

## Research Article

# Elevated Evolutionary Rates among Functionally Diverged Reproductive Genes across Deep Vertebrate Lineages

Christopher J. Grassa<sup>1</sup> and Rob J. Kulathinal<sup>2</sup>

<sup>1</sup>Department of Botany, University of British Columbia, 6270 University Boulevard, Vancouver, BC, Canada V6T 1Z4

<sup>2</sup>Department of Biology, Temple University, 1900 North 12th Street, Philadelphia, PA 19122, USA

Correspondence should be addressed to Rob J. Kulathinal, robkulathinal@temple.edu

Received 1 February 2011; Revised 17 May 2011; Accepted 23 May 2011

Academic Editor: Jose M. Eirin-Lopez

Copyright © 2011 C. J. Grassa and R. J. Kulathinal. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Among closely related taxa, proteins involved in reproduction generally evolve more rapidly than other proteins. Here, we apply a functional and comparative genomics approach to compare functional divergence across a deep phylogenetic array of egg-laying and live-bearing vertebrate taxa. We aligned and annotated a set of 4,986 1 : 1 : 1 : 1 orthologs in *Anolis carolinensis* (green lizard), *Danio rerio* (zebrafish), *Xenopus tropicalis* (frog), *Gallus gallus* (chicken), and *Mus musculus* (mouse) according to function using ESTs from available reproductive (including testis and ovary) and non-reproductive tissues as well as Gene Ontology. For each species lineage, genes were further classified as tissue-specific (found in a single tissue) or tissue-expressed (found in multiple tissues). Within independent vertebrate lineages, we generally find that gonadal-specific genes evolve at a faster rate than gonadal-expressed genes and significantly faster than non-reproductive genes. Among the gonadal set, testis genes are generally more diverged than ovary genes. Surprisingly, an opposite but nonsignificant pattern is found among the subset of orthologs that remained functionally conserved across *all* five lineages. These contrasting evolutionary patterns found between functionally diverged and functionally conserved reproductive orthologs provide evidence for pervasive and potentially cryptic lineage-specific selective processes on ancestral reproductive systems in vertebrates.

## 1. Introduction

Over the past 550 million years, evolutionary processes have generated a diverse array of vertebrate species. Taxa that include fishes, birds, reptiles, and mammals evolved unique suites of adaptations allowing them to prosper in the most extreme sea, air, and land environments. Vertebrate diversity spans morphological innovations, developmental programming, cellular responses, as well as behaviors and life histories, and such differences become increasingly evident when taxa are compared across deep phylogenies. Studying the evolutionary patterns of functional change across this subphylum provides an opportunity to understand the evolutionary processes that have been important throughout vertebrate evolution. Yet, to date, clear *common* functional signatures that are in rapid flux across all vertebrate taxa have not been identified indicating the historical presence of a variety of niche- and lineage-dependent selective processes.

While functional evolutionary signals are not apparent across diverse phylogenetic lineages, when more closely related species such as sister species or multiple species within a single genus are compared, reproductive traits consistently reveal high diversity among species. This reproductive signature has been known for centuries, beginning with Linnaeus' binomial classification system [1]. Charles Darwin, in his 1871 treatise on sexual selection, also catalogued highly differentiated secondary sexual organs between closely related bird and mammal species [2]. Over a century later, William Eberhard described the diversity of morphological differences found in male secondary sexual traits, including vertebrate genitalia [3]. Both Darwin and Eberhard explain this higher male variance as the result of female mate choice or male-male competition on sexually selected traits within populations. The last three decades have amassed more vertebrate examples including cichlids [4], frogs [5], and primates [6] indicating that selection on reproductive

traits may be a common underlying evolutionary process in vertebrates.

Studying rates of morphological character change demonstrates how certain functional classes evolve relative to others and provides a lens into evolutionary processes of the past. While this framework works well on closely related species, signatures diminish when applied to distantly related taxa due to the presence of lineage-specific rates of development, selective constraints, and genetic architectural differences [7]. In addition, there are many processes in which the selected phenotype may be hidden or cryptic to human observers. Such phenotypes often occur at the molecular level and include immune response [8], gametic interactions [9], and pheromonal exchange [10]. To systematically understand the relative roles of different functional classes in the evolutionary history of vertebrates, and hence the role of certain selective processes, it would be instructive to employ a common and unbiased framework on a representative sample of taxa.

With the availability of annotated genomic sequences across an ever-expanding number of taxa in addition to associated functional data (e.g., ESTs, GO) that can link genes to function, an operational framework is emerging that compares rates of functional change across varying degrees of phylogenetic relatedness [11, 12]. By applying this functional and comparative genomics approach, we now can use normalized information from sequences to infer how functional categories of genes have changed in the past. Combining the two domains of time and function can provide valuable information about the history of these lineages, in particular, how certain selective forces act upon certain reproductive processes such as gamete recognition, oogenesis, spermatogenesis, and adult behavior.

In this paper, we quantify the rates of change among reproductive and non-reproductive genes in five distantly related vertebrate lineages. We functionally categorize ~5,000 orthologs using available testis, ovary, and non-reproductive EST libraries in each species and find that individual vertebrate lineages generally follow a pattern of greater divergence in genes solely expressed in the gonads compared to genes expressed in non-reproductive tissue. In most cases, the testis appears to be driving gonadal divergence. However, an opposing pattern emerges when we compare evolutionary rates among the much smaller subset of tissue-expressed genes that have remained functionally conserved across vertebrates ( $dN_{\text{testis}} < dN_{\text{ovary}} < dN_{\text{non-reproductive}}$ ). Using this framework, we are beginning to unmask a pattern of rapid and cryptic molecular evolution on lineage-specific reproductive features that are part of conserved developmental processes, thus, providing a common underlying genetic basis of functional evolutionary change in the vertebrate subphylum.

## 2. Materials and Methods

**2.1. Orthology and Estimates of Divergence.** Protein coding genes from *A. carolinensis*, *G. gallus*, *D. rerio*, *M. musculus*, and *X. tropicalis* were used in this analysis. Orthologs for each species pair were obtained from BioMart

(<http://uswest.ensembl.org/biomart/index.html>). Orthologs were filtered so that only transitive sets of 1:1:1:1:1 orthologs remained, producing 4,986 sets of 5-species orthologs. We excluded all paralogous relationships (including 10,122 1:1:2:1:1 relationships, where “2” denotes paralogous sequences from the zebrafish lineage) in order to maintain a relatively ambiguous ortholog set. The protein coding CDS and amino acid sequence of each gene’s longest transcript were also obtained from BioMart: in the case of transcript length ties, the transcript with the lower incremental Ensemble ID number was used. Multiple sequence alignments for each orthologous set of proteins were generated using MUSCLE (version 3.8; [13]) and then back-translated using corresponding CDS and a custom Perl script (available from CJG on request). All 1:1:1:1:1 alignments in addition to their associated functional assignments will be made available via *lizardbase* (<http://www.lizardbase.org/>) as an active link to current *A. carolinensis* annotations in *lizardbase*’s genome browser, JBrowse, and *lizardbase*’s Resources Page. All alignments will also be made available on the Resources page in *lizardbase*.

A protein distance matrix was calculated for each protein alignment using the Jones-Taylor-Thornton (JTT) model in the prodist program from the Phylip suite of phylogenetic programs (version 3.69; [14]). Consensus phylogenetic trees were generated using concatenated sequences from both CDS and its associated protein sequences (See in Supplementary Material available online at doi:10.4061/2011/274975 Supplementary Figure 1). For a given gene from each species, the mean of its four orthologous protein distances was used as one of two estimates of sequence divergence. A matrix of nonsynonymous substitutions per nonsynonymous site, dN, was also estimated for each codon alignment using Nei and Gojobori’s method [15] using the SNAP Perl program [16], and its mean dN across four orthologs was used as an estimate of sequence divergence.

**2.2. Functional Annotation Using EST Libraries and Gene Ontologies.** ESTs from each of the five species were filtered as “normal adult” tissue from NCBI’s dbEST (downloaded in October 2009) and assigned to species-specific tissue libraries (see Supplementary Table 1) based on either organ or tissue fields in the Genbank record. EST sequences were locally indexed and aligned to genes from the same species using a standalone version of blastn (version 2.22; [17]). EST-to-gene alignments of at least 100 nucleotides, 90% identity, and an *E*-value of  $e - 20$  were used as alignment criteria. For each of the five species lineages, genes with at least three ESTs (i.e., hits) meeting the above alignment criteria were assigned to seven non mutually exclusive functional classes: (1) genes with hits in only the testis were classified as testis-specific; (2) genes with hits in the testis and another tissue(s) were classified as testis-expressed; (3) genes with hits in only the ovary were classified as ovary-specific; (4) genes with hits in the ovary and another tissue(s) were classified as ovary-expressed; (5) genes with hits in only the testis and/or ovary were classified as gonadal-specific; (6) genes with hits in the testis and/or ovary, in addition to non-reproductive tissue(s), were classified as gonadal-expressed;

TABLE 1: Functional classification of reproductive (testis, ovary, gonadal) and non-reproductive orthologs in vertebrate species. Genes were assigned to at least one of seven functional categories (see Section 2.2 for explanation). Non-reproductive genes are found in neither the testis nor ovary EST libraries but are present in other tissues.

Functional classification	<i>A. carolinensis</i>	<i>D. rerio</i>	<i>X. tropicalis</i>	<i>G. gallus</i>	<i>M. musculus</i>
Testis-specific	55	49	129	12	43
Testis-expressed	613	2511	2825	1011	1659
Ovary-specific	87	21	13	16	0
Ovary-expressed	889	2422	1367	2033	350
Gonadal-specific	243	126	187	53	45
Gonadal-expressed	1109	3142	3109	2447	1848
Non-reproductive	243	537	613	940	2159
Total annotated genes (out of 4986)	<b>1352</b>	<b>3679</b>	<b>3722</b>	<b>3387</b>	<b>4007</b>

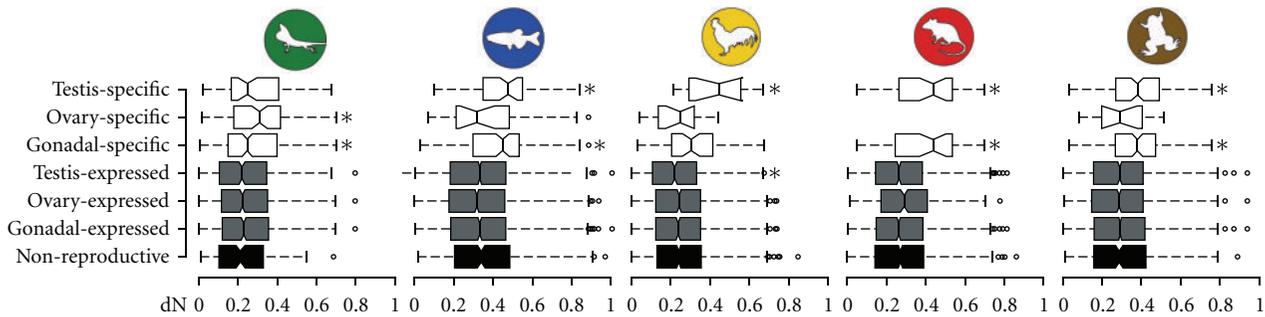


FIGURE 1: Protein divergence versus functional class across vertebrate lineages. Boxplots show the distribution of dN, nonsynonymous substitutions per nonsynonymous site in seven functional classes for each of the five species, *A. carolinensis*, *D. rerio*, *G. gallus*, *M. musculus*, and *X. tropicalis*. The three tissue-specific classes are found on the top (nonshaded), tissue-expressed classes are below in grey, and the non-reproductive functional class is indicated on the bottom, in black. Asterisks on the right-hand side of a boxplot signifies a highly significant ( $P < 0.001$ ) difference in mean, as given by the Wilcoxon rank sum test, when compared to the non-reproductive class. No ovary-specific genes were identified in *M. musculus*.

(7) genes with hits from an assortment of non-reproductive tissues (see Supplementary Table 1) that were neither testis nor ovary were classified as non-reproductive. Thus, for each of the five species, genes with sufficient EST coverage fell into at least one functional class (Table 1). The difference between the mean dN of each reproductive class and the mean dN of the non-reproductive class was tested using an unpaired two-sample two-sided Wilcoxon rank sum test ([18]; Figure 1).

We also compared evolutionary rates in *functionally conserved* genes, that is, those orthologs that do not change functional class across all five lineages, according to our EST annotations. Interestingly, we were not able to identify a single gonadal-specific gene, but were able to identify functionally conserved subsets of testis-expressed ( $n = 95$ ), ovary-expressed ( $n = 16$ ), and non-reproductive ( $n = 3$ ) orthologs. Figures 2(a)–2(g) provides Venn diagrams for all species combinations in each functional class. Figure 3 compares dN across four (nonzero) functional classes.

To complement the functional annotations generated by ESTs, we linked the 10% most diverged orthologs to the GO categories, Biological Process (BP) and Cellular Component (CC) in each species. GeneMerge [19] was used to test for statistically significant over-represented functional terms. A “word cloud” that relates the frequency of each

GO term to its font size was generated for four of the five species (Figure 4, Supplementary Figure 2). *X. tropicalis* was excluded from this analysis due to its sparse GO term set.

### 3. Results and Discussion

In this study, we chose five distantly related vertebrate species that fit the following criteria: (1) the presence of a well-assembled and freely available genome sequence, (2) the existence of well-curated gene models, (3) the availability of appreciable numbers of testis, ovary, and non-reproductive ESTs at dbEST, and (4) the condition that all five species, together, represent divergent clades thus presenting a deep vertebrate phylogeny with a diverse breadth of functional differences. After filtering out alignments that were of poor quality or had ambiguous orthologous relationships, a consensus tree-based off-concatenated CDS sequences from 4,986 orthologs was generated using the five vertebrate species. The tree’s topology was well supported in 100% of 1000 bootstrap replicates (Supplementary Figure 1). A concatenated protein tree-demonstrated the same topology and support (not shown) and mirrored published vertebrate phylogenies (e.g., [20]). We note that these ~5,000 orthologs represent a relatively “well-behaved” and conserved gene set

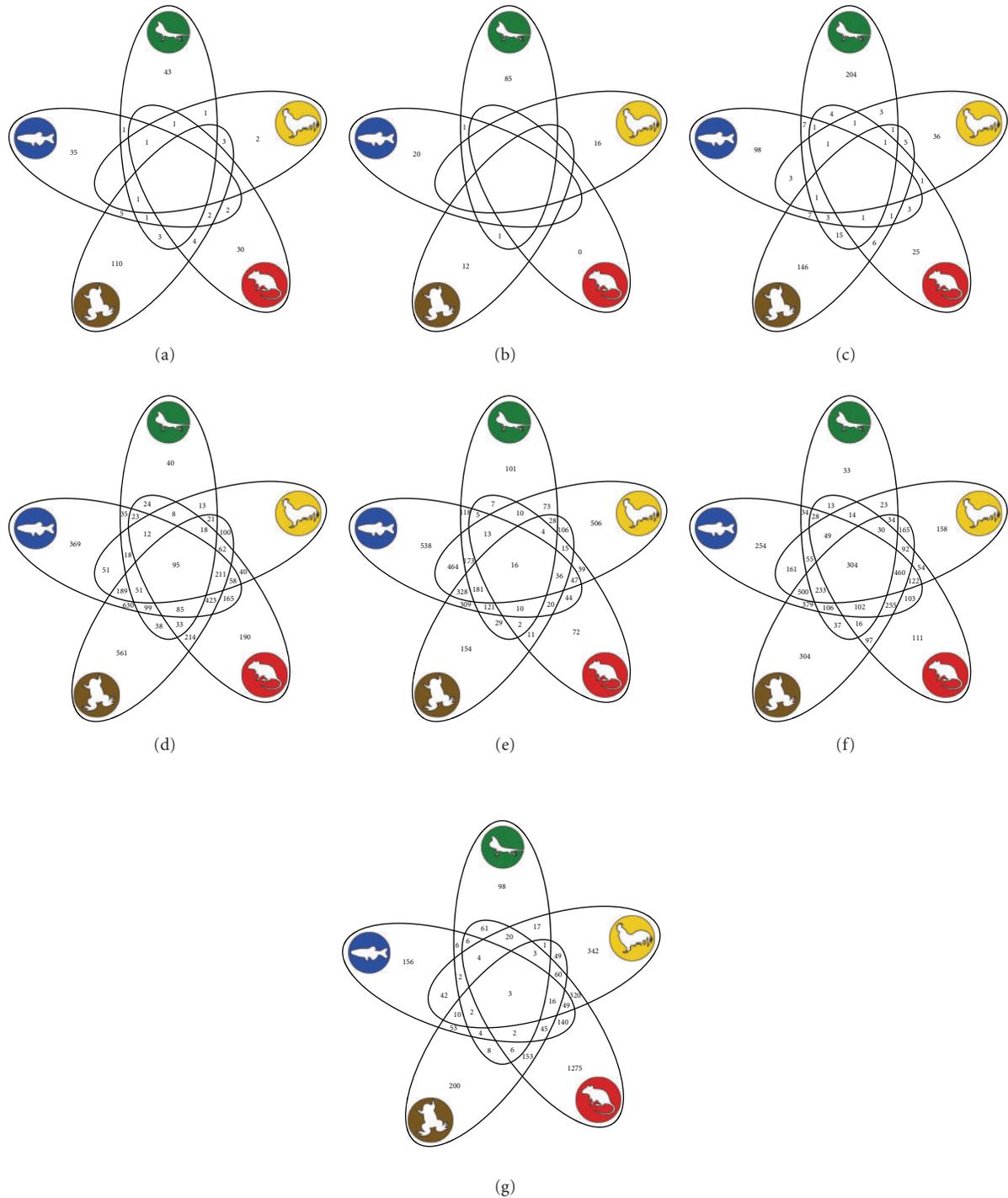


FIGURE 2: Venn diagrams of common functionally conserved genes across all five vertebrate species. For each of the seven functional classes, the number of genes found in all combination of species intersections and exclusions are listed. (a) testis-specific, (b) ovary-specific, (c) gonadal-specific, (d) testis-expressed, (e) ovary-expressed, (f) gonadal-expressed, (g) non-reproductive.

that do not possess paralogs in any of the five lineages. This study focuses on 1:1:1:1:1 orthologs and ignores complications arising from neo-/subfunctionalization caused by gene duplication events [21, 22], particularly those found in the zebrafish lineage after an ancient duplication event [23].

We used the extensive EST libraries publicly available for each species in order to categorize genes into functional classes. Our objective was to generate a standardized sample of genes in each of the reproductive and non-reproductive functional classes, for each species. Historically, EST libraries

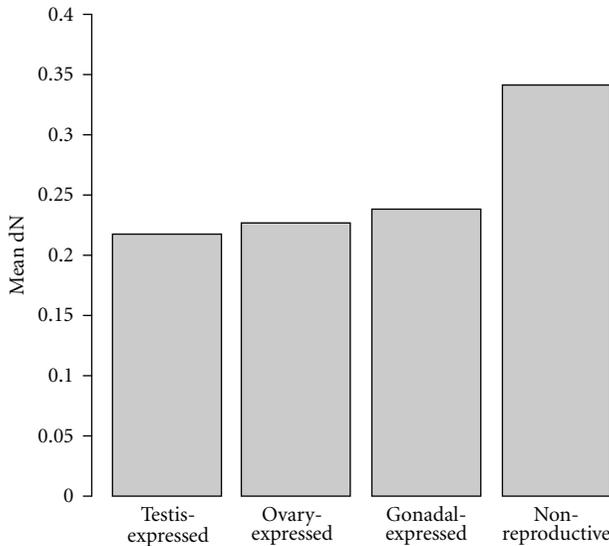


FIGURE 3: dN among functionally conserved classes across all five vertebrate species. Only four of the seven functional classes contained genes that were found in the same functional class across zebrafish, *Anolis*, *Xenopus*, chicken, and mouse. Functional classes were not significantly different from each other.

were originally developed to assist in the genome annotation process (e.g., [24]). The quantity, quality, specificity, and tissue-diversity of EST libraries vary considerably across species (see Supplementary Table 1) and are largely a function of each research community's priorities and preferences for each of the five sequenced genomes. Since our principal objective is to compare reproductive versus non-reproductive levels of molecular divergence in vertebrates, we sought to generate pooled gene samples derived from testis and ovary (i.e., reproductive) and non-reproductive tissue (any adult tissue that does not contain a sex-specific organ or tissue). In addition, genes from tissue-specific (or tissue-limited) classes were differentiated from "tissue-expressed" genes that are expressed more ubiquitously. This approach enables us to compare functional gene classes using relatively large sample sizes and ample statistical power.

A total of seven functional classes were assigned to genes in each of the five species (see Section 2). Table 1 summarizes the number of genes that are contained in each functional class for each species. It is important to note that the proportion of reproductive (e.g., testis, ovary) to non-reproductive genes in each species is not necessarily indicative of the total fraction of reproductive genes found in each genome but, again, reflects each community's specialized interests in generating certain libraries. In addition, overall EST library coverage can be different by an order of magnitude. For example, at last count, the mouse has nearly 5 million ESTs deposited in dbEST, while the green lizard has only 150,000 ESTs. The broader EST coverage in mouse may explain why our screen failed to identify any ovary-specific genes in this taxon. In contrast, since the anoles EST set includes only three non-reproductive tissues at a lower coverage than other species, this may also explain the relatively high number

of ovary-specific genes in this species. With such large differences in EST coverage in each of the five species, it is important to understand the limits of these analyses.

Overall, our results provide evidence of a general pattern of rapid reproductive change over deep vertebrate lineages. Each of the five vertebrates demonstrate significantly higher protein divergence in gonadal genes compared to non-reproductive genes (Figure 1). Rapidly evolving testis genes appear to be driving much of the pattern of higher gonadal-specific gene divergence in these lineages: four of the five taxa—zebrafish, *Xenopus*, chicken, and mouse—all share significantly higher testis-specific divergence. Interestingly, these three taxa include two of the more basal taxa, *Xenopus* and zebrafish (Supplementary Figure 1), supporting that this pattern spans broad phylogenetic groups across the vertebrate subphylum. In green lizards, we observe a contrasting pattern of gonadal divergence as ovary-specific genes appear to be driving the significantly higher divergence of gonadal genes (but see caveat above). Thus, while we see a general pattern of significantly higher divergence among reproduction-specific genes across all vertebrate lineages, there may be large differences in the subset of reproductive genes that are diverging.

In *Drosophila*, we also see a similar pattern of rapidly evolving gonadal genes from EST libraries. Reproductive genes from the testis and ovary and non-reproductive genes from the brain have been used to characterize sexually dimorphic expression patterns [25–27] as well as to compare the evolution of reproductive genes relative to non-reproductive genes [28–30]. A recent study using 12 genomes in *Drosophila* and an extensive EST set from *D. melanogaster* also found that rates of evolution among testis-expressed genes are significantly higher than genes expressed in the ovary or head [12]. A number of studies in mammals have also demonstrated a similar pattern of higher divergence rates in male reproductive genes [11, 31–33].

This higher divergence of reproductive genes, and in particular, male-specific proteins, supports the hypothesis that sexual selection may be an important driver of evolutionary change and extends sexual selection theory to the level of molecules such as those found in gametogenesis and fertilization [34–36]. The strength of this molecular signature indicates the pervasive and cryptic nature of this process: much of this pattern would remain hidden without a comparative and systematic treatment of genome-wide sequence data. We also note that reproductive proteins, particularly those regulating sperm development, are of particular interest to researchers studying mechanisms of reproductive isolation because hybrid male sterility may be the product of the rapid evolution of male reproductive genes: spermatogenesis appears to be a selected target of hybrid male fertility breakdown [37–41]. In addition, there is mounting evidence that positive selection drives the evolution of genes controlling key transitions during both spermatogenesis and oogenesis [42, 43].

Other functional classes of testis-associated genes have also been found in *Drosophila*. Genes encoding proteins secreted by male accessory glands (Acps), the ejaculatory

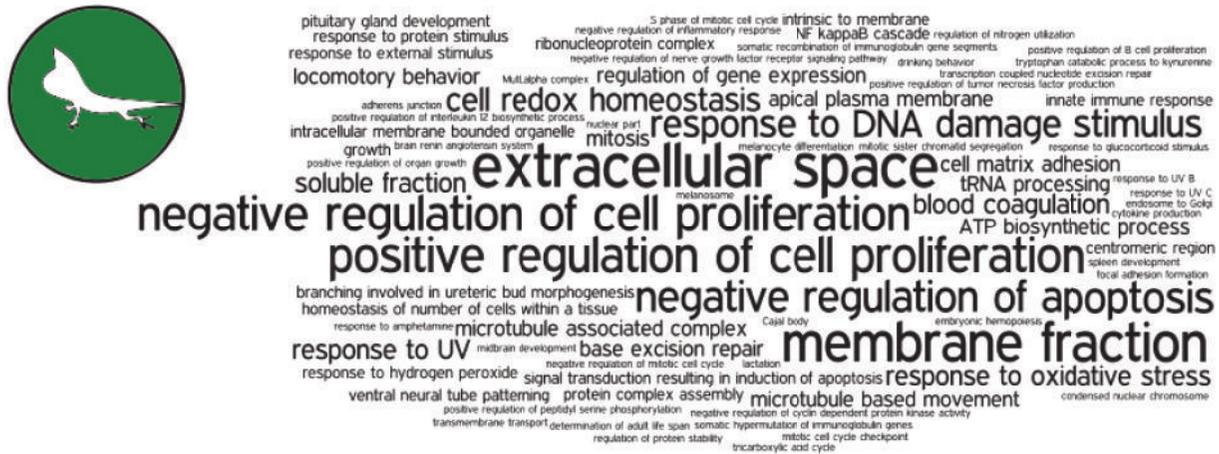


FIGURE 4: Word-size frequency distribution of Gene Ontology (GO) terms for the most diverged orthologs in *A. carolinensis*. Associated GO terms for the top 10% diverged ortholog subset are displayed according to size, based on the frequency of that term. GO terms from Biological Process (BP) and Cellular Component (CC) were used. Similar GO-based word-size frequencies based on 10% most diverged orthologs from *M. musculus*, *G. gallus*, and *D. rerio* are found in Supplementary Figure 2.

duct, and the ejaculatory bulb, as well as many components of *D. melanogaster* seminal fluid, are known to be rapidly evolving. These proteins are transferred from the male to the female along with sperm during mating and mediate a series of postmating events [44–46]. Furthermore, there is ample evidence of adaptive evolution at several loci that encode *D. melanogaster* seminal fluid proteins [47–52]. Whether a similar signal among secretory reproductive classes is found in vertebrate lineages is an intriguing question.

While a clear pattern of rapid testis-specific divergence emerges from our lineage-specific annotations, we then asked whether the same evolutionary pattern holds across genes that have maintained a similar function across all five vertebrate species. In other words, what are the relative rates of evolutionary change across *functionally conserved* classes? Surprisingly, the numbers of genes per class were drastically reduced to the point that only four classes—testis-expressed, ovary-expressed, gonadal-expressed, and non-reproductive genes—share genes in common across all five species (Figure 2). Furthermore, a decreasing but nonsignificant trend of evolutionary rates was found among these four functional classes:  $dN_{\text{non-reproductive}} > dN_{\text{gonad-expressed}} > dN_{\text{ovary-expressed}} > dN_{\text{testis-expressed}}$ . Overall, this functionally conserved group describes a subset of the data with a contrasting evolutionary pattern, thereby demonstrating that testis-specific genes are affected by a variety of evolutionary forces. In a recent study, Dean et al. [33] performed a genomic and proteomic study on six tissue types from the male reproductive tract of mouse (excluding testis) and found that one tract, the seminal vesicle, had significantly higher rates of divergence while the other five tracts showed significantly lower rates of divergence when compared to other proteins. Our results demonstrate a similar high variance of evolutionary rates within the testis.

*A. carolinensis* was the outlier of the five vertebrate taxa with a significantly higher divergence among ovary-specific genes. Ovaries have also been shown to be sites of rapid divergence in *D. melanogaster* as part of a molecular coevolutionary process between sperm and egg. A number of rapidly evolving genes have been found expressed in the female reproductive tract and potentially secreted [53] or induced in the female reproductive tract by mating [54–56]. Further characterization of the green anole genome, in addition to other Lepidosauria genomes and genomic resources that will soon be available, will allow us to address whether female lizards are indeed driving sexual selective processes and whether this is a common lineage-specific process among squamate reptiles or simply an artifact of EST functional annotation.

While aligning orthologs to ESTs offers a powerful approach for functional annotation, it is important to procure a more granular understanding of process, function and localization. Therefore, we took the 10% most diverged orthologs and associated each species' corresponding gene to its Gene Ontology (GO), namely, Biological Process (BP) and Cellular Component (CC). A word cloud in which the font size is a function of the frequency of statistically over-represented functional phrases in diverged orthologs is shown for *A. carolinensis* in Figure 4. GO-associated word frequencies for fish, mouse, and chicken are found in Supplementary Figure 2. The density of each word cloud for each species reflects the amount of curation effort in Gene Ontology within these species' communities. As expected, we don't see much overlap between the GO and EST approaches to functional annotation. Reproductive function is a poorly annotated ontological class harboring a level of characterization that will not substantially improve until

more geneticists and molecular biologists study reproductive loci in greater detail.

#### 4. Concluding Remarks

Patterns parsed from extant genomes can inform us about the underlying evolutionary processes that have acted upon lineages in the past. As a functional class, reproductive-specific genes are more rapidly evolving than other functional gene classes, and it appears that testis genes are driving this pattern of divergence in the majority of vertebrate lineages. This work sets the stage for a more nuanced analysis of divergence leveraged against function across diverse taxa. With more genomes and ESTs generated, greater effort can be afforded to better estimate the probability that a gene is a member of a particular functional class, even when the number of ESTs and libraries are quite different between species. Newer data types such as RNAseq will certainly help solve the sampling bias problem with better coverage and more tissues sampled. Future studies that include paralogous sequences to evaluate birth/death processes and *de novo* gene functionalization models (including incorporating the large number of paralogs from zebrafish) in the context of functional class will also be useful in addressing the role of reproductive genes in vertebrate evolution.

It is remarkable that across very distant phylogenetic lineages, we detect the same evolutionary patterns found among closely related species: high lineage-specific reproductive diversity and, in particular, a high variance in male reproductive characters. These parallel patterns support the contention that sexual selection on both morphological and molecular characters may be an important, common, and pervasive feature of vertebrate evolution.

#### Abbreviations

GO: Gene ontology  
 EST: Expressed sequence tag  
 CDS: Coding sequence  
 BP: Biological process  
 CC: Cellular component  
 dN: Nonsynonymous substitutions per nonsynonymous site.

#### Acknowledgments

The authors would like to thank Dr. Tonia Hsieh (Temple University) for guidance and support during all stages of this project. They also thank Dr. Ed Braun (University of Florida) for technical advice and the Beaty Biodiversity Research Centre at the University of British Columbia for use of their SciBorg cluster. C. J. Grassa dedicates his contributions to Thomas and Sarah Grassa.

#### References

[1] C. Linnaeus, *Systema Naturae*, Nieuwkoop, The Netherlands, 1735, reprinted in 1964 by B. de Graaf.

- [2] C. R. Darwin, *The Descent of Man, and Selection in Relation to Sex*, John Murray, London, UK, 1871.
- [3] W. G. Eberhard, *Sexual Selection and Animal Genitalia*, Harvard University Press, Cambridge, Mass, USA, 1985.
- [4] W. Salzburger and A. Meyer, "The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics," *Naturwissenschaften*, vol. 91, no. 6, pp. 277–290, 2004.
- [5] K. E. Boul, W. C. Funk, C. R. Darst, D. C. Cannatella, and M. J. Ryan, "Sexual selection drives speciation in an Amazonian frog," *Proceedings of the Royal Society B*, vol. 274, no. 1608, pp. 399–406, 2007.
- [6] A. H. Harcourt and J. Gardiner, "Sexual selection and genital anatomy of male primates," *Proceedings of the Royal Society B*, vol. 255, no. 1342, pp. 47–53, 1994.
- [7] J. Roux and M. Robinson-Rechavi, "Developmental constraints on vertebrate genome evolution," *PLoS Genetics*, vol. 4, no. 12, Article ID e1000311, 2008.
- [8] B. Korber, M. LaBute, and K. Yusim, "Immunoinformatics comes of age," *PLoS Computational Biology*, vol. 2, no. 6, article e71, 2006.
- [9] W. J. Swanson and V. D. Vacquier, "The rapid evolution of reproductive proteins," *Nature Reviews Genetics*, vol. 3, no. 2, pp. 137–144, 2002.
- [10] T. R. Shirangi, H. D. Dufour, T. M. Williams, and S. B. Carroll, "Rapid evolution of sex pheromone-producing enzyme expression in *Drosophila*," *PLoS Biology*, vol. 7, no. 8, Article ID e1000168, 2009.
- [11] C. I. Castillo-Davis, F. A. Kondrashov, D. L. Hartl, and R. J. Kulathinal, "The functional genomic distribution of protein divergence in two animal phyla: coevolution, genomic conflict, and constraint," *Genome Research*, vol. 14, no. 5, pp. 802–811, 2004.
- [12] W. Haerty, S. Jagadeeshan, R. J. Kulathinal et al., "Evolution in the fast lane: rapidly evolving sex-related genes in *Drosophila*," *Genetics*, vol. 177, no. 3, pp. 1321–1335, 2007.
- [13] R. C. Edgar, "MUSCLE: multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Research*, vol. 32, no. 5, pp. 1792–1797, 2004.
- [14] J. Felsenstein, *PHYLIP (Phylogeny Inference Package) Version 3.6*, Department of Genome Sciences, University of Washington, Seattle, Wash, USA, 2005.
- [15] M. Nei and T. Gojobori, "Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions," *Molecular Biology and Evolution*, vol. 3, no. 5, pp. 418–426, 1986.
- [16] B. Korber, "HIV signature and sequence variation analysis," in *Computational Analysis of HIV Molecular Sequences*, A. G. Rodrigo and G. H. Learn, Eds., chapter 4, pp. 55–72, Kluwer Academic, Dodrecht, The Netherlands, 2000.
- [17] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology*, vol. 215, no. 3, pp. 403–410, 1990.
- [18] D. F. Bauer, "Constructing confidence sets using rank statistics," *Journal of the American Statistical Association*, vol. 67, pp. 687–690, 1972.
- [19] C. I. Castillo-Davis and D. L. Hartl, "GeneMerge—post-genomic analysis, data mining, and hypothesis testing," *Bioinformatics*, vol. 19, no. 7, pp. 891–892, 2003.
- [20] S. B. Hedges, "Molecular evidence for the origin of birds," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, no. 7, pp. 2621–2624, 1994.

- [21] M. Lynch and A. Force, "The probability of duplicate gene preservation by subfunctionalization," *Genetics*, vol. 154, no. 1, pp. 459–473, 2000.
- [22] M. Lynch and J. S. Conery, "The evolutionary fate and consequences of duplicate genes," *Science*, vol. 290, no. 5494, pp. 1151–1155, 2000.
- [23] J. S. Taylor, Y. Van de Peer, I. Braasch, and A. Meyer, "Comparative genomics provides evidence for an ancient genome duplication event in fish," *Philosophical Transactions of the Royal Society B*, vol. 356, no. 1414, pp. 1661–1679, 2001.
- [24] M. G. Reese, G. Hartzell, N. L. Harris, U. Ohler, J. F. Abril, and S. E. Lewis, "Genome annotation assessment in *Drosophila melanogaster*," *Genome Research*, vol. 10, no. 4, pp. 483–501, 2000.
- [25] J. Andrews, G. G. Bouffard, C. Cheadle, J. Lü, K. G. Becker, and B. Oliver, "Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis," *Genome Research*, vol. 10, no. 12, pp. 2030–2043, 2000.
- [26] M. Parisi, R. Nuttall, P. Edwards et al., "A survey of ovary-, testis-, and soma-biased gene expression in *Drosophila melanogaster* adults," *Genome Biology*, vol. 5, no. 6, article R40, 2004.
- [27] R. S. Singh and R. J. Kulathinal, "Male sex drive and the masculinization of the genome," *BioEssays*, vol. 27, no. 5, pp. 518–525, 2005.
- [28] M. B. Coulthart and R. S. Singh, "High level of divergence of male-reproductive-tract proteins, between *Drosophila melanogaster* and its sibling species, *D. simulans*," *Molecular Biology and Evolution*, vol. 5, no. 2, pp. 182–191, 1988.
- [29] A. Civetta and R. S. Singh, "High divergence of reproductive tract proteins and their association with postzygotic reproductive isolation in *Drosophila melanogaster* and *Drosophila virilis* group species," *Journal of Molecular Evolution*, vol. 41, no. 6, pp. 1085–1095, 1995.
- [30] S. Jagadeeshan and R. S. Singh, "Rapidly evolving genes of *Drosophila*: differing levels of selective pressure in testis, ovary, and head tissues between sibling species," *Molecular Biology and Evolution*, vol. 22, no. 9, pp. 1793–1801, 2005.
- [31] J. M. Good and M. W. Nachman, "Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis," *Molecular Biology and Evolution*, vol. 22, no. 4, pp. 1044–1052, 2005.
- [32] L. M. Turner, E. B. Chuong, and H. E. Hoekstra, "Comparative analysis of testis protein evolution in rodents," *Genetics*, vol. 179, no. 4, pp. 2075–2089, 2008.
- [33] M. D. Dean, N. L. Clark, G. D. Findlay et al., "Proteomics and comparative genomic investigations reveal heterogeneity in evolutionary rate of male reproductive proteins in mice (*Mus domesticus*)," *Molecular Biology and Evolution*, vol. 26, no. 8, pp. 1733–1743, 2009.
- [34] A. Civetta and R. S. Singh, "Sex-related genes, directional sexual selection, and speciation," *Molecular Biology and Evolution*, vol. 15, no. 7, pp. 901–909, 1998.
- [35] R. S. Singh and R. J. Kulathinal, "Sex gene pool evolution and speciation: a new paradigm," *Genes and Genetic Systems*, vol. 75, no. 3, pp. 119–130, 2000.
- [36] R. J. Kulathinal and R. S. Singh, "The nature of genetic variation in sex and reproduction-related genes among sibling species of the *Drosophila melanogaster* complex," *Genetica*, vol. 120, no. 1–3, pp. 245–252, 2004.
- [37] R. J. Kulathinal and R. S. Singh, "Cytological characterization of premeiotic versus postmeiotic defects producing hybrid male sterility among sibling species of the *Drosophila melanogaster* complex," *Evolution*, vol. 52, no. 4, pp. 1067–1079, 1998.
- [38] P. Michalak and M. A. F. Noor, "Genome-wide patterns of expression in *Drosophila* pure species and hybrid males," *Molecular Biology and Evolution*, vol. 20, no. 7, pp. 1070–1076, 2003.
- [39] W. Haerty and R. S. Singh, "Gene regulation divergence is a major contributor to the evolution of Dobzhansky-Muller incompatibilities between species of *Drosophila*," *Molecular Biology and Evolution*, vol. 23, no. 9, pp. 1707–1714, 2006.
- [40] R. J. Kulathinal and R. S. Singh, "The molecular basis of speciation: from patterns to processes, rules to mechanisms," *Journal of Genetics*, vol. 87, no. 4, pp. 327–338, 2008.
- [41] D. C. Presgraves, L. Balagopalan, S. M. Abmayr, and H. A. Orr, "Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*," *Nature*, vol. 423, no. 6941, pp. 715–719, 2003.
- [42] A. Civetta, S. A. Rajakumar, B. Brouwers, and J. P. Bacik, "Rapid evolution and gene-specific patterns of selection for three genes of spermatogenesis in *Drosophila*," *Molecular Biology and Evolution*, vol. 23, no. 3, pp. 655–662, 2006.
- [43] V. L. Bauer DuMont, H. A. Flores, M. H. Wright, and C. F. Aquadro, "Recurrent positive selection at *Bgcn*, a key determinant of germ line differentiation, does not appear to be driven by simple coevolution with its partner protein bam," *Molecular Biology and Evolution*, vol. 24, no. 1, pp. 182–191, 2007.
- [44] T. Chapman and S. J. Davies, "Functions and analysis of the seminal fluid proteins of male *Drosophila melanogaster* fruit flies," *Peptides*, vol. 25, no. 9, pp. 1477–1490, 2004.
- [45] A. Wong and M. F. Wolfner, "Sexual behavior: a seminal peptide stimulates appetites," *Current Biology*, vol. 16, no. 7, pp. R256–R257, 2006.
- [46] K. R. Ram and M. F. Wolfner, "Sustained post-mating response in *Drosophila melanogaster* requires multiple seminal fluid proteins," *PLoS Genetics*, vol. 3, no. 12, article e23, 2007.
- [47] M. Aguadé, N. Miyashita, and C. H. Langley, "Polymorphism and divergence in the Mst26A male accessory gland gene region in *Drosophila*," *Genetics*, vol. 132, no. 3, pp. 755–770, 1992.
- [48] S. C. Tsaur, C. T. Ting, and C. I. Wu, "Positive selection driving the evolution of a gene of male reproduction, Acp26Aa, of *Drosophila*: II. Divergence versus polymorphism," *Molecular Biology and Evolution*, vol. 15, no. 8, pp. 1040–1046, 1998.
- [49] M. Aguadé, "Positive selection drives the evolution of the Acp29AB accessory gland protein in *Drosophila*," *Genetics*, vol. 152, no. 2, pp. 543–551, 1999.
- [50] D. J. Begun, P. Whitley, B. L. Todd, H. M. Waldrip-Dail, and A. G. Clark, "Molecular population genetics of male accessory gland proteins in *Drosophila*," *Genetics*, vol. 156, no. 4, pp. 1879–1888, 2000.
- [51] W. J. Swanson, A. G. Clark, H. M. Waldrip-Dail, M. F. Wolfner, and C. F. Aquadro, "Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 13, pp. 7375–7379, 2001.
- [52] J. L. Mueller, K. Ravi Ram, L. A. McGraw et al., "Cross-species comparison of *Drosophila* male accessory gland protein genes," *Genetics*, vol. 171, no. 1, pp. 131–143, 2005.
- [53] T. M. Panhuis and W. J. Swanson, "Molecular evolution and population genetic analysis of candidate female reproductive genes in *Drosophila*," *Genetics*, vol. 173, no. 4, pp. 2039–2047, 2006.

- [54] M. K. N. Lawniczak and D. J. Begun, "A genome-wide analysis of courting and mating responses in *Drosophila melanogaster* females," *Genome*, vol. 47, no. 5, pp. 900–910, 2004.
- [55] M. K. N. Lawniczak and D. J. Begun, "Molecular population genetics of female-expressed mating-induced serine proteases in *Drosophila melanogaster*," *Molecular Biology and Evolution*, vol. 24, no. 9, pp. 1944–1951, 2007.
- [56] L. A. McGraw, G. Gibson, A. G. Clark, and M. F. Wolfner, "Genes regulated by mating, sperm, or seminal proteins in mated female *Drosophila melanogaster*," *Current Biology*, vol. 14, no. 16, pp. 1509–1514, 2004.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

