University of Reading

Genes underlying the evolution of tetrapod testes size

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Baker, J., Meade, A. ORCID: https://orcid.org/0000-0001-7095-7711 and Venditti, C. ORCID: https://orcid.org/0000-0002-6776-2355 (2021) Genes underlying the evolution of tetrapod testes size. BMC Biology, 19. 162. ISSN 1741-7007 doi: 10.1186/s12915-021-01107-z Available at https://centaur.reading.ac.uk/130643/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1186/s12915-021-01107-z

Publisher: Springer

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

University of Reading

Central Archive at the University of Reading

Reading's research outputs online

Baker et al. BMC Biology (2021) 19:162
https://doi.org/10.1186/s12915-021-01107-z

BMC Biology

RESEARCH ARTICLE Open Access

Genes underlying the evolution of tetrapod testes size

Joanna Baker*, Andrew Meade and Chris Venditti*

Abstract

Background: Testes vary widely in mass relative to body mass across species, but we know very little about which genes underlie and contribute to such variation. This is partly because evidence for which genes are implicated in testis size variation tends to come from investigations involving just one or a few species. Contemporary comparative phylogenetic methods provide an opportunity to test candidate genes for their role in phenotypic change at a macro-evolutionary scale—across species and over millions of years. Previous attempts to detect genotype-phenotype associations across species have been limited in that they can only detect where genes have driven directional selection (e.g. brain size increase).

Results: Here, we introduce an approach that uses rates of evolutionary change to overcome this limitation to test whether any of twelve candidate genes have driven testis size evolution across tetrapod vertebrates—regardless of directionality. We do this by seeking a relationship between the rates of genetic and phenotypic evolution. Our results reveal five genes (*Alkbh5*, *Dmrtb1*, *Pld6*, *Nlrp3*, *Sp4*) that each have played unique and complex roles in tetrapod testis size diversity. In all five genes, we find strong significant associations between the rate of protein-coding substitutions and the rate of testis size evolution. Such an association has never, to our knowledge, been tested before for any gene or phenotype.

Conclusions: We describe a new approach to tackle one of the most fundamental questions in biology: how do individual genes give rise to biological diversity? The ability to detect genotype-phenotype associations that have acted across species has the potential to build a picture of how natural selection has sculpted phenotypic change over millions of years.

Keywords: Testes size, Genotype-phenotype associations, Evolutionary rates, Phylogenetic comparative methods

Background

Detecting which genes have driven phenotypic change is a fundamental goal in biology and the subject of many decades of research (e.g. [1, 2]). However, whilst we have a plethora of candidate genes implicated in phenotypes within individual (or a few) species, we still do not know which genes have driven phenotypic change across large taxonomic scales and over millions of years.

Hitherto, approaches for detecting continuous genotype-phenotype associations across species have sought a relationship between the strength of molecular adaptation (i.e. the rate of non-synonymous mutations, dN relative to the rate of synonymous mutations, dS) and the magnitude of a trait of interest, e.g. brain size [3, 4] or plumage coloration [5]. Such an association can reveal where adaptation in a given gene has driven phenotypic change over millions of years—for example, *SEMG2* (encoding a protein involved in semen coagulation) is linked to increasing testes size across primates [6]. However, where no relationship is found using these approaches, it merely indicates a lack of directional change rather than a lack of association per se. For example, *ASPM* (a gene involved in human microcephaly)

* Correspondence: j.l.a.baker@reading.ac.uk; c.d.venditti@reading.ac.uk
School of Biological Sciences, University of Reading, Reading RG6 6BX, UK

© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

has been shown using these phylogenetic approaches to drive both brain size increases and decreases across primates [7]. Current approaches are therefore inherently limited by the assumption that individual genes drive only unidirectional change. However, non-homogeneity in evolutionary change is now known to routinely occur at both molecular and phenotypic levels (e.g. [8, 9]).

It has recently been revealed that explosive bursts of rapid testes size change occurred during vertebrate evolutionary history [10] (Figure 1). Whilst testes size is a well-studied trait, we still know relatively little about which genes underlie the diversity observed across species. We use the substantial heterogeneity observed in the rate of testes mass evolution (Figure 1) to introduce and apply a novel approach which seeks to detect genes involved in testes size regardless of the directionality of change. Specifically, we test a suite of 12 genes previously implicated in mouse testes mass to determine whether they have played a wider role in the evolution of tetrapod testes size. Our results have the potential to reveal how natural selection at the genetic level can result in significant changes in species' phenotypes across millions of years.

Here, we introduce an approach to testing for genotype-phenotype associations acting across species and over millions of years that moves away from the limiting assumption of directionality. We harness the power of contemporary phylogenetic comparative approaches that reveal widespread and common heterogeneity in the rate of phenotypic evolution (e.g. [8, 11, 12]) reflecting variation in the strength of natural selection (e.g. [8, 11, 13–17]). What is of critical importance is that this variation in rate reflects periods of phenotypic change in any direction. With this in mind, we seek a relationship between the strength of molecular adaptation and the amount of historical testes size change measured by the relative rate of testes mass evolution (Figure 2).

In the absence of consistent directional phenotypic change, a relationship between molecular and phenotypic rates can reveal a link that would otherwise be hidden by studying only the magnitude of the phenotype (Figure 2a, c). A positive relationship between the rate of phenotypic evolution and the strength of molecular adaptation means that relatively high rates of protein-coding change are directly associated with rapid changes

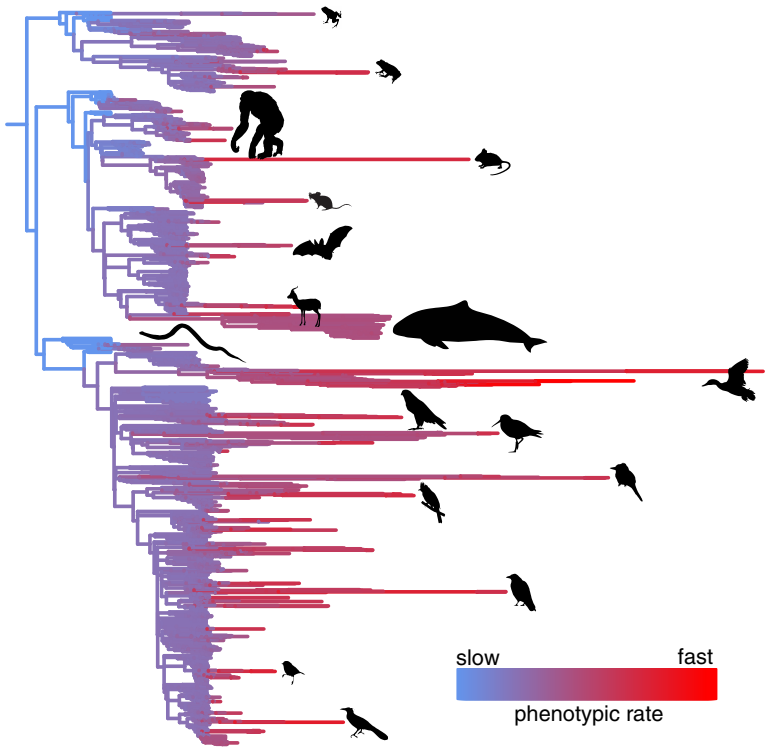

Fig. 1 Variation in the rate of testes size evolution across tetrapods. The branches of the tetrapod phylogeny ($N = 1845$) are measured to represent the rate of testes mass evolution (time multiplied by the median of the posterior distribution of rates estimated for each branch, methods). Lineages with comparatively fast rates of evolution in testes mass relative to body mass are represented by longer branches. Branches are additionally coloured according to their relative rate of phenotypic evolution (see scale). Silhouettes are added for purely illustrative purposes to indicate lineages with relatively rapid change in relative testes mass; they are not presented at any scale. All silhouettes are in the public domain and are obtained from phylopic.org

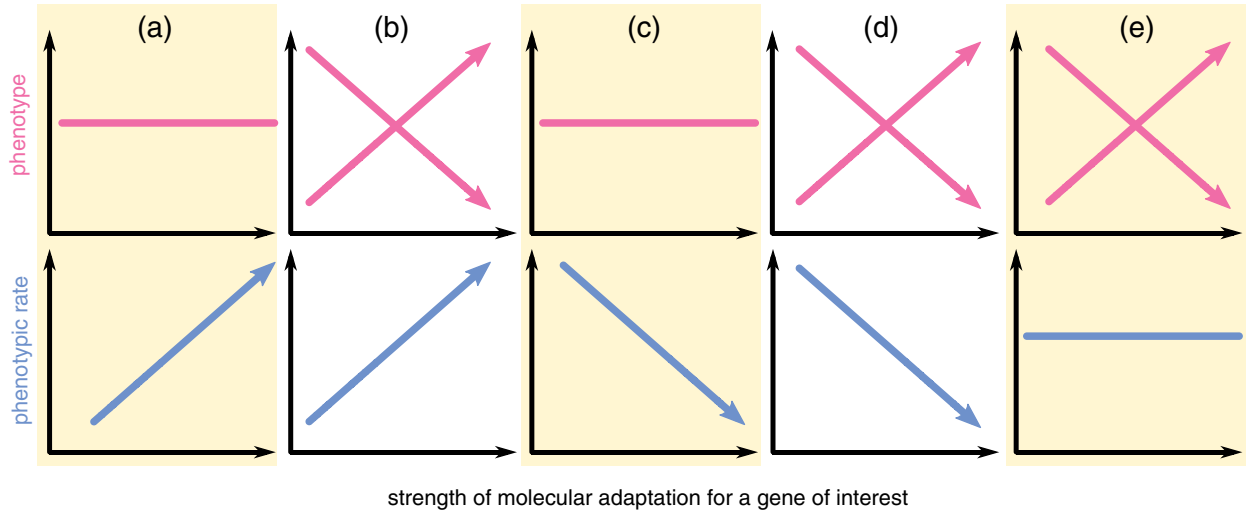

Fig. 2 Potential genotype-phenotype links for a hypothetical target gene. The relationship between strength of molecular adaptation (relative rate of protein-coding changes, dN given dS) are shown with magnitude of phenotype in pink and with rate of phenotypic evolution in blue. Crossed lines indicate an identical interpretation regardless of the direction. Where no association is found for the phenotype but is positive for rate (**a**), genetic selection has driven rapid (i.e. adaptive) phenotypic change in both directions during evolution (e.g. ASPM and primate brain size). If there is also an association with phenotype (**b**), interpretation is identical to (**a**), except the adaptation has been consistently directional. Where an association is not found for the phenotype but is negative for rate (**c**), molecular adaptation is explicitly associated with slower rates of phenotypic change—such genes can be thought of as phenotype moderators acting to minimise change. If there is an association also found for the phenotype (**d**), it means that although the gene is acting as a moderator, where molecular adaptation does occur it tends to have been in a consistent direction. Finally, where an association is found only for the phenotype and not the rate (**e**), it implies that this gene has driven directional phenotypic evolution, but phenotypic change has not occurred rapidly and thus may be associated with selection on another, associated phenotype

in the phenotype (Figure 2a, b). If protein-coding changes in the target gene have consistently acted to reduce phenotypic change, we would observe a negative relationship between the rate of phenotypic evolution and the strength of molecular adaptation (Figure 2c, d). Such a relationship means that a gene is actively reducing phenotypic variation: it is acting as some sort of phenotypic *moderator* and may have interactions or functional associations with other genes. In both cases, we can also study the phenotype to reveal directionality: where there is also a significant relationship between the phenotype and the strength of molecular adaptation, rapid change has been persistently unidirectional (Figure 2b, d, e.g. *SEMG2*); otherwise, the gene has driven change in both directions (Figure 2a, c). In a situation where we find an association with only the phenotype (Figure 1e), the gene of interest may have driven directional phenotypic change - though not rapidly.

Results and discussion

A recent genome-wide association study [18] revealed twelve genes within the mouse genome implicated in variation among testis weight. For each of the 12 target genes (Table 1), we downloaded [19, 20] and aligned [21] orthologous coding sequences for all tetrapods included in the time tree of life [22]. To maximise sample

Table 1 Target genes for testis size

Gene	D	df	N_{gene}	N_{codons}	N_{Taxa}	Outgroup
Alkbh5	1275.71*	375	193	373	101	*Xenopus*
Cdkn2c	582.93*	440	226	166	111	*Latimeria*
Dmrtb1	793.65*	388	200	151	97	*Latimeria*
Dnah11	5209.60*	311	159	4465	74	*Latimeria*
Lrp8	1790.75*	422	217	986	111	*Latimeria*
Ndc1	1266.98*	472	243	787	115	*Latimeria*
Nlrp3[a]	517.70*	232	120	1024	63	*Loxodonta*
Pld6	1057.41*	400	205	229	104	*Latimeria*
Sp4	1555.50*	464	239	757	117	*Latimeria*
Spata6	1179.64*	452	233	488	112	*Latimeria*
Sp8	1660.07*	319	163	506	87	*Latimeria*
Ubb	534.74*	300	156	126	79	*Xenopus*

Sample sizes, outgroups, and likelihood ratio (D) test results comparing a global vs. local model of molecular evolution are shown for each gene. N_{gene} is the number of taxa included in the model used to calculate dN and dS across the total number of codons (N_{codons}). N_{taxa} is the final number of taxa used in our regression models (Figure 3). *All comparisons are significant at the $p < 0.001$ level . [a]Note that Nlrp3 is restricted to mammals only—all other genes are analysed across tetrapods

size, we included all available sequences for each gene [see full sequence list in Table S1 in Additional file 1]. We used a maximum-likelihood codon-based substitution model with no site-to-site variation [23–25] implemented in HyPhY [26] to estimate branch-wise dN and dS values independently for each branch of the tetrapod phylogeny [22] (local model).

For all twelve target genes, we found significant evidence for branch-to-branch variation in molecular rates of evolution using standard likelihood ratio (D) tests (Table 1), comparing the fit of the local model for each gene to one which estimates only a single dN and dS across the whole alignment and phylogeny. To obtain a measure of molecular adaptation, for each species we summed all estimated rates along an evolutionary path of a species, leading from the root to the tip of the phylogeny. We use the root-to-tip dN (henceforth R_{dN}) accounting for the root-to-tip dS, R_{dS} (i.e. in Bayesian phylogenetic multiple regression analyses) as our measure of variation in the strength of molecular selection (Figure 2)—this allows for the potential that dS may be affected by fluctuating environments or ecological variability [27, 28].

We followed the approaches described in Baker et al. [10] in order to measure the rate of relative testes mass evolution across the branches of the tetrapod tree of life [22]. We used the variable rates model to measure rate heterogeneity in the phylogenetically structured residual error of the regression relationship between testes mass and body mass (n = 2036 vertebrate species, see Table S2 in Additional file 1 for data and references). The variable rates model simultaneously estimates three components [8, 11]: (i) the parameters of the regression model; (ii) an underlying background Brownian motion rate of evolutionary change (σ^2_b); and (iiii) a set of rates, r, which define whether each branch is evolving faster ($r > 1$) or slower ($r < 1$) than σ^2_b. We multiplied the original branch lengths of the phylogeny (measured in millions of years) by r to give a rate of testes size evolution along each individual branch relative to time (rt). Using Log Bayes Factors [29] (BF, see the "Methods" section), we found significant heterogeneity in the rate of testes size

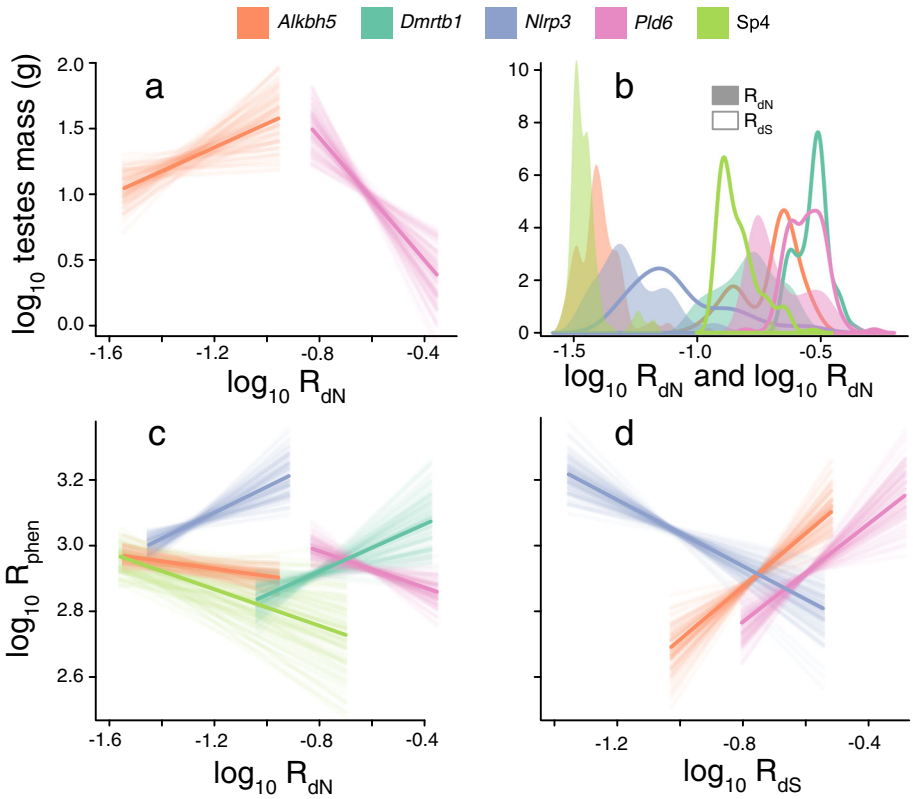

Fig. 3 Significant relationships between molecular adaptation and testes mass across tetrapods for five of the twelve studied target genes. The relationship between testes mass and R_{dN} (**a**) is significantly positive for Alkbh5 (the proportion of the parameter crossing zero, $p_x = 0.046$) and negative for Pld6 ($p_x = 0.002$). The relationship between R_{phen} and R_{dN} (**c**) is significant for all five genes ($p_x = 0.043$ for Alkbh5; $p_x = 0.036$ for Dmrtb1; $p_x = 0.006$ for Nlrp3; $p_x = 0.033$ for Pld6; $p_x = 0.048$ for Sp4). In both models, R_{dS} is included as a covariate, which is always larger than R_{dN} (**b**) in line with expectations, e.g. [28]. Whilst R_{dS} is not significantly associated with testes mass for any gene (not shown), it is significantly associated with R_{phen} for three genes (**d**): positively in both Alkbh5 ($p_x < 0.001$) and Pld6 ($p_x = 0.001$) and negatively for Nlrp3 ($p_x = 0.001$)

evolution across tetrapods (Figure 1): BF = 959.41 compared to a model estimating only a single rate of evolution. As with the molecular rates, we calculated the sum of all branch-wise *rt* for each species from root to tip as a measure of the total changes in phenotypic rate a species has experienced during its evolution, R_{phen}.

Our analyses provided three components of historical adaptation that have acted along the branches of the tetrapod phylogeny: (i) R_{dN}, (ii) R_{dS}, (iii) and R_{phen}. We used these components to test whether molecular adaptation (i.e. R_{dN} accounting for R_{dS}) in each of the twelve target genes has been linked to evolutionary change in testes mass. We conducted two sets of phylogenetic generalised least squares regression models implemented within a Bayesian framework [8], testing for an association between R_{dN} and either: (i) testes mass or (ii) R_{phen}. In all models, we include both R_{dS} and body mass as covariates; for the second set of models where R_{phen} is the response, we additionally include testes mass as another covariate.

We find significant associations in five of the twelve target genes (Figure 3). However, just two of the target genes have a directional association between R_{dN} and testes mass (after accounting for R_{dS}): natural selection has driven testes size increase via *Alkbh5* and decrease via *Pld6*. One of these genes (*Alkbh5*) has been previously heralded as a "*top candidate gene*" for testes size [18] and both genes have been demonstrated to directly affect testes mass using knock-out experiments [30–32]. On the other hand, all five of the genes in which we find significant associations demonstrate a significant relationship between R_{dN} and R_{phen} (Figure 3, after accounting for $R_{dS,}$ testes mass and body mass). This highlights the ability of our approach to overcome the limiting assumption of directionality [33], making it possible to find genotype-phenotype associations in the absence of any unidirectional relationship (e.g. *Nlrp3*, Dmrtb1 and *Sp4*, Figure 3a,c).

Three genes (*Alkbh5*, *Pld6*, *Sp4*) appear to be moderators of testes size (Figure 3c), i.e. they demonstrate a significant negative association between R_{dN} and R_{phen} (as in Figure 1c,d). This is the only association found for *Sp4* (Figure 3), indicating that this gene is a simple testes mass moderator (i.e. Figure 2c) acting to minimise evolutionary changes in testes mass across tetrapods. However, both *Alkbh5* and *Pld6* appear to act as moderators with specific directionality (e.g. Figure 2d). Both genes show a general link between testes size change and the mutation rate (a significant positive association between R_{phen} and R_{dS}, Figure 3d), but where there has been strong molecular adaptation (higher relative R_{dN}), testes size change is minimised (lower R_{phen}, Figure 3c). Where protein-coding changes do result in testes size change, these tend to be towards larger (for *Alkbh5*) or smaller (for *Pld6*) size (Figure 3a). The three genes for which we detect a moderator role are each known to act to affect the behaviour or function of other proteins. *Sp4* is a transcription factor and by its very nature involved in transcription. *Alkbh5* is an RNA demethylase that acts to moderate and remove m6a base modifications [31]. High levels of m6a modifications are implicated in male infertility [30], and they are also thought to play a key role in spermatogenesis [34]. *Pld6* is involved in piRNA biogenesis [32]—an essential molecule involved in gene silencing [35] that also has demonstrated links to spermatogenesis [36]. Our results imply that the three moderator genes we reveal here may have functional relationships with other genes acting on tetrapod testes sizes whilst also (for *Alkbh5* and *Pld6*) demonstrating a clear directional effect on testes themselves [30, 31] (Figure 3a).

In two genes, *Nlrp3* and *Dmrtb1*, protein-coding changes have driven increased testes size variability, but they have not driven directional change throughout tetrapod evolution (Figure 3a, c). That is, increased levels of protein-coding change are associated with increased levels of testes size change. Although little is known about the specific function of *Dmrtb1*, it has been linked with murine spermatogenesis and is co-expressed with DMRT1 [37, 38], a gene that is fundamental to male sex determination and that must be expressed throughout an individual's life to maintain sexual identity [39, 40]. Missense mutations in *Nlrp3*, a gene dominantly associated with innate immune system function [41], have been linked to altered fertility in mice [18, 42]. Elevated rates of molecular change in immunity genes in primates have previously been linked to promiscuity and the potentially increased risk of sexually transmitted diseases [43].

In accordance with the nearly neutral theory of molecular evolution [28, 44], all twelve genes had much faster rates of synonymous substitutions compared to non-synonymous rates (Figure 3b, both are calculated relative to each other, see the "Methods" section). There is also an expected ubiquitous strong significant correlation between both R_{dN} and R_{dS} as well as branch-wise dN and branch-wise dS. However, in *Nlrp3*, we find a negative relationship between R_{dS} and R_{phen} (Figure 3d). That is, after accounting for R_{dN} (and both body mass and testes mass), synonymous mutations have acted to reduce rapid evolutionary changes in testes mass. This unexpected negative association implies that, for this gene at least, synonymous substitutions are likely to not be silent to selection. This is in line with evidence that not only can the dN/dS ratio vary substantially across loci [45], but also that synonymous mutations can explicitly alter the expression and function of the translated protein [46, 47]. Indeed, elevated *Nlrp3* expression has recently

been detected in testes tissue compared with that of other organs for both rodents and primates [48].

Of the twelve genes we studied here, there were seven genes in which we found no significant relationship between the rate of molecular evolution and the evolution of testes mass across tetrapods (*cdkn2c*, *Dnah11*, *Lrp8*, *Ndc1*, *Sp8*, *Spata6*, *Ubb*). However, we do not interpret this to mean that these genes are not necessarily important in the evolution of tetrapod testes size. For example, they may still be implicated in the evolution of other phenotypic characteristics such as sperm morphology, Sertoli cell development, and sperm motility. For example, *Spata6* has been shown to be associated with "*pinhead sperm*" (a morphological abnormality wherein spermatozoa have no or very small heads) [49]; variability in phenotypes such as sperm morphology may not be reflected well by gross measures of testes morphology such as mass [50]. Increasing availability of phenotypic data may allow future studies to tease apart more nuanced roles for all the genes we study here—and many beyond.

The twelve target genes that we study here are not the only ones that have previously been implicated in testes size variation (e.g. [6]). We also know that there have been different patterns of testes size change in different vertebrate groups [10] and for species with varying intensities of sperm competition [51]. Using our approach alongside the increasing availability of molecular data, it is now possible to begin to tease apart which genes have played a role in testes size evolution as well as understanding their underlying ecological drivers.

Conclusions

We introduce and implement a novel approach to detecting genotype-phenotype associations in order to reveal five genes that are likely to have played a key role in driving the diversity in testes size across tetrapod vertebrates. Our approach demonstrates a way to detect which genes might have driven evolutionary change in the absence of directional selection. This provides a novel and exciting opportunity for detecting previously hidden gene-phenotype associations across long evolutionary scales. In the future, it should be possible to use this approach to screen suites of genes including potential whole-genome scans for impacts on relevant morphology, allowing us to build the picture of how natural selection has sculpted diversity—from genes to protein to phenotype over millions of years of evolutionary history.

Methods

Data

The phylogenetic tree used for all of our described analyses (molecular, phenotypic, and combined) was the vertebrate time tree of life [22] (downloaded January 2020 using the TimeTree web resource, http://www.timetree.org/). In each analysis, we limited the phylogeny to only include taxa with data (see below for more details, and also Tables S1-S3 in Additional File 1).

In a recent genome-wide association study, Yuan et al. [18] identified three quantitative trait loci (QTL) within the mouse genome (n = 502) that are associated with variation among testis weight [18]. Within these three regions, a total of 36 genes were found to have previously been implicated in reproductive phenotypes, but only a subset of these were found to have experienced natural selection (dN/dS ratio > 1). Whilst it is possible to reveal associations using our approach in the absence of positive selection (Figure 2), genes with known positive selection are more likely to have experienced significant heterogeneity in dN and dS during tetrapod evolution. We therefore identify 12 target genes from the analyses presented in Yuan et al. [18] to which we apply our approach to detect whether any have played a key role in the evolution of tetrapod testis size (Table 1).

For each of the twelve genes named in Yuan et al. [18], we downloaded orthologous sequences using the NCBI orthologue browser [19] and Ensembl [20], using mouse sequences as reference. We only included one-to-one orthologues (excluding species with multiple copies of the gene). Spurious sequences were identified both by visual assessment and automatically using trimAL [52], retaining only sequences where at least 80% of all sites overlap with at least 75% of sequences in the alignment. We also downloaded an outgroup sequence for each gene: the coelacanth was preferred, but where this was not available, we used an alternative suitable outgroup (Table 1). After removal of spurious sequences, multiple alignments were generated using MUSCLE [21]. Each multiple alignment was then matched to the time tree of life [22] and uninformative sites were removed using trimAL's [52] heuristic algorithm optimised for trimming alignments analysed by maximum likelihood phylogenetic analyses. A full list of accession numbers for all sequences used in our analyses can be found in Table S1 [see Additional file 1]. The sample size for each gene can be found in Table 1.

Testes mass (g) and body mass (g) data was obtained from Baker et al. [10] and updated to include 2036 species matched to the updated time tree of life [22]. We provide the full testes size dataset and reference list in Table S2 [see Additional file 1]. The final dataset comprised testes mass and body mass measurements for 992 birds, 621 mammals, 200 amphibians, 32 squamates and 191 fish. Taking a wider taxonomic perspective to estimate the phenotypic rate (i.e. across all vertebrates rather than across tetrapods) allowed us to more accurately estimate the background rate of evolution [8],

but all subsequent analyses were performed at tetrapod level owing to genome duplications in teleost fishes [53]. All data was $\log_{10}$-transformed before analysis.

For both genetic and phenotypic data, we only include species that are included in the tetrapod time tree of life. In some cases, species were included by genus matching (see Tables S1-S3 for details, all included in Additional file 1). However, our results do not differ if these species were excluded.

Phenotypic rate variation

Variation in the rate of testes mass evolution after accounting for body size was detected using the *variable rates regression model* [8, 11] following Baker et al. [10] and using the time tree of life [22] with branch lengths measured in millions of years. We used a regression model that estimated a separate relationship between testes mass and body mass for each of the five major clades studied (fish, frogs, birds, mammals and reptiles) in line with previously published results [10]. The variable rates regression model works within a phylogenetic Bayesian Markov Monte Carlo (MCMC) framework to estimate rate heterogeneity in the phylogenetically structured residual error of our regression model along the branches of the tree [8] . For each branch, the model estimates a posterior distribution of rate scalars r which define whether each branch is evolving faster ($r > 1$) or slower ($r < 1$) than the overall background rate of testes size evolution—which is estimated simultaneously as an underlying Brownian motion rate (σ^2_b). We then stretch (or compress) the original branch lengths of the phylogeny measured in millions of years by the median branch-specific r values to create a *phenotypic rate-scaled* phylogenetic tree where branches reflect the inferred rates of evolutionary change in testes size during the course of vertebrate evolutionary history. In this rate-scaled phylogeny, stretched branches reflect increased rates of morphological evolutionary change whereas compressed branches represent lineages where testes mass has changed less than would be expected given σ^2_b. We ran each model for a total 1 billion iterations after convergence, sampling every 500,000. All chains were run 5 times and checked visually to ensure convergence. Results are qualitatively identical across all replicates.

The presence of significant heterogeneity in the rate of phenotypic evolution was determined using a Log Bayes Factor ($BF = -2\log_e[m_1/m_0]$)—where m_1 is the marginal likelihood of our variable rates model and m_0 is the marginal likelihood of a model with a single underlying background rate, σ^2_b) assessed against the table provided by Raftery [29]. Marginal likelihoods were estimated in BayesTraits v3.0.1 (www.evolution.rdg.ac.uk/BayesTraitsV3.0.1) using a stepping-stone sampler [54] where we sampled 1 million iterations for each of 500 stones, drawing values from a beta-distribution ($\alpha = 0.40$, $\beta = 1$) [54]. Where we found significant phenotypic evolutionary rate heterogeneity, we obtained species-specific values of phenotypic rate variation. To do this, we created a phenotypic-rate scaled phylogeny using the median set of r values estimated in our variable rates model. We then summed all rate-scaled branches along the evolutionary path of each species from root to tip. This is a measure of the total amount of change in phenotypic rate a species has undergone during its evolutionary history [13], termed here R_{phen}. Any non-independence in R_{phen} among taxa is directly accounted for by using a phylogenetic comparative model which explicitly incorporates shared phylogenetic history (i.e. shared path lengths). Note that our path lengths are measured from the phylogenetic tree that is scaled using the median rate scalar acting along each branch. Whilst the median represents a suitable summary value, future advancements may allow us to incorporate the full posterior distribution of path lengths and rates acting throughout the phylogeny.

Molecular rate variation

Variation in the rate of molecular evolution was detected using codon-based substitution models as implemented in HyPhy [26]. All analyses were run on the time tree of life, removing species for which we did not have sequence data (see Table S1). The tree was then unrooted and branch-length information (in millions of years) was removed. For each gene, we ran two MG94xREV [24, 55] codon models to estimate separate independent synonymous (dS) and non-synonymous (dN) rates of evolution [56], but did not allow site-to-site variation. The first was a local model, where a single dN and dS was estimated across the full alignment and the second was a global model where a separate dN and dS was estimated for each individual branch. We assessed whether there was significant evidence for branch-to-branch variation in molecular rates of evolution using standard likelihood ratio (D) tests, comparing the fit of the local and global models.

Where we found significant evidence for branch-to-branch variation in dN and dS (likelihood ratio tests, see below), we obtained species-specific values of molecular rate variation in the same way as described for phenotypic rate variation above. We created two new trees, where branch lengths are respectively measured as the branch-specific dN and dS as estimated by our codon models. These trees were rooted on the outgroup taxa which was then removed. The branch lengths of these molecular rate trees are proportional to the instantaneous rate and are therefore time-independent. However, time is accounted for in all our models both as a part of

the phylogenetic tree (where branches are measured in time) and within the phenotypic rate (see above for more details). We then summed all branches from root to tip for each species to obtain R_{dN} and R_{dS}: measures of the total amount of change in dN and dS a species has experienced during its evolution. We did this independently for each gene. As with R_{phen}, non-independence in R_{dN} and R_{dS} is directly accounted for by using a phylogenetic comparative model (see below). Using these root-to-tip measures and accounting for the shared evolutionary history among them allows us to study substitution rates (and phenotypic rates) that are much more in line with taxon-level properties (testes size, body size, etc.) than by using alternative approaches such as by studying terminal branches alone [57]. Studying the rate of evolution (be it molecular or phenotypic) throughout the whole tree can also give increased statistical power [57].

Detecting genotype-phenotype associations

We used our data on testes mass, body mass, and phenotypic and molecular rate heterogeneity in order to determine whether any of the proposed candidate genes [18] have driven changes in testes mass during the course of tetrapod evolutionary history.

We determine whether each gene has had any directional effect on testes mass evolution using a PGLS regression model estimating the relationship between testes mass, R_{dN} and R_{dS}—whilst also accounting for body size. Sample sizes were limited to those species which had both molecular and phenotypic data (see Table 1) and were not large enough to test this relationship independently in different taxonomic groups. We then determined whether any changes in selection pressure on each individual gene have led to any adaptive response in testes mass by conducting a second PGLS regression in which we test the effects of testes mass, body mass, R_{dn} and R_{dS} on R_{phen}. All data were $\log_{10}$-transformed before analysis.

All PGLS models are conducted in a Bayesian MCMC framework using BayesTraits and estimate lambda [58] as a measure of phylogenetic signal. All models were run on the time tree of life including only those species for which we had both molecular and phenotypic data. All chains are run for 1 million iterations, sampling every 500 thousand iterations after convergence. We assess parameter significance by calculating the proportion of the posterior distribution of a parameter estimate that crosses zero (p_x)—where $p_x < 0.05$, less than 5% of the distribution crosses zero and we consider the parameter to be significant.

Our approach is uniquely placed to detect genotype-phenotype associations that have acted across species. A recently proposed method, *RERconverge* [59] seeks to link trait change to rates of molecular evolution, but differs from our approach in several important ways: firstly, our approach allows us to control for multiple explanatory factors. Here, we account for both testes mass and body mass to rule out any spurious associations owing to a possible general association between the rate of molecular evolution and testes size and body size. Secondly, *RERconverge* relies on non-specific rates of genetic evolution that do not distinguish between non-synonymous and synonymous changes. Finally, *RERconverge* depends on ancestral state reconstruction to estimate phenotypic rate heterogeneity.

Ensuring robustness of results

Following the rationale and logic of Montgomery et al. [4], we conducted additional analyses to determine the specificity of the associations we find and exclude the possibility of type 1 errors. We repeated our main analyses (detecting molecular rate variation and genotype-phenotype associations) on nine additional loci identified from human orthologues with no known implications for testes size (*PAX6*, *MC1R*, *MAOA*, *FOXP2*, *ENAM*, *GATA2*, *MATK*, *SHH*, *TYRP1*—see table S3 [Additional file 1] for all accessions). This control set includes genes that have been previously demonstrated to be under positive selection among various vertebrate lineages and had sequences available for a large subset of taxa overlapping with the testes size data. In line with this, we find evidence for molecular rate heterogeneity ($p < 0.001$, see table S4 in Additional file 1) for all nine control genes. However, when we test for an association between testes mass or the rate of testes mass evolution and the rate of molecular evolution, we find no evidence for any significant relationship. That is, we found no additional loci with significant implications for the evolution of tetrapod testes. This is good evidence that the significant associations we find in our candidate genes represent real biological effects and are unlikely to be type I errors.

For each of our regression models, we excluded each species one-by-one (for each gene we therefore a number of models equal to the number of species included in the full model, i.e. N_{taxa} in Table 1) to ensure no individual values for any of the measured data were driving the relationships (or lack thereof) we detected. Our results are robust to this process: we find identical results regardless of which species is excluded. We also repeated the analyses excluding amphibians: a general paucity of sequence data for amphibians means that for most genes we study here, there are only one or two sequences available (Tables S2, S3 in Additional File 1). We find no differences in the results: it seems that amphibians are not unusual with regards to the genes involved in tetrapod testes mass evolution. However, as

more genomic data for amphibians becomes available, future investigations could verify this.

For the twelve target genes, several orthologues are classified as "low-confidence" orthologues by ENSEMBL [20] (e.g. *ENSANIP00000026038*, see Table S1 in Additional file 1), meaning that they are below the threshold of one of the three following criteria: (i) Whole Genome Alignment scores which measure how well orthologous genes fall within aligned genomic regions, (ii) a gene order conservation score measures how much an orthologue falls within a block of genes and in the same order amongst its closest relatives, and (iii) and the percentage identity compared with the original sequence (in this case, the mouse). Note that the values of these thresholds vary depending on the taxa being studied; see ENSEMBL help pages [20]. Our results remained completely unchanged when we repeated our analyses excluding all species that are low-confidence orthologues from the calculation of molecular rates in HyPhy—including all post-hoc regression models. Our results are also unaffected if we exclude "low-quality" sequences from GenBank [19], i.e. those that have been altered to correct for potential mismatches between the nucleotide and protein sequences (e.g. Accession *XM_003732882.3*, see Table S1 in Additional file 1).

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12915-021-01107-z.

Additional file 1.

Acknowledgements

Not applicable.

Authors' contributions

All authors contributed to the conception and design of this work, as well as the acquisition, analysis, and interpretation of data. All authors read and approved the final manuscript.

Funding

This work was supported by the Leverhulme Trust (ECF-2017-022 to JB, RPG-2017-071 and RL-2019-012 to CV).

Availability of data and materials

All sequence data is available as described in the "Methods" of this paper. We provide all accessions and IDs for all sequence data used in this study in Additional File 1 (Tables S1, S3). Testes mass (g) and body mass (g) data was obtained from Baker et al. [10] and updated to include 2036 species matched to the updated time tree of life [22]. This updated dataset is also included within Additional File 1 (Table S2).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 23 April 2021 Accepted: 19 July 2021
Published online: 18 August 2021

References

1. Thompson S, Clarke AR, Pow AM, Hooper ML, Melton DW. Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. Cell. 1989;56(2):313–21. https://doi.org/10.1016/0092-8674(89)90905-7.
2. Pontikos N, Murphy C, Moghul I, Arno G, Fujinami K, Fujinami Y, et al. Phenogenon: gene to phenotype associations for rare genetic diseases. PLoS ONE. 2020;15(4):e0230587. https://doi.org/10.1371/journal.pone.0230587.
3. McGowen MR, Montgomery SH, Clark C, Gatesy J. Phylogeny and adaptive evolution of the brain-development gene microcephalin (MCPH1) in cetaceans. BMC Evol Biol 2011;11(1):1, DOI: https://doi.org/10.1186/1471-2148-11-98.
4. Montgomery SH, Capellini I, Venditti C, Barton RA, Mundy NI. Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. Mol Biol Evol. 2011;28(1):625–38. https://doi.org/10.1093/molbev/msq237.
5. Nadeau NJ, Burke T, Mundy NI. Evolution of an avian pigmentation gene correlates with a measure of sexual selection. Proc R Soc Lond B Biol Sci. 2007;274(1620):1807–13. https://doi.org/10.1098/rspb.2007.0174.
6. Dorus S, Evans PD, Wyckoff GJ, Choi SS, Lahn BT. Rate of molecular evolution of the seminal protein gene SEMG2 correlates with levels of female promiscuity. Nat Genet. 2004;36(12):1326–9. https://doi.org/10.1038/ng1471.
7. Montgomery SH, Mundy NI. Evolution of ASPM is associated with both increases and decreases in brain size in primates. Evolution. 2012;66(3):927–32. https://doi.org/10.1111/j.1558-5646.2011.01487.x.
8. Baker J, Meade A, Pagel M, Venditti C. Positive phenotypic selection inferred from phylogenies. Biol J Linn Soc. 2016;118(1):95–115. https://doi.org/10.1111/bij.12649.
9. Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Pond SLK. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. Mol Biol Evol. 2015;32(5):1342–53. https://doi.org/10.1093/molbev/msv022.
10. Baker J, Humphries S, Ferguson-Gow H, Meade A, Venditti C. Rapid decreases in relative testes mass among monogamous birds but not in other vertebrates. Ecol Lett. 2019;23(2):283–92. https://doi.org/10.1111/ele.13431.
11. Venditti C, Meade A, Pagel M. Multiple routes to mammalian diversity. Nature. 2011;479(7373):393–6. https://doi.org/10.1038/nature10516.
12. Gardner JD, Laurin M, Organ CL. The relationship between genome size and metabolic rate in extant vertebrates. Philosophical Transactions of the Royal Society B. 2020;375(1793):20190146. https://doi.org/10.1098/rstb.2019.0146.
13. Baker J, Meade A, Pagel M, Venditti C. Adaptive evolution toward larger size in mammals. Proc Natl Acad Sci U S A. 2015;112(16):5093–8. https://doi.org/10.1073/pnas.1419823112.
14. Kutsukake N, Innan H. Detecting phenotypic selection by Approximate Bayesian Computation in phylogenetic comparative methods. In: Garamszegi LZ, editor. Modern phylogenetic comparative methods and their application in evolutionary biology. Berlin: Springer-Verlag; 2014. p. 409–24. https://doi.org/10.1007/978-3-662-43550-2_17.
15. Rabosky DL. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. PLoS ONE. 2014;9(2):e89543. https://doi.org/10.1371/journal.pone.0089543.
16. Duchen P, Leuenberger C, Szilágyi SM, Harmon L, Eastman J, Schweizer M, et al. Inference of evolutionary jumps in large phylogenies using Lévy processes. Syst Biol. 2017;66(6):950–63. https://doi.org/10.1093/sysbio/syx028.
17. Baker J, Venditti C. Rapid change in mammalian eye shape is explained by activity pattern. Curr Biol. 2019;29(6):1082–8. https://doi.org/10.1016/j.cub.2019.02.017.
18. Yuan JT, Gatti DM, Philip VM, Kasparek S, Kreuzman AM, Mansky B, et al. Genome-wide association for testis weight in the diversity outbred mouse population. Mamm Genome. 2018;29(5):310–24. https://doi.org/10.1007/s00335-018-9745-8.

19. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2009;37(Database issue):D26–31. https://doi.org/10.1093/nar/gkn723.
20. Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, et al. Ensembl 2020. Nucleic Acids Res. 2020;48(D1):D682–D8. https://doi.org/10.1093/nar/gkz966.
21. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7. https://doi.org/10.1093/nar/gkh340.
22. Kumar S, Stecher G, Suleski M, Hedges SB. TimeTree: a resource for timelines, timetrees, and divergence times. Mol Biol Evol. 2017;34(7):1812–9. https://doi.org/10.1093/molbev/msx116.
23. Yang Z, Swanson WJ, Vacquier VD. Maximum-likelihood analysis of molecular adaptation in abalone sperm lysin reveals variable selective pressures among lineages and sites. Mol Biol Evol. 2000;17(10):1446–55. https://doi.org/10.1093/oxfordjournals.molbev.a026245.
24. Muse SV, Gaut BS. A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. Mol Biol Evol. 1994;11(5):715–24. https://doi.org/10.1093/oxfordjournals.molbev.a040152.
25. Kosakovsky Pond S, Delport W, Muse SV, Scheffler K. Correcting the bias of empirical frequency parameter estimators in codon models. PLoS ONE. 2010;5(7):e11230. https://doi.org/10.1371/journal.pone.0011230.
26. Pond SLK, Muse SV. HyPhy: hypothesis testing using phylogenies. Statistical methods in molecular evolution: Springer; 2005. p. 125–81. https://doi.org/10.1007/0-387-27733-1_6.
27. Lartillot N, Poujol R. A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. Mol Biol Evol. 2011;28(1):729–44. https://doi.org/10.1093/molbev/msq244.
28. Ohta T, Gillespie JH. Development of neutral and nearly neutral theories. Theor Popul Biol. 1996;49(2):128–42. https://doi.org/10.1006/tpbi.1996.0007.
29. Raftery AE. Hypothesis testing and model selection. In: Gilks WR, Richardson S, Spiegelhalter DJ, editors. Markov Chain Monte Carlo in practice. London: Chapman & Hall; 1996. p. 163–87.
30. Tang C, Klukovich R, Peng H, Wang Z, Yu T, Zhang Y, et al. ALKBH5-dependent m6A demethylation controls splicing and stability of long 3′-UTR mRNAs in male germ cells. Proc Natl Acad Sci U S A. 2018;115(2):E325–E33. https://doi.org/10.1073/pnas.1717794115.
31. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang C-M, Li CJ, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell. 2013;49(1):18–29. https://doi.org/10.1016/j.molcel.2012.10.015.
32. Watanabe T, Chuma S, Yamamoto Y, Kuramochi-Miyagawa S, Totoki Y, Toyoda A, et al. MITOPLD is a mitochondrial protein essential for nuage formation and piRNA biogenesis in the mouse germline. Dev Cell. 2011;20(3):364–75. https://doi.org/10.1016/j.devcel.2011.01.005.
33. Montgomery SH, Mundy NI, Barton RA. *ASPM* and mammalian brain evolution: a case study in the difficulty in making macroevolutionary inferences about gene–phenotype associations. Proc R Soc Lond B Biol Sci. 2014;281(1778).
34. Lin Z, Tong M-H. m6A mRNA modification regulates mammalian spermatogenesis. Biochim Biophys Acta. 2019;1862(3):403–11. https://doi.org/10.1016/j.bbagrm.2018.10.016.
35. Aravin AA, Hannon GJ, Brennecke J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science. 2007;318(5851):761–4. https://doi.org/10.1126/science.1146484.
36. Carmell MA, Girard A, van de Kant HJG, Bourc'his D, Bestor TH, de Rooij DG, et al. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell. 2007;12(4):503–14. https://doi.org/10.1016/j.devcel.2007.03.001.
37. Hilbold E, Bergmann M, Fietz D, Kliesch S, Weidner W, Langeheine M, et al. Immunolocalization of DMRTB1 in human testis with normal and impaired spermatogenesis. Andrology. 2019;7(4):428–40. https://doi.org/10.1111/andr.12617.
38. Zhang T, Murphy MW, Gearhart MD, Bardwell VJ, Zarkower D. The mammalian Doublesex homolog DMRT6 coordinates the transition between mitotic and meiotic developmental programs during spermatogenesis. Development. 2014;141(19):3662–71. https://doi.org/10.1242/dev.113936.
39. Fahrioglu U, Murphy M, Zarkower D, Bardwell V. mRNA expression analysis and the molecular basis of neonatal testis defects in Dmrt1 mutant mice. Sexual Development. 2007;1(1):42–58. https://doi.org/10.1159/000096238.
40. Zarkower D. DMRT Genes in Vertebrate Gametogenesis. In: Wassarman, PM, editor. Current Topics in Developmental Biology. Massachussetts: Academic Press; 2013:327-356. https://doi.org/10.1016/B978-0-12-416024-8.00012-X.
41. Sharma D, Kanneganti TD. The cell biology of inflammasomes: mechanisms of inflammasome activation and regulation. J Cell Biol. 2016;213(6):617–29. https://doi.org/10.1083/jcb.201602089.
42. Meng G, Zhang F, Fuss I, Kitani A, Strober W. A mutation in the Nlrp3 gene causing inflammasome hyperactivation potentiates Th17 cell-dominant immune responses. Immunity. 2009;30(6):860–74. https://doi.org/10.1016/j.immuni.2009.04.012.
43. Wlasiuk G, Nachman MW. Promiscuity and the rate of molecular evolution at primate immunity genes. Evolution. 2010;64(8):2204–20. https://doi.org/10.1111/j.1558-5646.2010.00989.x.
44. Ohta T. Slightly deleterious mutant substitutions in evolution. Nature. 1973;246(5428):96–8. https://doi.org/10.1038/246096a0.
45. Yang Z, Nielsen R, Hasegawa M. Models of amino acid substitution and applications to mitochondrial protein evolution. Mol Biol Evol. 1998;15(12):1600–11. https://doi.org/10.1093/oxfordjournals.molbev.a025888.
46. Spencer PS, Siller E, Anderson JF, Barral JM. Silent substitutions predictably alter translation elongation rates and protein folding efficiencies. J Mol Biol. 2012;422(3):328–35. https://doi.org/10.1016/j.jmb.2012.06.010.
47. Hunt RC, Simhadri VL, Iandoli M, Sauna ZE, Kimchi-Sarfaty C. Exposing synonymous mutations. Trends Genet. 2014;30(7):308–21. https://doi.org/10.1016/j.tig.2014.04.006.
48. Walenta L, Schmid N, Schwarzer JU, Köhn F-M, Urbanski HF, Behr R, et al. NLRP3 in somatic non-immune cells of rodent and primate testes. Reproduction. 2018;156(3):231–8. https://doi.org/10.1530/REP-18-0111.
49. Yuan S, Stratton CJ, Bao J, Zheng H, Bhetwal BP, Yanagimachi R, et al. *Spata6* is required for normal assembly of the sperm connecting piece and tight head–tail conjunction. Proceed Nat Acad Sci. 2015;112(5):E430–E9. https://doi.org/10.1073/pnas.1424648112.
50. Ramm SA, Schärer L. The evolutionary ecology of testicular function: size isn't everything. Biol Rev. 2014;89(4):874–88. https://doi.org/10.1111/brv.12084.
51. Harcourt AH, Purvis A, Liles L. Mating system, not breeding season, affects testes size of primates. Funct Ecol. 1995;9(3):468–76. https://doi.org/10.2307/2390011.
52. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25(15):1972–3. https://doi.org/10.1093/bioinformatics/btp348.
53. Hoegg S, Brinkmann H, Taylor JS, Meyer A. Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. J Mol Evol. 2004;59(2):190–203. https://doi.org/10.1007/s00239-004-2613-z.
54. Xie W, Lewis PO, Fan Y, Kuo L, Chen M-H. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. Syst Biol. 2010;60(2):150–60. https://doi.org/10.1093/sysbio/syq085.
55. Goldman N, Yang Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol. 1994;11(5):725–36. https://doi.org/10.1093/oxfordjournals.molbev.a040153.
56. Pond SK, Muse SV. Site-to-site variation of synonymous substitution rates. Mol Biol Evol. 2005;22(12):2375–85. https://doi.org/10.1093/molbev/msi232.
57. O'Connor TD, Mundy NI. Evolutionary modeling of genotype-phenotype associations, and application to primate coding and non-coding mtDNA rate variation. Evol Bioinform Online. 2013;9:301–16. https://doi.org/10.4137/EBO.S11600.
58. Pagel M. Inferring evolutionary processes from phylogenies. Zoologica Scripta. 1997;26(4):331–48. https://doi.org/10.1111/j.1463-6409.1997.tb00423.x.
59. Kowalczyk A, Meyer WK, Partha R, Mao W, Clark NL, Chikina M. RERconverge: an R package for associating evolutionary rates with convergent traits. Bioinformatics. 2019;35(22):4815–7. https://doi.org/10.1093/bioinformatics/btz468.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.