200	43000	9.37	11.44	167.70 [aa]
Pan troglodytes		35200	42000	11.94	15.30	118.80
Gorilla gorilla		97700	159200	15.09	23.67	29.60
Pongo abelii		38300 [h]	62140	13.37 [r]	19.50	34.80
Pongo pygmaeus	*	38700	86300	13.33	17.50	35.50 [ab]
Hoolock leuconedys [a]	*	6880	6870	7.69	8.42	
Hylobates moloch		6250 [i]	6580	7.60	7.40	6.10
Hylobates pileatus	*	5440	5500	7.62 [r]	7.42	
Symphalangus syndactylus	*	10700	11880	8.45	9.72	11.95 [n]
Nomascus leucogenys		7620 [i]	7790	7.69 [s]	8.14	5.50 [ab]
Nomascus siki	*	5800 [j]	5600			
Cercopithecus albogularis	*	4150	6900	5.22 [t]	8.29	
Chlorocebus aethiops	*	3560	5080	5.17 [u]	7.37	13.00
Erythrocebus patas	*	6500	12400	6.88	10.58	7.20 [n]
Cercocebus atys [b]		6100 [k]	8680	6.46	10.35	25.10
Mandrillus leucophaeus		10000	17000	8.60	17.89	41.05 [ac]
Mandrillus sphinx		11500	25000	8.54 [r]	19.40	68.00 [ac]
Lophocebus aterrimus	*	5760 [i]	7840	5.92	8.25	
Theropithecus gelada		13950	26100	7.88	14.89	17.10
Papio anubis		14700	24000	8.83	16.10	93.50
Papio hamadryas	*	9400	21500	7.85	14.14	27.10 [n]
Macaca assamensis	*	6700 [l]	11300			
Macaca nemestrina		7800	10400	7.42	12.29	66.70
Macaca silenus	*	5000	6800	6.25 [r]	11.04	42.00 [n]
Macaca fascicularis		3500	5500	6.21	9.73	35.20
Macaca mulatta		3000	6200	6.10	8.67	46.20
Colobus angolensis		8200	10200	7.05	9.30	10.70 [ab]
Colobus guereza	*	8700	10600	8.19	9.84	2.98 [ac]
Piliocolobus tephrosceles						
Pygathrix nigripes [c]	*	8440 [i]	11000	5.82 [r]	8.04	8.40 [n]
Rhinopithecus bieti						
Rhinopithecus roxellana		11600 [m]	17900	6.31 [v]	8.80	35.60 [ab]
Rhinopithecus strykeri	*					
Trachypithecus francoisi		7350 [n]	7700	6.40 [r]	8.20	3.98 [n]
Trachypithecus phayrei	*	6300 [i]	7870			
Alouatta palliata		5350	7150	6.46	8.16	23.00
Ateles geoffroyi		7460	8210	5.71	6.38	13.40
Aotus nancymaae		957 [o]	998	3.13 [w]	3.19	0.50 [ab]
Callithrix jacchus		240	260	2.11 [x]	2.06	1.30
Saguinus midas	*	430	590	2.84	2.80	1.83 [ac]
Cebus albifrons		1810	2480	4.79 [y]	5.91	
Cebus imitator [d]		2670	3870	5.28 [y]	6.42	9.10 [ab]
Sapajus apella [e]		2390	3050	6.32	7.49	4.64 [ac]
Saimiri boliviensis		700	1020	2.88 [r]	3.54	3.20 [ab]
Pithecia pithecia		1520	1730	4.26	5.00	0.92 [ac]
Plecturocebus donacophilus [f]		887 [p]	1017	2.89	2.84	0.54 [ad]
Tarsius syrichta		120	130			0.70 [ab]

124

Cassini (2020) reviewed several life history traits that are also expected to covary with the intensity of sexual selection, including mating systems that range from social groups with monogamous male-female pairs (and limited scope for competition between males or among sperm) to social groups with multimale mating by females (and greater scope for sexual selection arising from competition for fertilizations) (see Lindenfors & Tullberg, 1998; Adret et al., 2018; Spence-Aizenberg et al., 2018). Because sexual size dimorphism is especially pronounced in primates relative to other mammals (Plavcan, 2001; Tombak et al., 2024), and traits that are closely associated with male competition for matings or for fertilizations are expected to covary with the molecular evolution of genes that encode fertilization molecules (Prothmann et al., 2012; Ports & Jensen-Seaman, 2024), we focused on those five morphological variables in our analyses of *ZP3R* pseudogenization.

For 36 species we found complete morphological data for all five variables (9 apes, 12 macaques and baboons, 5 colobines and 10 New World monkeys). For those species we summarized covariation among the five morphological variables in a principal components analysis using covariances and standardized values (Appendix S3).

Figure 1. *ZP3R* functionality and morphological variables from 31 primate genomes analyzed in Carlisle et al. (2024) and 17 other primate genomes analyzed in this study (marked by asterisks). Cladogram from 10kTrees (Arnold et al., 2010). New *ZP3R* data come from genomes in Shao et al. (2023). Red edges in the cladogram show 11 parallel origins of *ZP3R* pseudogenes in humans, *Pan*, *Gorilla gorilla*, gibbons, *Colobus angolensis* and *C. guereza*, *Pygathrix nigripes*, *Alouatta palliata*, *Aotus nancymaae*, the most recent common ancestor of *Callithrix jacchus* and *Saguinus midas,* and in *Plecturocebus donacophilus*. Twelve species names in plain type show species for which some morphological data were missing; those species were not included in the principal components analysis. Grey edges in the cladogram show species or lineages that were not included in the comparative analysis (calculation of relative size as residuals from lineage regressions, BayesTraits estimate of ancestral states for relative sizes, or one-sample tests) because those species were descended from the same single pseudogenization event or because data were missing. Superscript letters indicate footnotes that explain species names or give data sources for five morphological variables that were used in quantitative analyses: a, *Bunopithecus hoolock* in 10kTrees; b, *C. torquatus atys* in 10kTrees; c, *P. nemaeus* in 10kTrees (two species were recent split, probably hybridize); d, *C. capucinus* in 10kTrees; e, *Cebus apella* in 10kTrees; f, *Callicebus donacophilus* in 10kTrees; g, body masses from Lindenfors & Tullberg (1998) unless otherwise noted; h, Markham & Groves (1990) for "Sumatran" individuals; i, Thorén et al. (2006); j, https://primate.wisc.edu/primate-info-net/pin-factsheets/pin-factsheet-white-cheeked-gibbon; k, Kennard & Willner (1941); l, Gordon (2006); m, Smith & Jungers (1997); n, Lupold et al. (2019); o, Spence-Aizenberg et al. (2018); p, Gron (2007); q, canine lengths from Thorén et al. (2006) unless otherwise noted; r, Plavcan & Ruff (2008), similar to M. Skinner (unpublished data); s, as *Hylobates concolor*; t, as *C. mitis*;u, as *Cercopithecus aethiops*; v, Pan & Oxnard (2001); w, to include this lineage in the analysis we used canine sizes for *A. trivirgatus*, which have adult body sizes similar to *A. nancymaae*; x, Leutenegger & Cheverud (1982); y, Masterson (2003); z, testis masses from Harcourt et al. (1981) unless otherwise noted; aa, Dixson (2012); ab, Ports & Jensen-Seaman (2024); ac, Harcourt et al. (1995); ad, to include this lineage in the analysis we used testis sizes for *P. cupreus* from Conley et al. (2022), in which adult body sizes and life histories are similar to *P. donacophilus*

We used a regression method like that of Ports & Jensen-Seaman (2024) to characterize relative male size for each of those 36 species as the residual from a linear regression against female size (male body mass = 1.61 × female body mass – 958). We estimated relative male canine length as the residual from a linear regression against female canine length (male canine length = 1.58 × female canine length – 1.11). And we estimated relative testis mass as the residual from a linear regression against male size (log[testis mass] = 0.82 × log[male body mass] – 2.11).

We then used those three traits in a comparative analysis. We found independent origins of different pseudogenizing mutations in 11 lineages (red in Figure 1; see Carlisle et al., 2024), including *Alouatta palliata*, *Aotus nancymaae*, *Colobus angolensis* and *C. guereza* (with four different unshared mutations), *Gorilla gorilla*, *Homo sapiens*, *Pygathrix nigripes*, and *Plecturocebus donacophilus* (Figure 1). We also found shared pseudogenizing mutations in the marmoset *Callithrix jacchus* and the tamarin *Saguinus midas*; in the two species of *Pan*; and in all six species of gibbons. Each of those represent a single origin of a *ZP3R* pseudogene in the most recent common ancestor of each of those species groups.

We used the phyloglm function in the R package phylolm 2.6.5 (Ho & Ane, 2014) to fit the phylogenetic logistic regression of Ives & Garland (2010) for a binary response variable (a pseudogene or a functional *ZP3R*) and a continuous predictor variable (relative male mass, relative male canine length, or relative testis mass) with 1000 bootstrap replicates. We used trait data and the phylogram with branch lengths for all 36 species from 10kTrees (the same topology shown in Figure 1; Appendix S4). We used the logistic maximum likelihood estimator in phyloglm (for canine size and testis size) or the generalized estimating equations for a Poisson regression with penalized likelihood (for male size, where the MLE method failed to converge on a solution).

A functional *ZP3R* is the ancestral state shared among primates and rodents, and pseudogenization appears to be irreversible (see Carlisle et al., 2024). Thus trait values on internal edges in the phylogeny can be associated with known *ZP3R* states, including a pseudogene that evolved in the common ancestor of *Pan* species, a pseudogene that evolved in the common ancestor of *Callithrix* and *Saguinus*, a pseudogene that evolved in the common ancestor of the gibbons (shown in red in Figure 1), and a functional gene for all other internal edges (shown in black in Figure 1). This makes the phylogenetic logistic regressions problematic because they model evolution of the dependent and independent variables along all edges of the phylogeny, including edges on which no evolution of *ZP3R* is expected and evolution of the morphological traits is not informative (e.g., the 10 edges descended from the MRCA of gibbons, which all have the irreversible derived pseudogene).

To complement those results, we added a second simpler test. We used BayesTraits v4.1.1 (Barker et al., 2007) to infer the ancestral value of the trait on internal branches or edges of the phylogeny under a random walk model using the MCMC method (Appendix S5); similar ancestral state estimates are one analytical step in the

phylogenetic logistic regression analysis as well. We used the same phylogram from the phylogenetic logistic regressions.

We then used a one-sample *t* test to compare the mean and variance for each trait in the sample of edges (3) and species (8) that independently evolved a pseudogene (shown in red in Figure 1) to a test value based on the phylogeny-wide expectation for that trait (the average for all 51 other internal edges and species with a functional *ZP3R*, shown in black in Figure 1).

The one-sample tests avoid modeling trait evolution along edges with irreversible pseudogenes, including the 10 internal and terminal edges among the gibbons that all descended from a single origin of a pseudogene in their common ancestor (Figure 1), the two terminal edges leading to *Pan paniscus* and *P. troglodytes*, and the two terminal edges leading to *C. jacchus* and *S. midas*. However, the one-sample tests do not account for autocorrelation among trait values along the internal edges used to calculate the phylogeny-wide expectation.

Results

Most pseudogenes were found in species with small males, small canines, and small testes (Figure 2). Both females and males were larger in species with a pseudogene (mean 15.6 kg, 21.0 kg, respectively) than in species with a functional *ZP3R* (8.5 kg, 14.7 kg), mainly because our sample of species (12) and observations of *ZP3R* pseudogenes (10) included many apes with relatively large body sizes (Figure 1). However, when plotted against female size, males were relatively larger among species with a functional *ZP3R* (Figure 2), and this difference was significant in other analyses (see below).

The pattern of smaller size associated with pseudogenes was strongest in the data for canine length (Figure 2): males of species with a pseudogene had smaller canines (mean 8.6 mm) than males of species with a functional *ZP3R* (10.7 mm); and the difference was reversed in females (7.2 vs. 6.9 mm). The difference in relative male canine length was partly caused by the large male canines found in most macaques, baboons, and other cercopithecines, in which no species was found to have evolved a pseudogene (Figure 1). Notably, that group includes many species with promiscuous or multi-male mating systems and high expected competition among males for access to females or for fertilization of eggs (Lindenfors & Tullberg, 1998; Cassini, 2020).

The association between pseudogenes and sexually selected traits was weakest in data for testis size. Males of species with (28.6 g) and without (28.4 g) a pseudogene had nearly identical mean testis mass. This similarity was caused in part by the wide variation in absolute and relative testis mass among apes that have (*Pongo*) or lack (*Homo*, *Gorilla,* gibbons) a functional *ZP3R* and in particular by the very large testes for both species of chimpanzees (Figure 1).

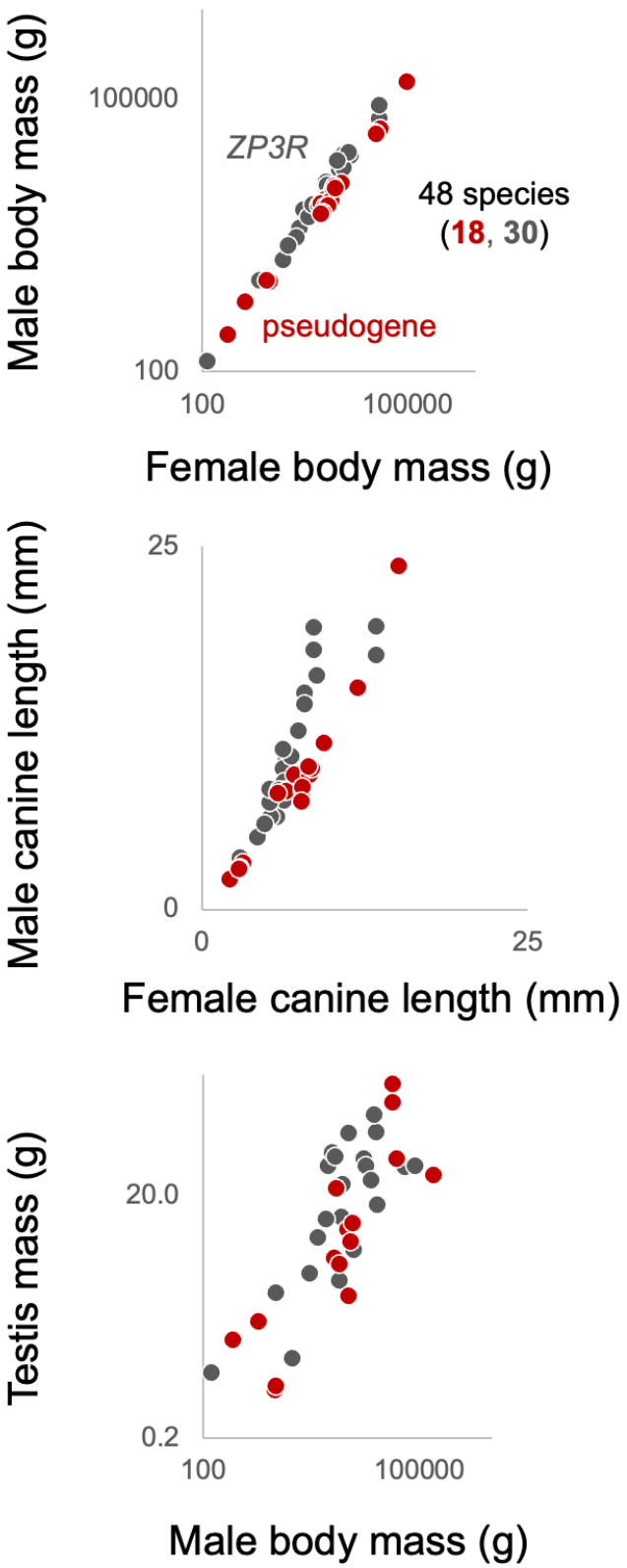


Figure 2. Female and male body masses (on logarithmic axes), canine lengths (on linear axes), and male body mass and testis mass (on logarithmic axes) for all 30 primate species with a functional *ZP3R* (including some species with some missing data) and all 18 species with a pseudogene (including some such as the gibbons species that share a single pseudogenization event from a common ancestor) (see Figure 1)

220

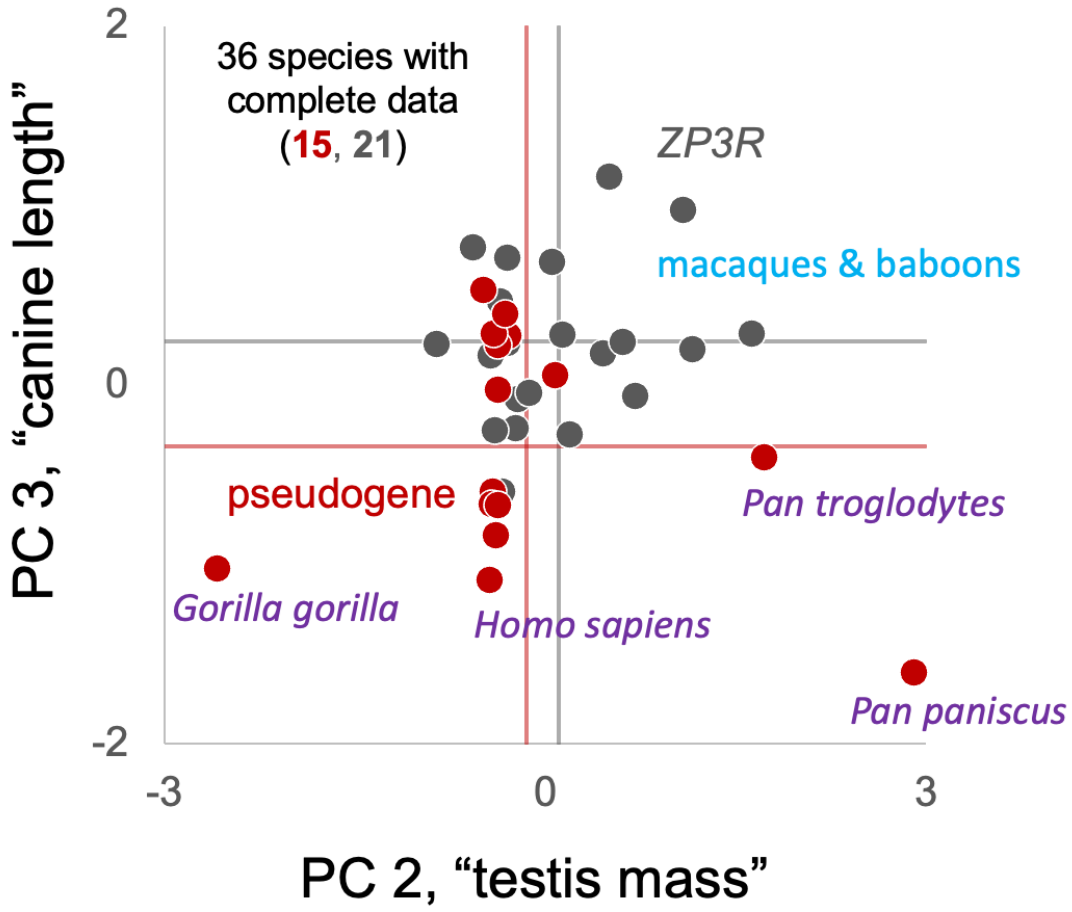
The principal components analysis helped to reduce the dimensionality of that variation to three simple vectors (Table 1): a compound size variable for which all five variables had positive coefficients of similar magnitude (PC1; 74% of variance); a contrast between testis mass versus male and female body mass (PC2; 16%); and a contrast between male and female canine length versus body masses and testis mass (PC3; 7%). In the space defined by PC2 and PC3, species with a functional *ZP3R* tended to have higher or positive scores for PC2 (large testes, small male and female body sizes) and PC3 (large female and male canines, small body and testis sizes) (Figure 3). As in the broader data set including all species (Figure 2), the difference was larger for PC3 (canine size relative to body size) than for PC2 (testis size) in part due to the wide variation in testis mass among apes.

Table 1. Principal components analysis of five morphological variables for 36 primate species. Bold values for some loadings in two eigenvectors show the variable(s) that contribute(s) most to positive scores for the two vectors that define a space in which sexually selected traits vary among species with and without a *ZP3R* pseudogene

ALT TEXT: Results from a clustering analysis of five morphological variables from 36 primate species. Three eigenvectors explain most of the variation, including a contrast between testis size and body size, and a contrast canine size with body and testis size.

	PC_1	PC_2	PC_3	PC_4	PC_5
Eigenvalue	3.72	0.810	0.359	0.0868	0.0144
% of variance	74.5	16.2	7.19	1.73	0.288
Eigenvectors					
Female canine length	0.491	-0.0181	**0.343**	0.794	-0.0902
Male canine length	0.465	0.0896	**0.664**	-0.555	0.157
Female body mass	0.477	**-0.273**	**-0.475**	-0.0183	0.685
Male body mass	0.479	**-0.334**	**-0.327**	-0.242	-0.701
Testis mass	0.288	**0.897**	**-0.326**	-0.0246	-0.0704

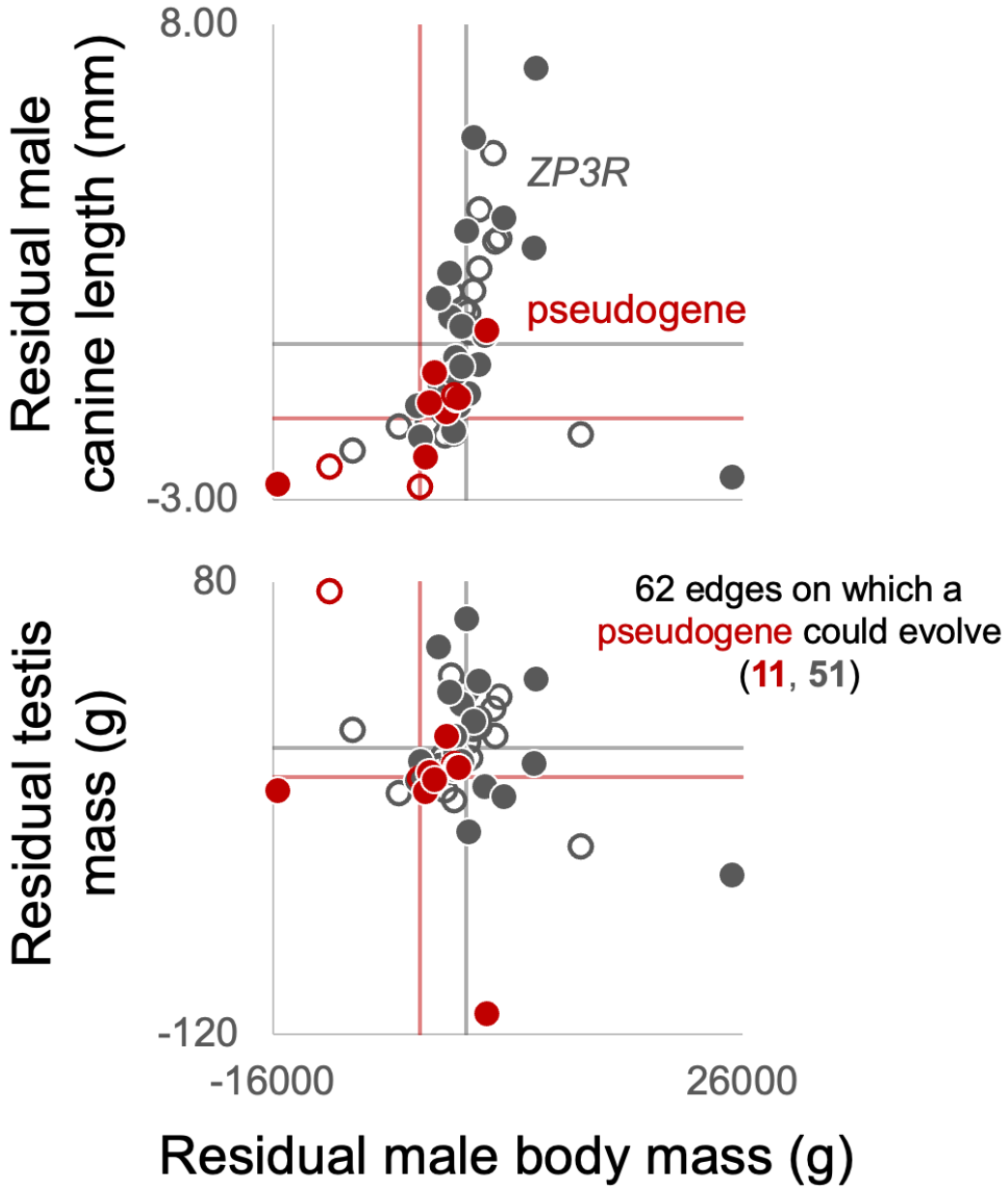
The evolution of small male size was a good predictor of the evolution of a pseudogene. The phylogenetic logistic regression coefficient for relative male mass was negative (b = -4.3e-4), and the standardized effect was large (z = -2.17) and significant (p = .029). That result was consistent with the one-sample test. Among 11 lineages with a *ZP3R* pseudogene, mean relative male mass (range) was -2.9 kg (-15.6 to 3.1 kg) (i.e., males were smaller than expected) (Figure 4). By contrast, the phylogeny-wide expected relative male mass (range) was 1.3 kg. That difference was in the same direction as the phylogenetic logistic regression and was also significant (t = -2.52, df = 10, p = .03).



248

Figure 3. Scores on the second and third eigenvectors of a principal components analysis of five morphological variables from 21 primate species with a functional *ZP3R* (grey) and 15 species with a pseudogene (red); the analysis included only species for which no morphological data were missing. See Table 1 for the coefficients of each variable in each vector. Vertical and horizontal lines show mean scores for the two species groups

249 Among the three variables, small male canines were the strongest predictor of the
250 evolution of a pseudogene. The phylogenetic logistic regression coefficient was
251 negative and the effect of relative male canine length on pseudogenization was
252 significant (b = -0.737, z = -2.37, p = .017). In 11 lineages with a *ZP3R* pseudogene,
253 mean relative male canine length was -1.1 mm (-2.7 to 0.92 mm) (i.e., male canines
254 were smaller than expected) (Figure 4). By contrast, the phylogeny-wide expected
255 canine size was 0.60 mm (Figure 4). In the one-sample test, that large difference in
256 canine size (-1.71 mm) between lineages with a pseudogene and the phylogeny-wide
257 expected value was highly significant (t = -4.97, df = 10, p < .001).



258

Figure 4. Relative male body mass, relative male canine length, and relative testis mass estimated as residuals from linear regressions. Each symbol shows values for a terminal edge of the phylogeny leading to one of 29 primate species (closed symbols) or values inferred using BayesTraits for one of 33 internal edges of the phylogeny (open symbols). Vertical and horizontal lines show mean values for the 11 total edges on which a pseudogene evolved (red) or the phylogeny-wide expected value on all other edges with a functional *ZP3R* (the known ancestral state). Species and edges included in the analysis are shown in Figure 1

Those two patterns among measures of relative male body mass and relative male canine length were also consistent across the four major primate lineages in our dataset. A *ZP3R* pseudogene evolved 3 or 4 times each among apes, colobines, and New World monkeys with relatively small males and small canines (Figure 5). Those patterns were consistent despite the broad range of male sizes among apes and the range of canine sizes among macaques and baboons. Among the apes, pseudogenes arose mainly in lineages with small males and small canines (humans, chimps, gibbons). No pseudogene evolved among the macaques and baboons, which all have large males with large canines.

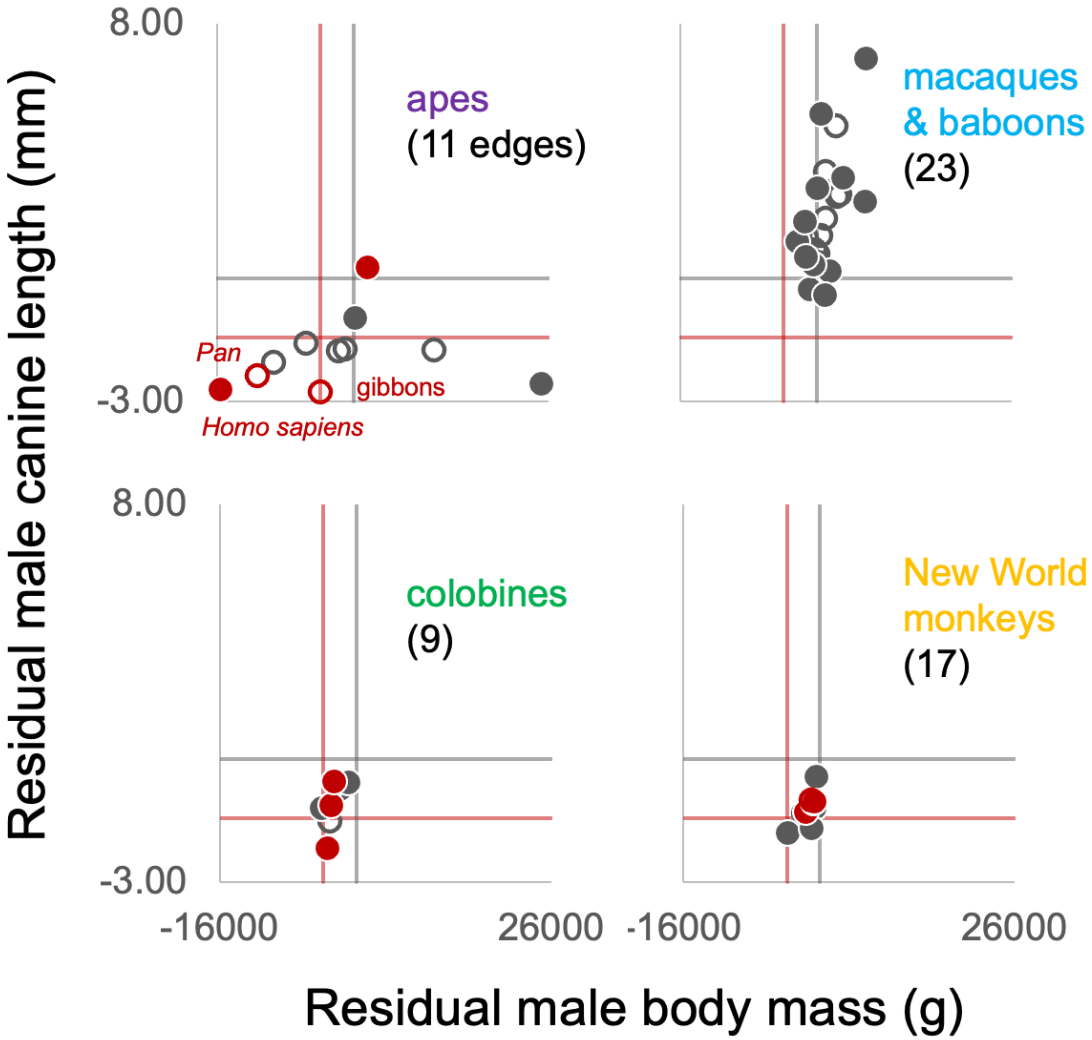


Figure 5. Relative male body mass and relative male canine length from Figure 4 disambiguated by the four major primate lineages (see Figure 1). Each symbol shows values for a terminal edge on the phylogeny leading to one of 29 primate species (closed symbols) or values inferred using BayesTraits for one of 33 internal edges of the phylogeny (open symbols). Vertical and horizontal lines show mean values for the 11 total edges on which a pseudogene evolved (red) or the phylogeny-wide expected value on all other edges with a functional *ZP3R* (the known ancestral state). Species and edges included in the analysis are shown in Figure 1.

Surprisingly, testis size did not predict pseudogenization. Mean relative testis mass was similar among lineages that did (-6.22 g) and did not (6.70 g) evolve a pseudogene, relative to the broad scale of relative testis sizes among the species in our sample (range -110 g to 75 g), and there was no obvious relationship between relative testis mass and relative male body mass (Figure 4). The phylogenetic logistic regression coefficient was small, not significantly different from zero (b = -8.66e-5, z = 0.0162, p = .98), and similar to the results from the one-sample test (t = -1.01, df = 10, p = .34).

Discussion

We found that a *ZP3R* pseudogene tends to evolve in lineages that have evolved smaller values for sexually-selected male traits. A reasonable interpretation of those patterns is that a change in the strength of sexual selection on male competition for access to females (before copulation) or access to eggs (after copulation) allows for the evolution of relatively smaller male body and canine sizes, and that such circumstances also permit or favour the loss of gene function in systems with redundancy in sperm adhesion to the egg coat. We conclude that loss of *ZP3R* function may also be a consequence of relaxed sexual selection on males. The patterns are similar to the insertion of regulatory regions that enhance sex-specific gene expression among some of the same primates in response to changes in sexual selection inferred from analysis of the same three morphological traits by Anderson & Jones (2024).

Carlisle et al. (2024) showed that *ZP3R* evolves under positive selection (a high relative rate of nonsynonymous substitutions) among primate species with a functional gene. They argued that the parallel origins of pseudogenization, the divergence of functional genes under positive selection, and fixation of the human pseudogene mutation all argue against neutral processes (such as relaxed selection in humans and other lineages) and favour some form of strong selection as the cause of pseudogenization. Our results are consistent with that argument and suggest that sexual selection causes both the loss of *ZP3R* function (in species that are experiencing weaker sexual selection or have resolved sexual conflicts) and amino acid sequence divergence among functional *ZP3R* genes (in species that are experiencing strong sexual selection on males or sexual conflict between males and females; Gavrilets et al., 2000).

We found consistent quantitative associations between male body size or canine size and *ZP3R* pseudogenization, but less evidence linking relative testis size and *ZP3R* evolution. Those mixed results suggest that some other traits might account for among-lineage variation in the selective conditions that favour loss of *ZP3R* function in fertilization. One good candidate is behavioural interactions that result in monogamous male-female pairs or that permit single males to monopolize matings with several females (Harcourt et al., 1995; Lindenfors & Tullberg, 1998; Cassini, 2020). Some of our observations are consistent with a role for such traits in accounting for variation among some primate lineages with different mating systems and potential for sperm competition, sexual selection among males, and sexual conflicts of interest between mates. Future studies could also extend the comparison of *ZP3R* structure and life history variation to the lemurs and lorises, in which complex patterns of variation in life

history and mating system traits (Kappeler, 1997; Kappeler et al., 2017) are expected to parallel those found among apes, monkeys, and tarsiers and predict the evolution of *ZP3R* structure and function. Unfortunately, we were not able to find and unambiguously annotate *ZP3R* in available genomes from lemurs and lorises (Shao et al., 2023).

We emphasize that these phenotypic traits associated with behavioural interactions between males or between mates may predict *ZP3R* pseudogenization but the causal mechanism of *ZP3R* pseudogenization is more likely to involve the number, durability, and timing of molecular interactions between sperm and eggs at fertilization. These molecular interactions are cryptic in primates but can be observed and quantified in aquatic animals with planktonic gametes and fertilization (Evans & Lymbery, 2020). Changes in the ecology of gamete interactions in the ocean cause the molecular evolution of sperm–egg binding genes (e.g., Levitan, 2012; Patiño et al., 2016). Comparable mechanisms based on the ecology of sperm–egg interactions (e.g., Weber & Fisher, 2023) are likely to cause the evolutionary loss of gene expression such as the repeated parallel pseudogenization of primate *ZP3R*. We know of few previous studies (e.g., Goudet et al., 2008) that have discovered specific genes lost from a fertilization repertoire (as in Carlisle et al., 2024) and then tried to identify features of the selective environment that are associated with loss of gene function and expression (as in this study). But incidental discoveries of loss of gene expression in comparative transcriptomic studies of species with large qualitative differences in life history traits (e.g., species with external versus internal fertilization; Guerra et al., 2021) suggest that loss of function may also be common in comparisons among species with quantitative differences in mating systems (e.g., primates with stronger or weaker precopulatory competition for mates; Carlisle et al., 2022; Rivera & Swanson, 2022).

Carlisle et al. (2024) showed that *ZP3R* originated from a gene duplication of *C4BPA* in a common ancestor that was older than the divergence between Glires and Primates. Genes with other functions can be recruited into gamete binding, possibly as evolutionary solutions to complement or overcome adaptations in the other sex in the context of sexual conflicts of interest (Gavrilets, 2000) over the frequency or intensity of mating or reproductive allocation (Carlisle et al., 2022; Rivera & Swanson, 2022). How old was this gene duplication event? Extending these analyses to the other major lineages of mammals could identify the age and origin of *ZP3R* and increased redundancy in gene function at fertilization.

These future studies could also help to understand the causes of decreased redundancy in gene function when *ZP3R* (and other fertilization genes) is lost by pseudogenization. Our analysis suggests that altered sexual selection is associated with this phenomenon. That proposed relationship implies an evolutionary cost to maintaining redundancy in the suite of primate fertilization genes under conditions of relaxed sexual selection on males or resolution of sexual conflict between mates. The nature of this cost is not known but might include high risk of fatal polyspermy relative to the reduced benefits of high-fidelity sperm binding in species with lower sperm competition among males (Okamoto, 2016; Levitan et al., 2019; Weber & Fisher, 2023;

354 but see Dale, 2024). Under this hypothesis, males in species that have resolved sexual
355 conflict or reduced male competition under less promiscuous mating systems may
356 experience selection for lower sperm binding affinity to the egg coat and reduced
357 likelihood that two sperm from the same male fertilize the same egg (resulting in the
358 deaths of all three genome copies).

359 Removing some male-expressed loci from the egg-binding repertoire of sperm could be
360 one response to such selection, similar to changes in allele frequencies for single loci in
361 response to changes in the male optimum for rates of sperm–egg contact (Levitan,
362 2012). That hypothesis also predicts that, in species that have not evolved this kind of
363 resolution of sexual conflict, females may also experience selection for lowering sperm
364 binding affinity to the egg coat by adding new female-expressed genes to the egg coat
365 repertoire in order to fend off supernumerary sperm and avoid polyspermy (e.g.,
366 Aagaard et al., 2013).

367 Reducing the size of testes and the number of sperm per ejaculate seems like another
368 obvious potential response to selection on males to avoid polyspermy, and might be
369 associated with genomic changes that reduce gene expression that mediates sperm–
370 egg binding. Anderson & Jones (2024) also found phylogenetic correlations between
371 changes in testis size and the number of genomic elements that affect androgen-
372 dependent gene expression. By contrast, our results do not suggest that adaptive
373 change in testis size is often associated with the evolution of a *ZP3R* pseudogene, and
374 we are not able to explain why the evolution of some sexually selected male traits (body
375 size, canine length) but not others (testis size) might be associated with this recurrent
376 gene loss. Lupold et al. (2019) reviewed the evidence for a tradeoff between male
377 investment in weapons, badges of status, or other mechanisms of precopulatory mate
378 competition versus investment in testis size for postcopulatory sperm competition. One
379 speculative hypothesis is that testis size (and number of sperm) may show a similar
380 evolutionary tradeoff with molecular traits of sperm (and the expression of genes that
381 moderate the affinity of each sperm for the egg coat).

382 The causes of decreased redundancy in human fertilization (through the evolution of the
383 human *ZP3R* pseudogene) are of particular interest. The timing of the human
384 pseudogenization event could point to its evolutionary cause. Because the premature
385 stop codon in the second CCP domain of human *ZP3R* is not a polymorphism in dbSNP
386 (Sherry et al., 1999), and because both the divergence and loss of *ZP3R* among
387 primates appears to be driven by selection (Carlisle et al., 2024), the origin and fixation
388 of the human pseudogene is expected to be associated with a change in selection
389 acting on the early human lineage. Among the Neanderthal and Denisovan genomes in
390 the UCSC genome browser (GRCh37/hg19), three Neanderthal sequence reads and 25
391 Denisovan reads map to the same site (chr1:207340753) as the human stop codon in
392 *ZP3R*; all have the human TAA stop mutation. That similarity relative to the depth of
393 coverage suggests that the human pseudogene is shared between archaic and early
394 modern humans, and that the causes of the human pseudogenization event were older
395 than the divergence between them and unrelated to any life history or mating system
396 differences among them.

Author contributions

JAC compiled and analyzed data, designed the figures, and wrote the paper; ALB compiled and analyzed data and wrote the paper; MWH conceived the project, compiled and analyzed data, designed the figures, and wrote the paper.

Acknowledgments

Members of the FAB/Crawford lab at Simon Fraser University provided constructive criticism of the project. David Fisher, Carlos Schrager, Jon Slate, and anonymous reviewers provided helpful reviews that greatly improved on early versions of the study and manuscript. Thanks to Doug Brandon, Mark Collard, Maria Creighton, Mick Elliot, Dan Greenberg, Michael Krützen, Ruth Linsky, Fanny Mazzamurro, Brea McCauley, Mark Skinner, Larry Ulibarri, and Rutger Vos for help with data and analyses.

Funding

The data analysis and writing were supported by Natural Sciences and Engineering Research Council Discovery Grant R611839 to MWH.

Conflict of interest statement

The authors declare no conflicts of interest.

Data and code availability

All data analyzed in this study are available as supplementary data files.

References

Aagaard, J. E., Springer, S. A., Soelberg, S. D., & Swanson, W. J. (2013). Duplicate abalone egg coat proteins bind sperm lysin similarly, but evolve oppositely, consistent with molecular mimicry at fertilization. *PLoS Genetics, 9*, e1003287. https://doi.org/10.1371/journal.pgen.1003287

Adret, P., Dingess, K. A., Caselli, C. B., Vermeer, J., Martínez, J., Luna Amancio, J. C., van Kuijk, S. M., Hernani, Lineros, L. M., Wallace, R. B., Fernandez-Duque, E., & Di Fiore, A. (2018). Duetting patterns of titi monkeys (Primates, Pitheciidae: Callicebinae) and relationships with phylogeny. *Animals*, *8,* 178. https://doi.org/10.3390/ani8100178

Allingham, J., & Floriano, W. B. (2021) Genetic diversity in the IZUMO1-JUNO protein-receptor pair involved in human reproduction. *PLoS ONE*, *16*, e0260692. https://doi.org/10.1371/journal.pone.0260692

Anderson, A. P., & Jones, A. (2024) The relationship between sexual dimorphism and androgen response element proliferation in primate genomes. *Evolution*, *78*, 195–208. https://doi.org/10.1093/evolut/qpad181

Arenzana, N., Rodríguez de Córdoba, S., & Rey-Campos, J. (1995). Expression of the human gene coding for the α -chain of C4b-binding protein, C4BPA, is controlled by an HNF1-dependent hepatic-specific promoter. *The Biochemical Journal*, *308*, 613–621. https://doi.org/10.1042/bj3080613

Arnold, C., Matthews, L. J., & Nunn, C. L. (2010). The 10kTrees website: A new online resource for primate phylogeny. *Evolutionary Anthropology, 19*, 114–118. https://doi.org/10.1002/evan.20251

Avella, M. A., Xiong, B., & Dean, J. (2013). The molecular basis of gamete recognition in mice and humans. *Molecular Human Reproduction, 19,* 279–289. https://doi.org/10.1093/molehr/gat004

Barker, D., Meade, A., & Pagel, M. (2007). Constrained models of evolution lead to improved prediction of functional linkage from correlated gain and loss of genes, *Bioinformatics*, *23*, 14–20. doi.org/10.1093/bioinformatics/btl558

Brukman, N. G., Li, X. H., & Podbilewicz, B. (2022). Fusexins, HAP2/GCS1 and evolution of gamete fusion. *Frontiers in Cell and Developmental Biology, 9*, 824024. https://doi.org/10.3389/fcell.2021.824024

Buffone, M. G., Zhuang, T., Ord, T. S., Hui, L., Moss, S. B., & Gerton, G. L. (2008). Recombinant mouse sperm ZP3-binding protein (ZP3R/sp56) forms a high order oligomer that binds eggs and inhibits mouse fertilization in vitro. *The Journal of Biological Chemistry*, *283*, 12438–12445. https://doi.org/10.1074/jbc.M706421200

Carithers, L. J., & Moore, H. M. (2015). The genotype-tissue expression (GTEx) project. *Biopreservation and Biobanking, 13*, 307–308. https://doi.org/10.1089/bio.2015.29031.hmm

Carlisle, J. A, Gurbuz, D. H, & Swanson, W. J. (2024). Recurrent independent pseudogenization events of the sperm fertilization gene ZP3r in apes and monkeys. *Journal of Molecular Evolution, 92*, 695–702. https://doi.org/10.1007/s00239-024-10192-x

Carlisle, J. A., Glenski, M. A., & Swanson, W. J. (2022). Recurrent duplication and diversification of acrosomal fertilization proteins in abalone. *Frontiers in Cell and Developmental Biology, 10*, 795273. https://doi.org/10.3389/fcell.2022.795273

Cassini, M. H. (2020). Sexual size dimorphism and sexual selection in primates. *Mammal Review, 50,* 231–239. https://doi.org/10.1111/mam.12191

Conley, A. J., Berger, T., Arias del Razo, R., Cotterman, R. F., Sahagúnc, E., Goetze, L. R., Jacob, S., Weinstein, T. A. R., Dufek, M. E., Mendoza, S. P., & Bales, K. L. (2022). The onset of puberty in colony-housed male and female titi monkeys (*Plecturocebus cupreus*): Possible effects of oxytocin treatment during peri-adolescent development. *Hormones and Behavior, 142,* 105157. https://doi.org/10.1016/j.yhbeh.2022.105157

Dale, B. (2024). Has the concept of polyspermy prevention been invented in the laboratory? *Zygote, 32,* 103–108. doi.org/10.1017/S0967199424000030

Dixson, A. F. (2012). *Primate sexuality: Comparative studies of the prosimians, monkeys, apes, and humans* (2nd ed.). Oxford University Press. doi.org/10.1093/acprof:osobl/9780199544646.003.0008

Evans, J. P., & Lymbery, R. A. (2020). Sexual selection after gamete release in broadcast spawning invertebrates. *Philosophical Transactions of the Royal Society B – Biological Sciences, 375*, 1813. http0s://doi.org/10.1098/rstb.2020.0069

Gavrilets, S. (2000). Rapid evolution of reproductive barriers driven by sexual conflict. *Nature, 403*, 886–889. https://doi.org/10.1038/35002564

Gordon, A. D. (2006). Scaling of size and dimorphism in primates. II. Macroevolution. *International Journal of Primatology*, *27*, 63–105. https://doi.org/10.1007/s10764-005-9004-1

Goudet, G., Mugnier, S., Callebaut, I., & Monget, P. (2008). Phylogenetic analysis and identification of pseudogenes reveal a progressive loss of zona pellucida genes during evolution of vertebrates. *Biology of Reproduction*, *78*(5), 796–806. https://doi.org/10.1095/biolreprod.107.064568

Gron, K. J. (2007). Primate factsheets: Dusky titi (*Callicebus moloch*) taxonomy, morphology, & ecology . <http://pin.primate.wisc.edu/factsheets/entry/dusky_titi/taxon>. Accessed 3 July 2020.

Harcourt, A. H., Harvey, P. H., Larson, S. G., & Short, R. V. (1981). Testis weight, body weight and breeding system in primates. *Nature, 293*, 55–57. https://doi.org/10.1038/293055a0

Harcourt, A. H., Purvis, A., & Liles, L. (1995). Sperm competition: Mating system, not breeding season, affects testes size of primates. *Functional Ecology, 9,* 468–476. doi.org/10.2307/2390011

Hart, M. W., Stover, D. A., Guerra, V., Mozaffari, S. V., Ober, C., Mugal, C. F., & Kaj, I. (2018). Positive selection on human gamete-recognition genes. *PeerJ, 6*, e4259. https://doi.org/10.7717/peerj.4259

Hirohashi, J., Kamei, N., Kubo, H., Sawada, H., Matsumoto, M., & Hoshi, M. (2008). Egg and sperm recognition systems during fertilization. *Development Growth and Differentiation, 50*, S221–S238. https://doi.org/10.1111/j.1440-169X.2008.01017.x

Ho, L. S. T., & Ane, C. (2014). A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Systematic Biology*, *63*, 397–408.

Huang, J.-L., Lv, C., Zhao, Y.-C., Li, W., He, X.-N., Li, P., Sha, A.-G., Tian, X., Papasian, C. J., Deng, H.-W., Lu, G.-X., & Xiao, H.-M. (2014). Mutant ZP1 in familial infertility. *The New England Journal of Medicine*, *370*, 1220–1226. https://doi.org/10.1056/NEJMoa1308851

Ikawa, M., Inoue, N., Benham, A. M., & Okabe, M. (2010). Fertilization: A sperm's journey to and interaction with the oocyte. *Journal of Clinical Investigation, 120,* 984–994. https://doi.org/ 10.1172/JCI41585

Inoue, N., Hamada, D., Kamikubo, H., Hirata, K., Kataoka, M., Yamamoto, M., Ikawa, M., Okabe, M., & Hagihara, Y. (2013). Molecular dissection of IZUMO1, a sperm protein essential for sperm-egg fusion. *Development*, *140*, 3221–3229. https://doi.org/10.1242/dev.094854

Ives, A. R., & Garland, T., Jr. (2010). Phylogenetic logistic regression for binary dependent variables. *Systematic Biology*, *59*, 9–6.

Kappeler, P. M. (1997). Intrasexual selection and testis size in strepsirhine primates. *Behavioral Ecology, 8*, 10–19. https://doi.org/10.1093/beheco/8.1.10

Kappeler, P. M., Cuozzo, F. P., Fichtel, C., Ganzhorn, J. U., Gursky-Doyen, S., Irwin, M. T., Ichino, S., Lawler, R., Nekaris, K. A.-I., Ramanamanjato, J.-B., Radespiel, U., Sauther, M. L., Wright, P. C., & Zimmermann, E. (2017). Long-term field studies of lemurs, lorises, and tarsiers, *Journal of Mammalogy*, 98*,* 661–669. https://doi.org/10.1093/jmammal/gyx013

Kennard, M. A., & Willner, M. D. (1941). Weights of brains and organs of 132 New and Old World monkeys. *Endocrinology, 28*, 977–984.

Levitan, D. R. (2012). Contemporary evolution of sea urchin gamete-recognition proteins: Experimental evidence of density-dependent gamete performance predicts shifts in allele frequencies over time. *Evolution, 66,* 1722–1736. https://doi.org/10.1111/j.1558-5646.2012.01608.x

Levitan, D. R., Buchwalter, R., & Hao, Y. (2019). The evolution of gametic compatibility and compatibility groups in the sea urchin *Mesocentrotus franciscanus*: An avenue for speciation in the sea. *Evolution, 73,* 1428–1442. doi.org/10.1111/evo.13766

Lindenfors, P., & Tullberg, B. S. (1998). Phylogenetic analyses of primate size evolution: The consequences of sexual selection. *Biological Journal of the Linnean Society*, *64*, 413–447. https://doi.org/10.1006/bijl.1998.0237

Lu, Y., Oura, S., Matsumura, T., Oji, A., Sakurai, N., Fujihara, Y., Shimada, K., Miyata, H., Tobita, T., Noda, T., Castaneda, J. M., Kiyozumi, D., Zhang, Q., Larasati, T., Young, S., Kodani, M., Huddleston, C. A., Robertson, M. J., Coarfa, C., Isotani, A., … & Ikawa, M. (2019). CRISPR/Cas9-mediated genome editing reveals 30 testis-enriched genes dispensable for male fertility in mice. *Biology of Reproduction, 101*, 501–511. https://doi.org/10.1093/biolre/ioz103

Luetenegger, W. & Cheverud, J. (1982). Correlates of sexual dimorphism in primates. *International Journal of Primatology, 3,* 387–402.

Lupold, S., Simmons, L. W., & Grueter, C. C. (2019). Sexual ornaments but not weapons trade off against testes size in primates. *Proceedings of the Royal Society B*, *286*, 20182542. http://doi.org/10.1098/rspb.2018.2542

Markham, R., & Groves, C. P. (1990). Brief communication: Weights of wild orang utans. *American Journal of Physical Anthropology*, *81,*1–3.

Masterson, T. J. (2003). Canine dimorphism and interspecific canine form in *Cebus. International Journal of Primatology*, *24*, 159–178.

McLure, C. A., Dawkins, R. L., Williamson, J. F., Davies, R. A., Berry, J., Longman-Jacobsen, N., Laird, R., & Gaudieri, S. (2004). Amino acid patterns within short consensus repeats define conserved duplicons shared by genes of the RCA complex. *Journal of Molecular Evolution, 59,* 143–157. https://doi.org/10.1007/s00239-004-2609-8

Meeusen, E. N. T., Walker, J., Peters, A., Pastoret, P.-P., & Jungersen, G. (2007). Current status of veterinary vaccines. *Clinical Microbiology Reviews, 20*, 489–510. https://doi.org/10.1128/CMR.00005-07

Metz, E. C., & Palumbi, S. R. (1996). Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin. *Molecular Biology and Evolution, 13,* 397-406. https://doi.org/10.1093/oxfordjournals.molbev.a025598.

Morgan, C. C., & Hart, M. W. (2019). Molecular evolution of mammalian genes with epistatic interactions in fertilization. *BMC Evolutionary Biology*, *19*, 154. https://doi.org/10.1186/s12862-019-1480-6

Muro, Y., Buffone, M. G., Okabe, M., & Gerton, G. L. (2012). Function of the acrosomal matrix: Zona pellucida 3 receptor (ZP3R/sp56) is not essential for mouse fertilization. *Biology of Reproduction*, *86*, 1–6. http://doi.org/10.1095/biolreprod.111.095877.

Okamoto, D. K. (2016). Competition among eggs shifts to cooperation along a sperm supply gradient in an external fertilizer. *American Naturalist, 187,* E129–E142. doi.org/10.1086/685813

Pan, R., & Oxnard, C. (2001). Metrical dental analysis on Golden Monkey (*Rhinopithecus roxellana*). *Primates*, *42*, 75-89.

Patiño, S., Keever, C. C., Sunday, J. M., Popovic, I., Byrne, M., & Hart, M. W. (2016). Sperm bindin divergence under sexual selection and concerted evolution in sea stars. *Molecular Biology and Evolution, 33*, 1988–2001. https://doi.org/10.1093/molbev/msw081

Plavcan, J. M. (2001). Sexual dimorphism in primate evolution. *Yearbook of Physical Anthropology, 44*, 25–53.

Plavcan, J. M., & Ruff, C. B. (2008). Canine size, shape, and bending strength in primates and carnivores. *American Journal of Physical Anthropology, 136,* 65–84. https://doi.org/10.1002/ajpa.20779

Ports, B. L., & Jensen-Seaman, M. I. (2024). Convergent rates of protein evolution identify novel targets of sexual selection in primates. *Evolution*, *78*, 364–377. doi.org/10.1093/evolut/qpad188

Prothmann, A., Laube, I., Dietz, J. Roos, C., Mengel, K., Zischler, H., & Herlyn, H. (2012). Sexual size dimorphism predicts rates of sequence evolution of SPermAdhesion Molecule 1 (SPAM1, also PH-20) in monkeys, but not in hominoids (apes including humans). *Molecular Phylogenetics and Evolution, 63,* 52-63. doi.org/10.1016/j.ympev.2011.12.008

Rivera, A. M., & Swanson, W. J. (2022). The importance of gene duplication and domain repeat expansion for the function and evolution of fertilization proteins. *Frontiers in Cell and Developmental Biology, 10,* 827454. https://doi.org/10.3389/fcell.2022.827454

Rohlfs, R. V., Swanson, W. J., & Seir, B. S. (2010) Detecting coevolution through allelic association between physically unlinked loci. *American Journal of Human Genetics*, *86*, 674–685. https://doi.org/10.1016.j.ajhg.2010.03.001

Shao, &., Zhou, L., Fang, L., Zhao, L., Zhang, B.-L., Shao, F., Chen, J.-W., Chen, C.-Y., Bi, X., Zhuang, X.-L., Zhu, H.-L., Hu, J., Sun, Z., Li, X., Wang, D., Rivas-González, I., Wang, S., Wang, Y.-M., Chen, W.,…,Wu, D.-D. (2023). Phylogenomic analyses provide

insights into primate evolution. *Science, 380*, 913–924. doi.org/10.1126/science.abn6919

Sherry, S. T., Ward, M. & Sirotkin, K. (1999) dbSNP—Database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Research*, *9*, 677–679.

Smith, R. J., & Jungers, W. L. (1997). Body mass in comparative primatology. *Journal of Human Evolution*, *32*, 523–559. doi.org/10.1006/jhev.1996.0122

Spence-Aizenberg, A., Williams, L. E., & Fernandez-Duque, E. (2018). Are olfactory traits in a pair-bonded primate under sexual selection? An evaluation of sexual dimorphism in *Aotus nancymaae*. *American Journal of Physical Anthropology, 166*, 884–894. https://doi: 10.1002/ajpa.23487

Swanson, W. J., & Vacquier, V. D. (2002). The rapid evolution of reproductive proteins. *Nature Reviews Genetics, 3,* 137–144. https://doi.org/10.1038/nrg733

Thorén, S., Lindenfors, P., & Kappeler, P. M. (2006). Phylogenetic analyses of dimorphism in primates: Evidence for stronger selection on canine size than on body size. *American Journal of Physical Anthropology, 130*, 50–59. https://doi.org/10.1002/ajpa.20321

Tombak, K. J., Hex, S. B. S. W., & Rubenstein, D. I. (2024). New estimates indicate that males are not larger than females in most mammal species. *Nature Communications, 15*, 1872. https://doi.org/10.1038/s41467-024-45739-5

Turner, L. M., & Hoekstra, H. E. (2008). Causes and consequences of the evolution of reproductive proteins. *The International Journal of Developmental Biology*, *52*, 769–780. https://doi.org/10.1387/ijdb.082577lt

Uy, J. A. C., & Borgia, G. (2000). Sexual selection drives rapid divergence in bowerbird display traits. *Evolution*, *54*, 273–278. https://doi.org/10.1111/j.0014-3820.2000.tb00027.x

Vacquier, V. D. (1998). Evolution of gamete recognition proteins. *Science, 281,* 1995–1998. https://doi.org/10.1126/science.281.5385.1995

Van Doorn, G. S., Luttikhuizen, P. C., & Weissing, F. J. (2001). Sexual selection at the protein level drives the extraordinary divergence of sex-related genes during sympatric speciation. *Proceedings of the Royal Society B Biological Sciences, 268*, 2155–2161. https://doi.org/10.1098/rspb.2001.1780

Wassarman, P. M. (2009). Mammalian fertilization: The strange case of sperm protein 56. *Bioessays*, *31*, 153–158. doi:10.1002/bies.2008

632 Weber, W. D., & Fisher, H. S. (2023). Sexual selection drives the coevolution of male
633 and female reproductive traits in *Peromyscus* mice. *Journal of Evolutionary Biology*, *36,*
634 67–81. doi.org/10.1111/jeb.14126

635 Supplementary Data

636 Appendix S1. Supplementary methods for identifying *ZP3R* pseudogenes in 17
637 genomes from Shao et al. (2023).

638 Appendix S2. Data and sources for five morphological variables in 48 primate species
639 (including incomplete data for 12 species).

640 Appendix S3. Complete data for five morphological variables in 36 primate species, with
641 R code and results for the principal components analysis (eigenvalues, eigenvectors,
642 scores for each species).

643 Appendix S4. Phylogenetic logistic regression analyses of *ZP3R* pseudogenization as a
644 function of relative male size.

645 Appendix S5. BayesTraits inferences of relative male size on internal edges of the
646 primate phylogeny.

Figure legends

Figure 1. *ZP3R* functionality and morphological variables from 31 primate genomes analyzed in Carlisle et al. (2024) and 17 other primate genomes analyzed in this study (marked by asterisks). Cladogram from 10kTrees (Arnold et al., 2010). New *ZP3R* data come from genomes in Shao et al. (2023). Red edges in the cladogram show 11 parallel origins of *ZP3R* pseudogenes in humans, *Pan*, *Gorilla gorilla*, gibbons, *Colobus angolensis* and *C. guereza*, *Pygathrix nigripes*, *Alouatta palliata*, *Aotus nancymaae*, the most recent common ancestor of *Callithrix jacchus* and *Saguinus midas,* and in *Plecturocebus donacophilus*. Twelve species names in plain type show species for which some morphological data were missing; those species were not included in the principal components analysis. Grey edges in the cladogram show species or lineages that were not included in the comparative analysis (calculation of relative size as residuals from lineage regressions, BayesTraits estimate of ancestral states for relative sizes, or one-sample tests) because those species were descended from the same single pseudogenization event or because data were missing. Superscript letters indicate footnotes that explain species names or give data sources for five morphological variables that were used in quantitative analyses: a, *Bunopithecus hoolock* in 10kTrees; b, *C. torquatus atys* in 10kTrees; c, *P. nemaeus* in 10kTrees (two species were recent split, probably hybridize); d, *C. capucinus* in 10kTrees; e, *Cebus apella* in 10kTrees; f, *Callicebus donacophilus* in 10kTrees; g, body masses from Lindenfors & Tullberg (1998) unless otherwise noted; h, Markham & Groves (1990) for "Sumatran" individuals; i, Thorén et al. (2006); j, https://primate.wisc.edu/primate-info-net/pin-factsheets/pin-factsheet-white-cheeked-gibbon; k, Kennard & Willner (1941); l, Gordon (2006); m, Smith & Jungers (1997); n, Lupold et al. (2019); o, Spence-Aizenberg et al. (2018); p, Gron (2007); q, canine lengths from Thorén et al. (2006) unless otherwise noted; r, Plavcan & Ruff (2008), similar to M. Skinner (unpublished data); s, as *Hylobates concolor*; t, as *C. mitis*;u, as *Cercopithecus aethiops*; v, Pan & Oxnard (2001); w, to include this lineage in the analysis we used canine sizes for *A. trivirgatus*, which have adult body sizes similar to *A. nancymaae*; x, Leutenegger & Cheverud (1982); y, Masterson (2003); z, testis masses from Harcourt et al. (1981) unless otherwise noted; aa, Dixson (2012); ab, Ports & Jensen-Seaman (2024); ac, Harcourt et al. (1995); ad, to include this lineage in the analysis we used testis sizes for *P. cupreus* from Conley et al. (2022), in which adult body sizes and life histories are similar to *P. donacophilus*

ALT TEXT: A phylogeny for 48 primate species with data for body mass, canine length, and testis mass. Eighteen species have evolved a *ZP3R* pseudogene; others have a functional *ZP3R*.

Figure 2. Female and male body masses (on logarithmic axes), canine lengths (on linear axes), and male body mass and testis mass (on logarithmic axes) for all 30 primate species with a functional *ZP3R* (including some species with some missing data) and all 18 species with a pseudogene (including some such as the gibbons species that share a single pseudogenization event from a common ancestor) (see Figure 1)

ALT TEXT: Three scatter plots for sexually selected traits including male body mass plotted against female body mass, male canine length against female canine length, and testis mass against male body mass.

Figure 3. Scores on the second and third eigenvectors of a principal components analysis of five morphological variables from 21 primate species with a functional *ZP3R* (grey) and 15 species with a pseudogene (red); the analysis included only species for which no morphological data were missing. See Table 1 for the coefficients of each variable in each vector. Vertical and horizontal lines show mean scores for the two species groups

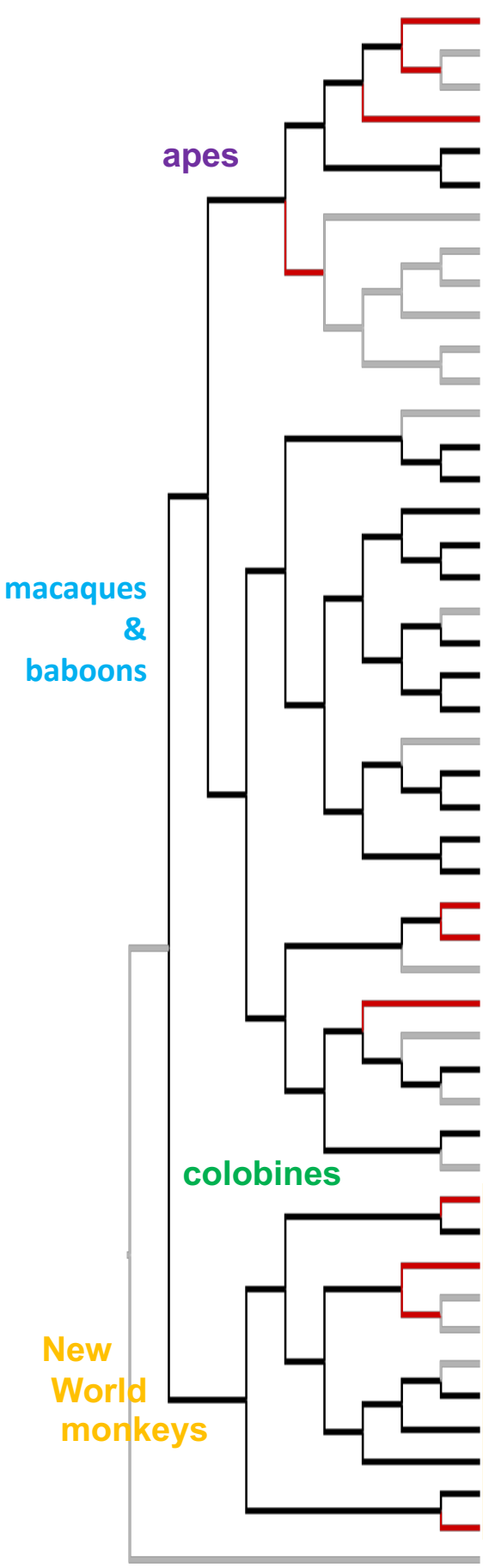
ALT TEXT: A scatter plot of two sexually selected traits based on results from the clustering analysis.

Figure 4. Relative male body mass, relative male canine length, and relative testis mass estimated as residuals from linear regressions. Each symbol shows values for a terminal edge of the phylogeny leading to one of 29 primate species (closed symbols) or values inferred using BayesTraits for one of 33 internal edges of the phylogeny (open symbols). Vertical and horizontal lines show mean values for the 11 total edges on which a pseudogene evolved (red) or the phylogeny-wide expected value on all other edges with a functional *ZP3R* (the known ancestral state). Species and edges included in the analysis are shown in Figure 1

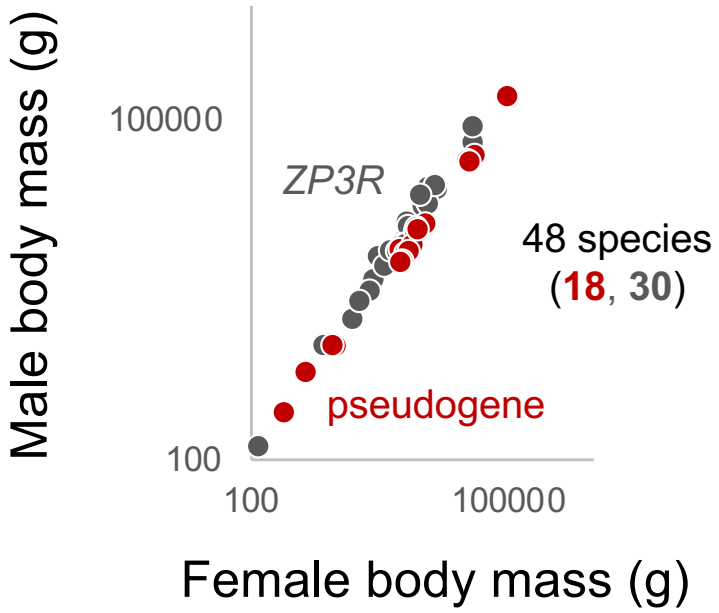
ALT TEXT: Two scatter plots of sexually selected traits based on estimates of ancestral trait values from the primate phylogeny.

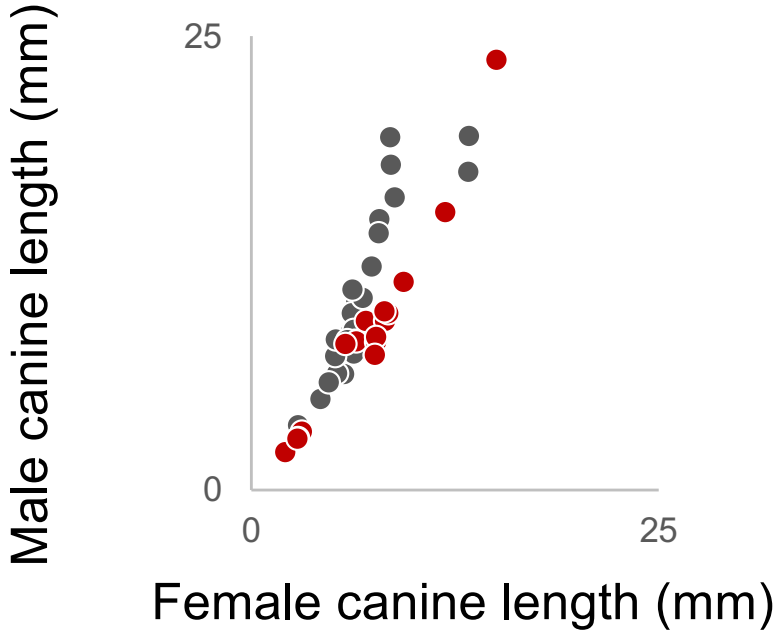
Figure 5. Relative male body mass and relative male canine length from Figure 4 disambiguated by the four major primate lineages (see Figure 1). Each symbol shows values for a terminal edge on the phylogeny leading to one of 29 primate species (closed symbols) or values inferred using BayesTraits for one of 33 internal edges of the phylogeny (open symbols). Vertical and horizontal lines show mean values for the 11 total edges on which a pseudogene evolved (red) or the phylogeny-wide expected value on all other edges with a functional *ZP3R* (the known ancestral state). Species and edges included in the analysis are shown in Figure 1.

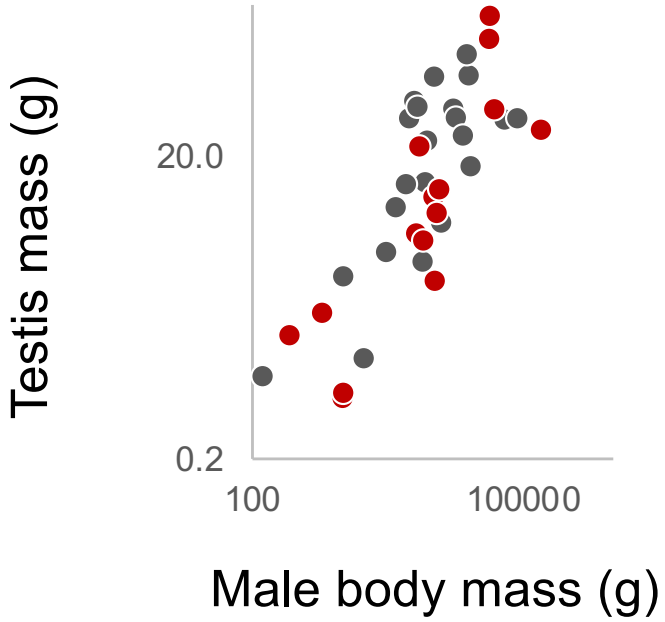
ALT TEXT: For scatter plots of relative male canine length plotted against relative male body mass shown separately for the apes, the macaques and baboons, the colobines, and the New World monkeys.

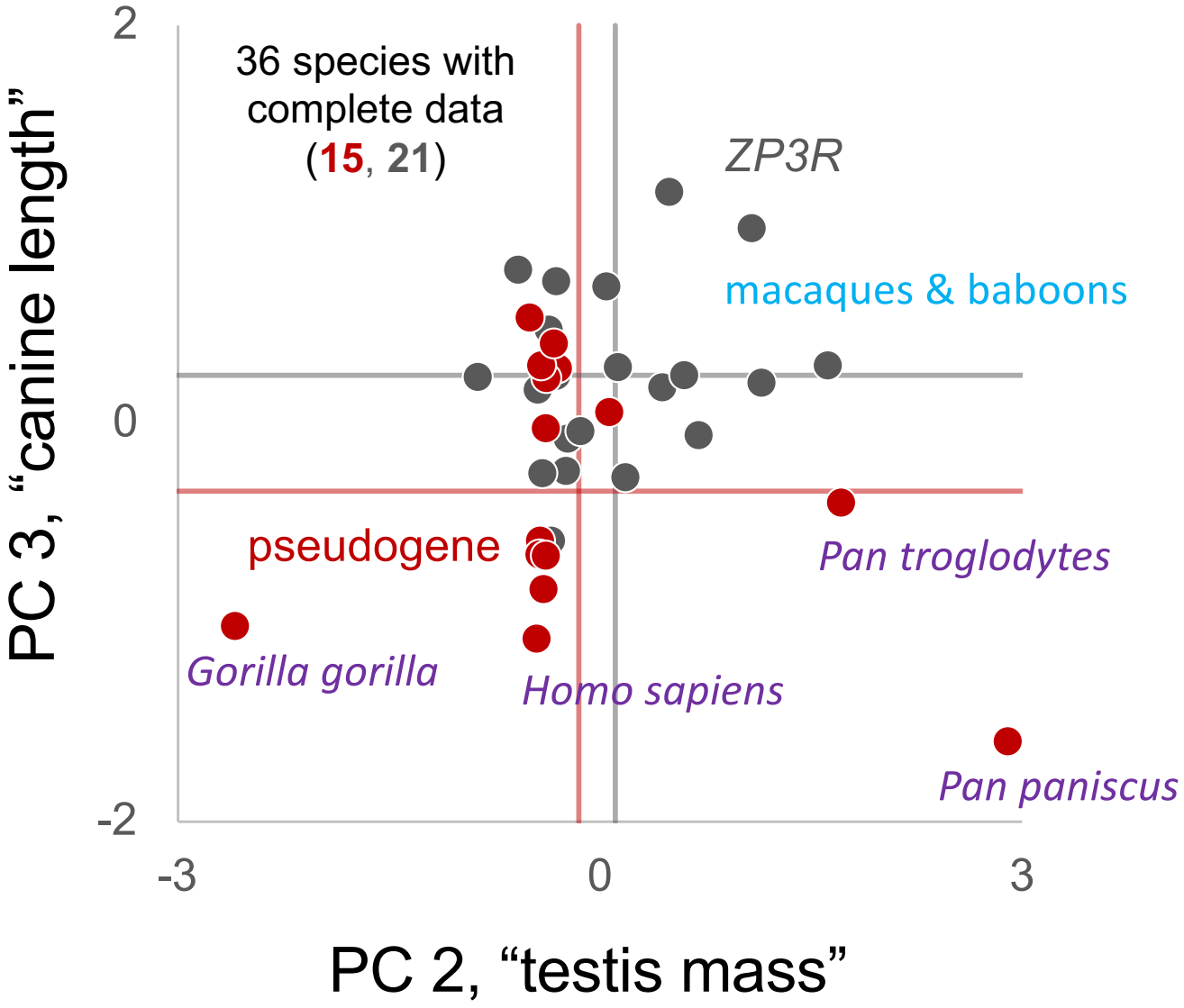


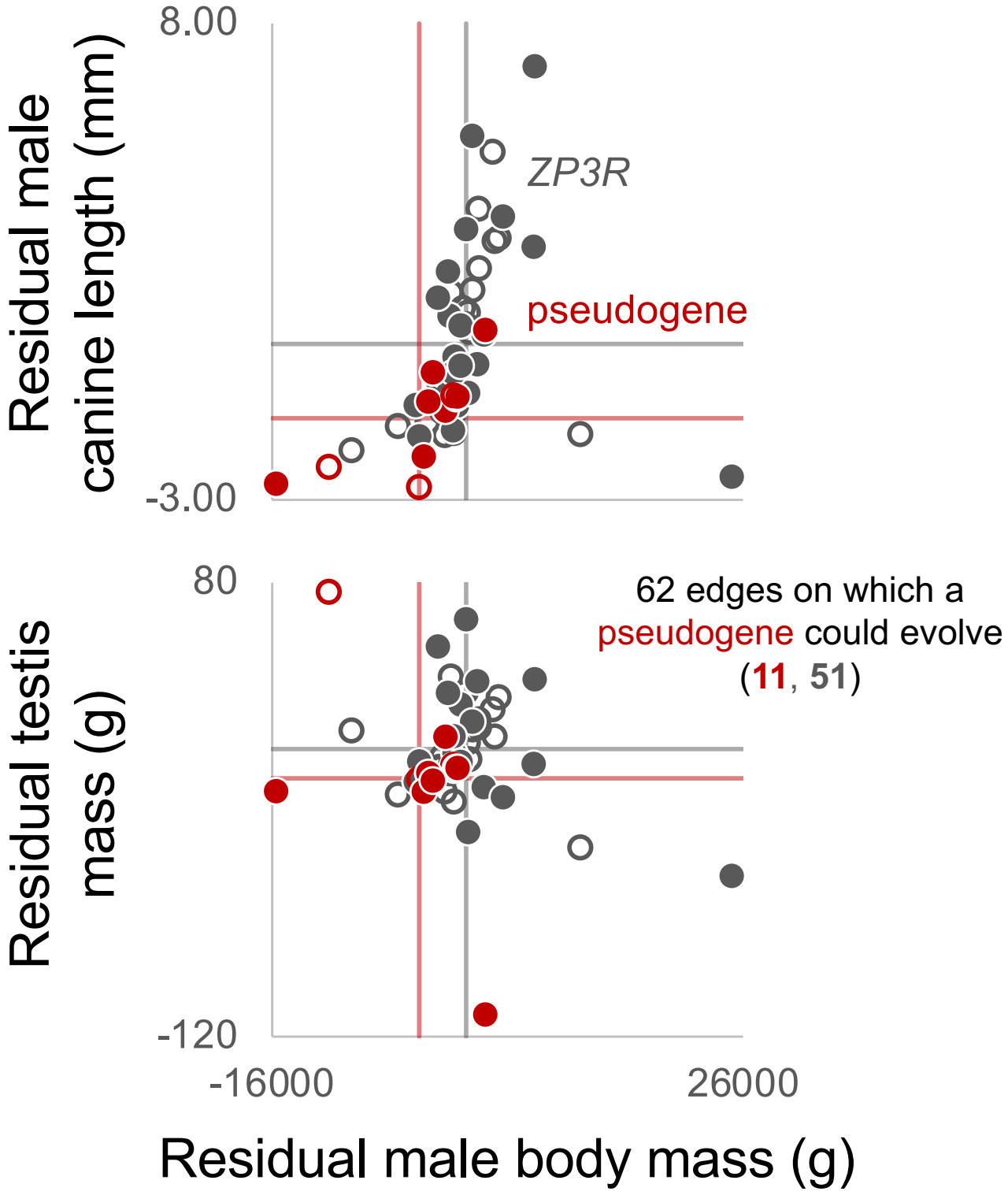
Species		Body mass (g)		Maxillary canine mesiodistal length (mm)		Testis mass (g)
		Female	Male	Female	Male	
Homo sapiens		40100 [g]	47900	8.24 [q]	9.29	40.50 [z]
Pan paniscus		33200	43000	9.37	11.44	167.70 [aa]
Pan troglodytes		35200	42000	11.94	15.30	118.80
Gorilla gorilla		97700	159200	15.09	23.67	29.60
Pongo abelii		38300 [h]	62140	13.37 [r]	19.50	34.80
Pongo pygmaeus	*	38700	86300	13.33	17.50	35.50 [ab]
Hoolock leuconedys [a]	*	6880	6870	7.69	8.42	
Hylobates moloch		6250 [i]	6580	7.60	7.40	6.10
Hylobates pileatus	*	5440	5500	7.62 [r]	7.42	
Symphalangus syndactylus	*	10700	11880	8.45	9.72	11.95 [n]
Nomascus leucogenys		7620 [i]	7790	7.69 [s]	8.14	5.50 [ab]
Nomascus siki	*	5800 [j]	5600			
Cercopithecus albogularis	*	4150	6900	5.22 [t]	8.29	
Chlorocebus aethiops	*	3560	5080	5.17 [u]	7.37	13.00
Erythrocebus patas	*	6500	12400	6.88	10.58	7.20 [n]
Cercocebus atys [b]		6100 [k]	8680	6.46	10.35	25.10
Mandrillus leucophaeus		10000	17000	8.60	17.89	41.05 [ac]
Mandrillus sphinx		11500	25000	8.54 [r]	19.40	68.00 [ac]
Lophocebus aterrimus	*	5760 [i]	7840	5.92	8.25	
Theropithecus gelada		13950	26100	7.88	14.89	17.10
Papio anubis		14700	24000	8.83	16.10	93.50
Papio hamadryas	*	9400	21500	7.85	14.14	27.10 [n]
Macaca assamensis	*	6700 [l]	11300			
Macaca nemestrina		7800	10400	7.42	12.29	66.70
Macaca silenus	*	5000	6800	6.25 [r]	11.04	42.00 [n]
Macaca fascicularis		3500	5500	6.21	9.73	35.20
Macaca mulatta		3000	6200	6.10	8.67	46.20
Colobus angolensis		8200	10200	7.05	9.30	10.70 [ab]
Colobus guereza	*	8700	10600	8.19	9.84	2.98 [ac]
Piliocolobus tephrosceles						
Pygathrix nigripes [c]	*	8440 [i]	11000	5.82 [r]	8.04	8.40 [n]
Rhinopithecus bieti						
Rhinopithecus roxellana		11600 [m]	17900	6.31 [v]	8.80	35.60 [ab]
Rhinopithecus strykeri	*					
Trachypithecus francoisi		7350 [n]	7700	6.40 [r]	8.20	3.98 [n]
Trachypithecus phayrei	*	6300 [i]	7870			
Alouatta palliata		5350	7150	6.46	8.16	23.00
Ateles geoffroyi		7460	8210	5.71	6.38	13.40
Aotus nancymaae		957 [o]	998	3.13 [w]	3.19	0.50 [ab]
Callithrix jacchus		240	260	2.11 [x]	2.06	1.30
Saguinus midas	*	430	590	2.84	2.80	1.83 [ac]
Cebus albifrons		1810	2480	4.79 [y]	5.91	
Cebus imitator [d]		2670	3870	5.28 [y]	6.42	9.10 [ab]
Sapajus apella [e]		2390	3050	6.32	7.49	4.64 [ac]
Saimiri boliviensis		700	1020	2.88 [r]	3.54	3.20 [ab]
Pithecia pithecia		1520	1730	4.26	5.00	0.92 [ac]
Plecturocebus donacophilus [f]		887 [p]	1017	2.89	2.84	0.54 [ad]
Tarsius syrichta		120	130			0.70 [ab]

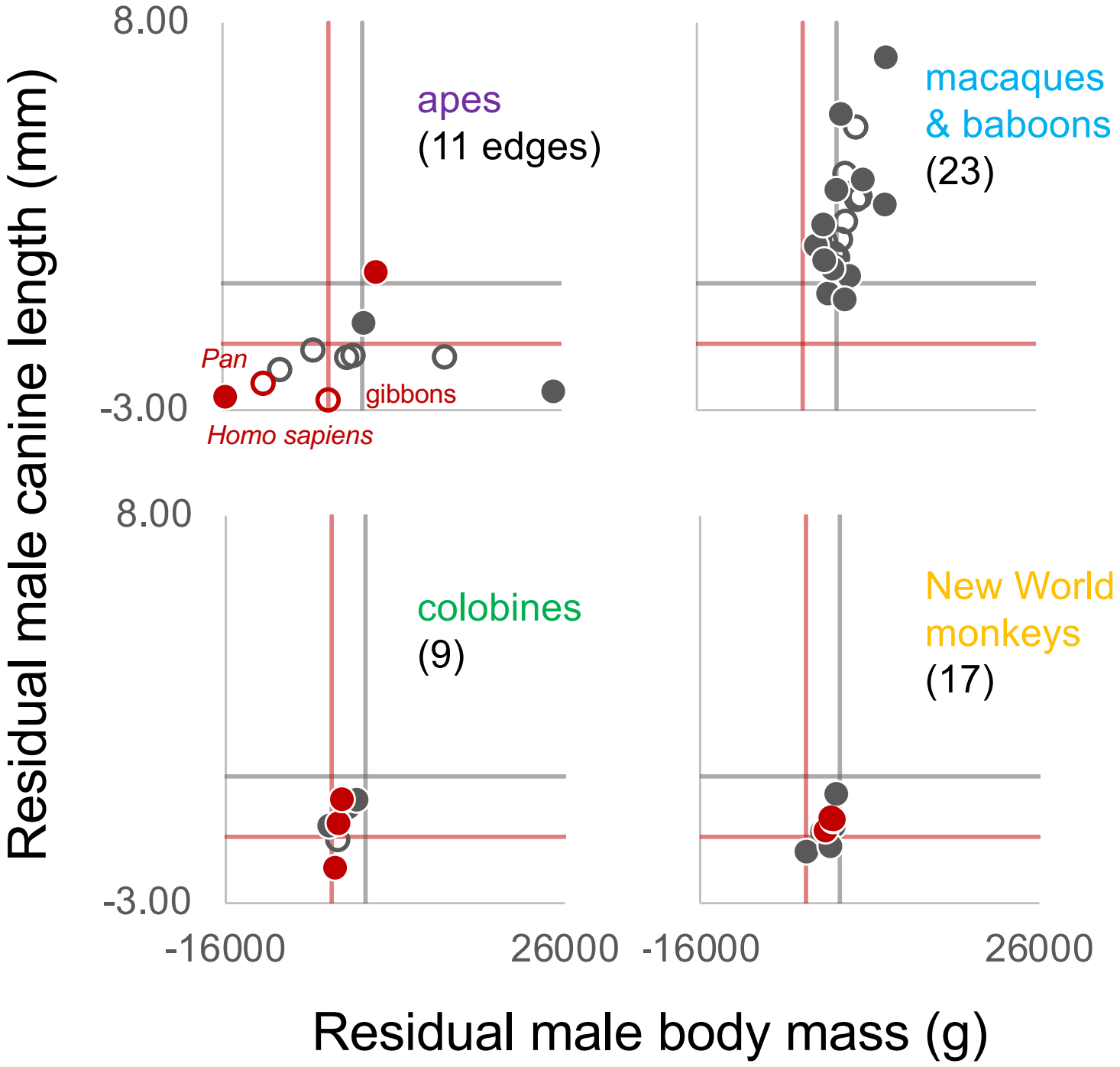

Male body mass (g)
100000
100
ZP3R
48 species
(18, 30)
pseudogene
100
100000
Female body mass (g)

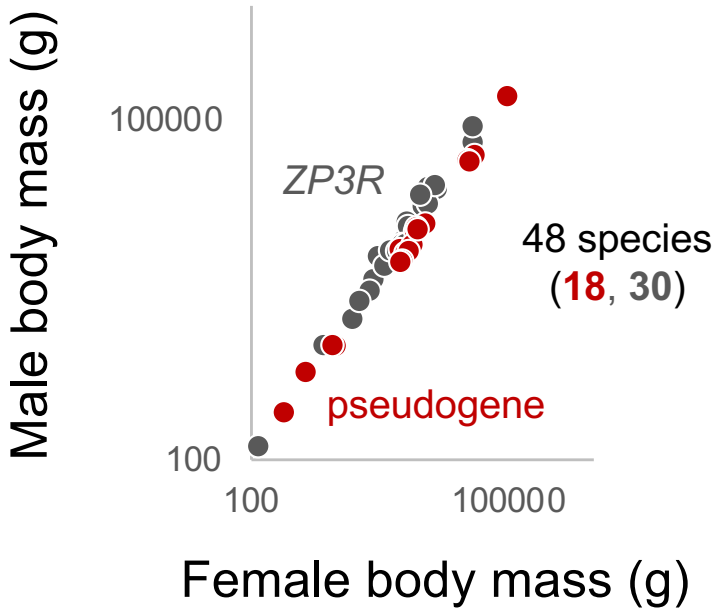

Male canine length (mm)
25
0
0
25
Female canine length (mm)

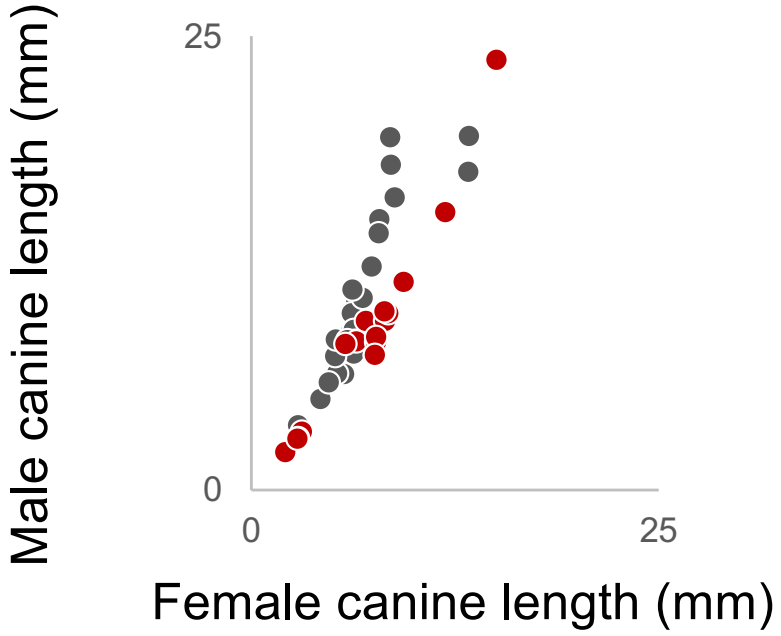

Testis mass (g)
20.0
0.2
100
100000
Male body mass (g)

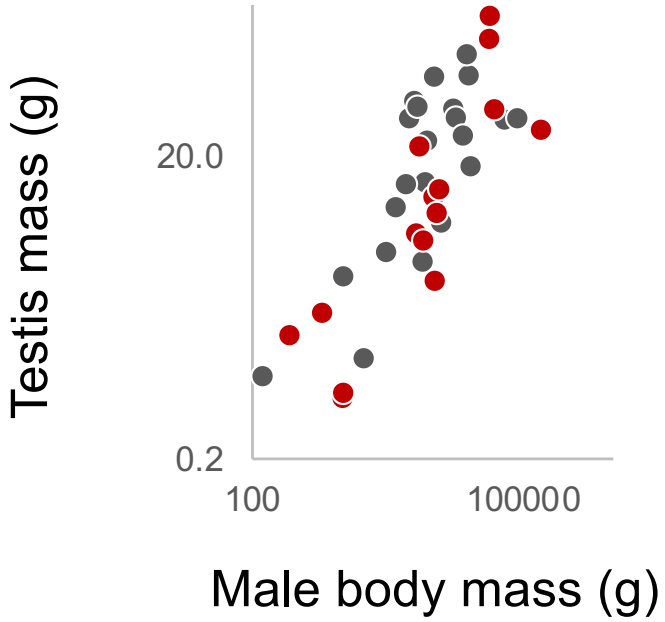

36 species with complete data (15, 21)
ZP3R
macaques & baboons
pseudogene
Pan troglodytes
Gorilla gorilla
Homo sapiens
Pan paniscus
2
0
-2
-3
0
3
PC 3, "canine length"
PC 2, "testis mass"

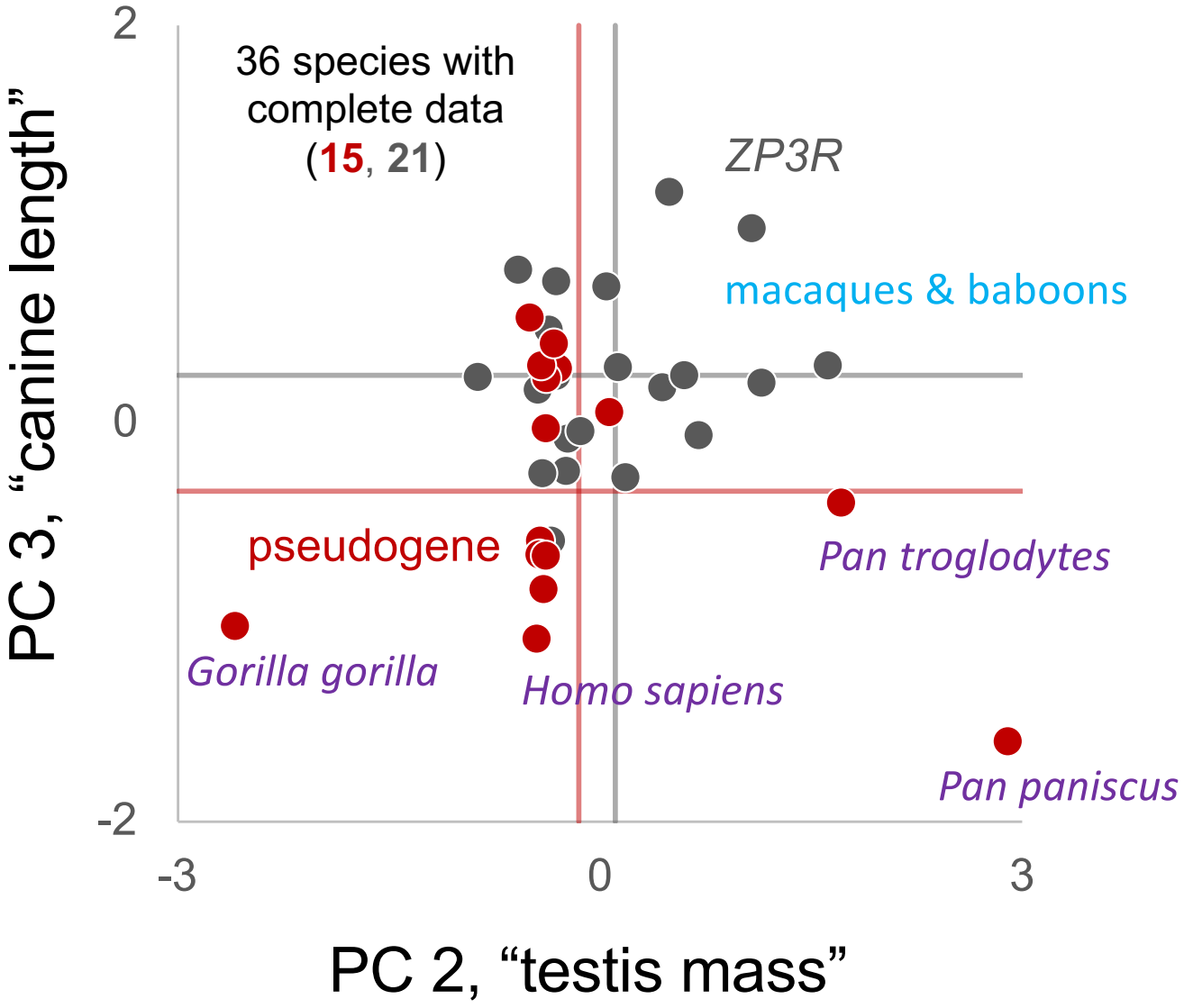

8.00
-3.00
Residual male canine length (mm)
ZP3R
pseudogene
80
-120
Residual testis mass (g)
62 edges on which a pseudogene could evolve (11, 51)
-16000
26000
Residual male body mass (g)

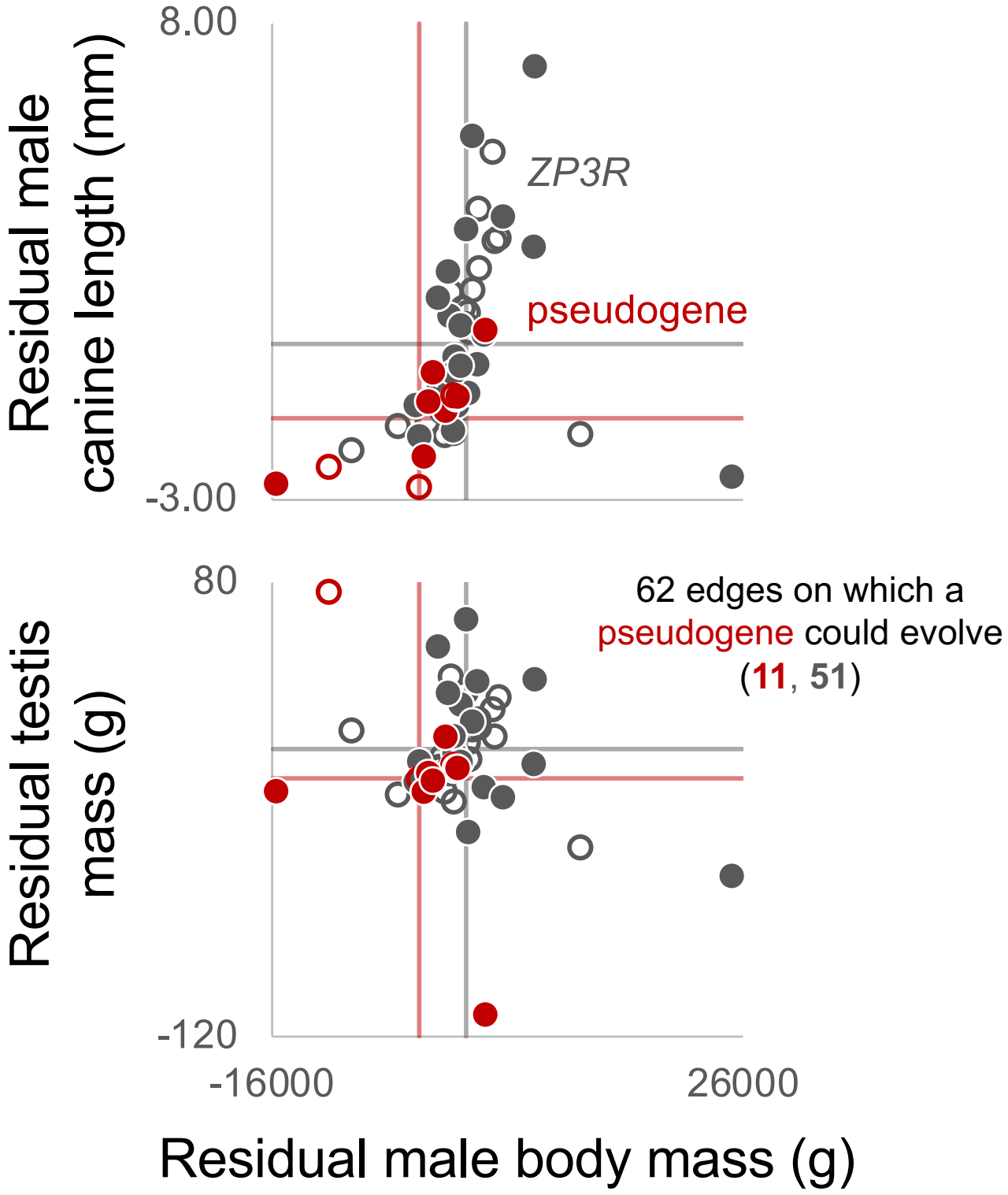

8.00
-3.00
apes
(11 edges)
Pan
gibbons
Homo sapiens
macaques
& baboons
(23)
8.00
-3.00
colobines
(9)
New World
monkeys
(17)
-16000
26000
-16000
26000
Residual male canine length (mm)
Residual male body mass (g)

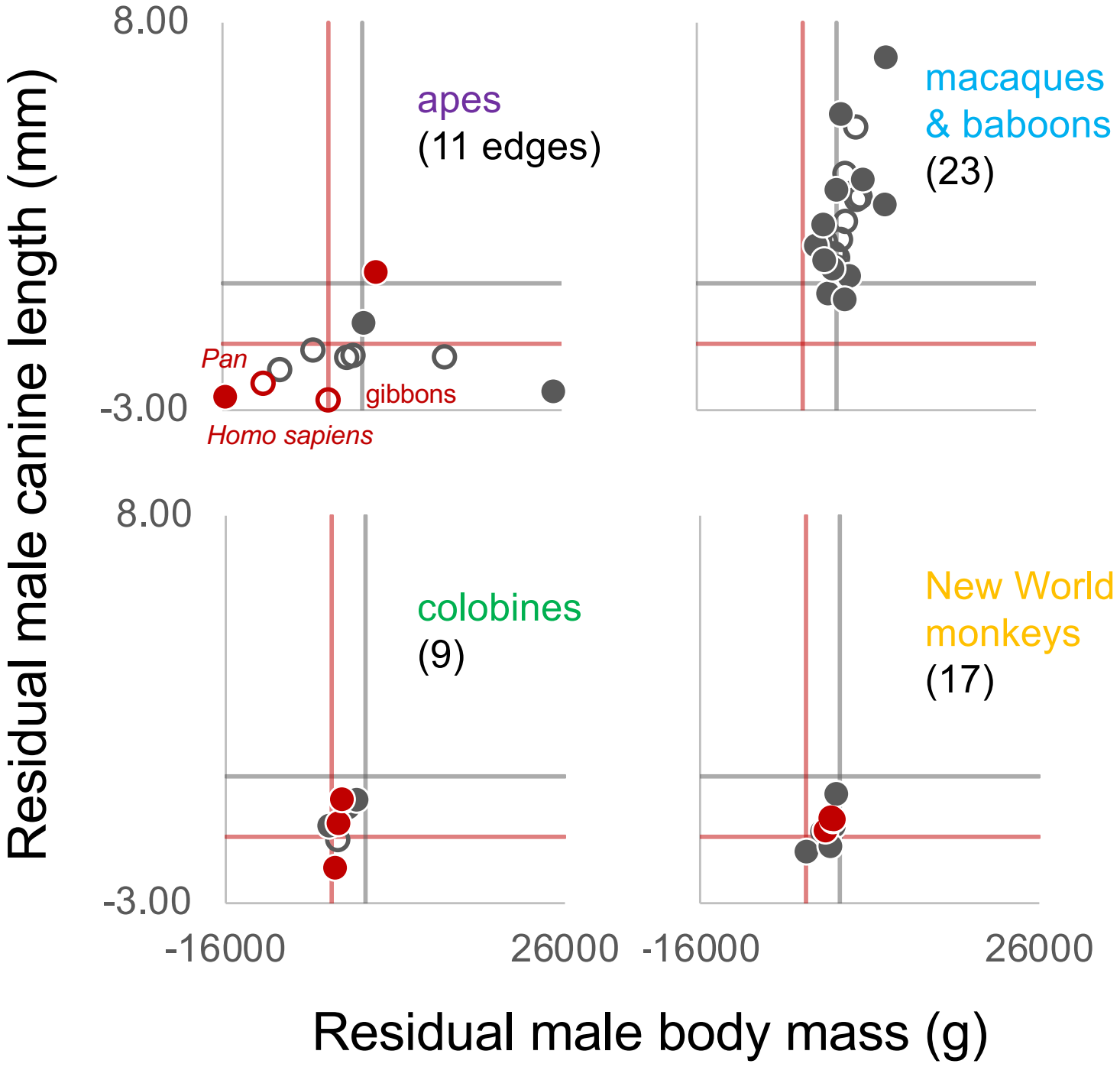

Male body mass (g)
100000
100
ZP3R
48 species
(18, 30)
pseudogene
100
100000
Female body mass (g)


Male canine length (mm)
25
0
0
25
Female canine length (mm)


Testis mass (g)
20.0
0.2
100
100000
Male body mass (g)


36 species with complete data (15, 21)
ZP3R
macaques & baboons
pseudogene
Pan troglodytes
Gorilla gorilla
Homo sapiens
Pan paniscus
2
0
-2
-3
0
3
PC 3, "canine length"
PC 2, "testis mass"


8.00
-3.00
Residual male canine length (mm)
ZP3R
pseudogene
80
-120
Residual testis mass (g)
62 edges on which a pseudogene could evolve (11, 51)
-16000
26000
Residual male body mass (g)


8.00
-3.00
apes
(11 edges)
Pan
gibbons
Homo sapiens
macaques
& baboons
(23)
8.00
-3.00
colobines
(9)
New World
monkeys
(17)
-16000
26000
-16000
26000
Residual male canine length (mm)
Residual male body mass (g)