Selective Constraint Dominates the Evolution of Genes Expressed in a Novel Reproductive Gland

Findley R. Finseth,*[1,2] Eliana Bondra,[1,3] and Richard G. Harrison[1]

[1]Department of Ecology and Evolutionary Biology, Cornell University

[2]Division of Biological Sciences, University of Montana, Missoula

[3]Department of Biomolecular Chemistry, University of Wisconsin, Madison

***Corresponding author:** E-mail: far25@cornell.edu.

Associate editor: Willie Swanson

Abstract

One striking pattern in molecular evolution is that genes encoding proteins involved in reproduction tend to evolve rapidly. Seminal fluid proteins frequently exhibit this pattern and directly affect multiple reproductive processes including enhancing sperm performance and mediating postmating sexual selection. Here, we investigate molecular evolutionary patterns of genes expressed in the foam gland of Japanese quail (*Coturnix japonica*), a novel reproductive phenotype. Foam provides an interesting contrast to seminal fluid because it plays a similar functional role, yet is produced, stored, and transferred to females independent of semen. We combined RNA-Seq and comparative genomics to examine evolutionary rates of genes with enriched expression in the foam gland of Japanese quail and those that exhibit enriched expression in two other tissues (testis and liver) and with broadly expressed genes. Overall, we found pronounced heterogeneity in evolutionary rates. Foam gland genes evolved under strong evolutionary constraint, whereas testis genes evolved rapidly and sometimes adaptively. These striking differences were robust to variation in gene expression. Genes with enriched expression in the foam gland did not show major shifts in selective pressure after the quail and chicken lineages split; in contrast, testis-expressed genes experienced a burst of accelerated evolution specifically along the *Coturnix* lineage. Our work demonstrates that, as a class, genes expressed in the novel foam gland experience different selection regimes than genes expressed in many other tissues producing seminal fluid proteins. Our results also highlight the importance of selective constraint in shaping the evolution of male reproductive genes.

Key words: **sperm competition, molecular evolution, seminal fluid, male reproduction, RNA-Seq, comparative genomics.**

Article

Introduction

Across diverse taxa, genes involved in sexual reproduction are often targets of positive selection, producing a pattern of rapid evolutionary change compared with nonreproductive genes (Swanson and Vacquier 2002a, 2002b; Clark et al. 2006; Panhuis et al. 2006; Turner and Hoekstra 2008). In internally fertilizing species, the complex mixture of proteins transferred along with sperm to the female exhibit, on average, particularly rapid and adaptive divergence (e.g., Swanson et al. 2001; Clark and Swanson 2005; Andrés et al. 2006; Findlay et al. 2008; Dean et al. 2009; Walters and Harrison 2010). These seminal fluid proteins have profound effects on fitness through mediating sperm performance, influencing the outcome of postmating sexual selection, and modulating female reproductive physiology (Poiani 2006; Pitnick et al. 2008; Sirot et al. 2009). In addition to being integral to reproductive fitness, seminal fluid proteins may also play a role in the speciation process, as their rapid divergence could contribute to the formation of barriers to gene flow (e.g., Andrés et al. 2008; Dean and Nachman 2009).

What selective forces drive the pattern of relatively rapid divergence of seminal fluid proteins? Positive selection due to repeated functional turnover in response to postmating sexual selection of some form (sperm competition, sexual conflict, and/or cryptic female choice) is frequently cited as responsible for this phenomenon (Swanson and Vacquier 2002a; Turner and Hoekstra 2008; Wong 2011). Although the contribution of postmating sexual selection to the divergence of a wide range of reproductive phenotypes is clear (e.g., Hosken 1998; Anderson and Dixson 2002; Ramm and Stockley 2010), empirical evidence for its role in the evolution of reproductive proteins is mixed (reviewed in Wong 2011). Investigations of single reproductive genes find some, though surprisingly few, significant associations between intensity of postmating sexual selection and the rate of molecular evolution (Dorus et al. 2004; Herlyn and Zischler 2007; Hurle et al. 2007; Ramm et al. 2008; Martin-Coello et al. 2009; O'Connor and Mundy 2009; Finn and Civetta 2010). Studies adopting a multigene approach generally confirm the prediction that average evolutionary rates of reproductive genes are higher in taxa with higher levels of postmating sexual selection (Wagstaff 2005; Kelleher et al. 2007; Almeida and DeSalle 2008; Wong 2010; Schumacher et al. 2013; but see Walters and Harrison 2011; Carnahan-Craig and Jensen-Seaman 2013; Good et al. 2013). In addition to adaptive evolution in response to sexual selection, relaxed constraint, host-pathogen dynamics, gene duplication, and reinforcement during speciation have all been suggested to contribute to increased divergence of reproductive proteins (reviewed in Swanson and Vacquier 2002a).

© The Author 2014. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

Although recent studies often emphasize the high incidence of positive selection acting on seminal fluid proteins, many seminal fluid proteins are actually quite conserved. Indeed, substantial heterogeneity in evolutionary rates occurs across groups of reproductive genes expressed in particular tissues (Dean et al. 2009; Grassa and Kulathinal 2011; Arunkumar et al. 2013), at different times during development (Good and Nachman 2005), with varying degrees of tissue specificity (Good and Nachman 2005; Dean et al. 2008; Grassa and Kulathinal 2011; Parsch and Ellegren 2013) or species specificity (Marshall et al. 2010; Grassa and Kulathinal 2011), and from different functional classes (Dorus et al. 2006; Turner et al. 2008; Dorus et al. 2010; Carnahan-Craig and Jensen-Seaman 2013; Good et al. 2013). One striking example exists in mice, where the male reproductive tract is highly compartmentalized. Here, tissues from different compartments that contribute material to seminal fluid reveal marked heterogeneity in evolutionary rates; genes with enriched expression in tissues likely involved in postmating sexual selection (e.g., testis, seminal vesicles) evolve rapidly, whereas genes from other tissues evolve much more slowly than the genomic average, revealing an unappreciated degree of selective constraint acting on male reproductive genes (Dean et al. 2009). Although much research on male reproductive genes has focused on a single tissue (accessory glands) in insects, the mouse example highlights how novel insights can be gained by expanding the field to focus on more diverse taxa and tissues.

One tissue with the potential to contribute to our understanding of the molecular evolution of male reproductive genes is the foam gland of male Japanese quail. Males of the genus *Coturnix* are unique among birds in possessing a well-developed, sexually dimorphic foam gland (also known as the "proctodeal gland" or "cloacal gland"; Klemm et al. 1973). This large, red, external protuberance is interdigitated with the cloacal sphincter muscle and lies dorsal to the cloaca in sexually mature males (McFarland et al. 1968; Klemm et al. 1973). The gland secretes a viscous mixture of proteins that is whipped into an airy, meringue-like foam by rhythmic motions of the cloacal muscle, the frequency of which increase upon detection of a female (Klemm et al. 1973; Seiwert and Adkins-Regan 1998). Females possess rudimentary foam glands, but these have limited secretory activity and do not make foam (McFarland et al. 1968; Klemm et al. 1973; King 1981). During copulation, males deposit semen along with a large quantity of foam in the female reproductive tract (Klemm et al. 1973). The presence of foam improves fertilization efficiency at certain stages in a female ovulatory cycle (Cheng, Hickman, et al. 1989; Adkins-Regan 1999), extends the window of time that a male can achieve fertilization following a single insemination (Cheng, Hickman, et al. 1989; Singh et al. 2012), and increases sperm motility, viability, and transport in the oviduct (Cheng, McIntyre, et al. 1989; Singh, Sastry, Shit, et al. 2011; Singh et al. 2012). Foam also mediates the outcome of sperm competition by conferring fertility benefits to a focal male's sperm at a cost to a rival's fertility (Cheng, Hickman, et al. 1989; Adkins-Regan 1999; Finseth et al. 2013).

As a male reproductive tissue, the foam gland may offer unique insights because it is not strictly a seminal fluid per se and offers an intriguing contrast to tissues that produce traditional seminal fluid proteins. In Japanese quail males, seminal fluid is manufactured by the seminiferous tubules, testis epithelia, or ductus deferens, which comprise an entirely separate tubule system from the foam–foam gland complex (Lake 1981). Unlike seminal fluid proteins, foam is stored separately from sperm, never packaged together with sperm or seminal fluid inside of males, and not mixed with an ejaculate until inside the female reproductive tract. The foam–foam gland complex also represents a novel phenotype. *Coturnix* males are the only avian lineage in which the dorsal proctodeal gland is reddened and swollen, noticeably protrudes from the cloacal region, and secretes a copulatory fluid with a similar foamy quality (Klemm et al. 1973; King 1981; Fujihara 1992). Additionally, in contrast to seminal fluid, foam is not required for fertilization (Ikeda and Taji 1954; Marks and Lepore 1965; King 1981). Yet, like seminal fluid proteins, foam is deposited in the female reproductive tract during copulation, enhances sperm function, improves fertility outcomes, and plays a role in postmating sexual selection. Thus, the foam–foam gland complex presents an opportunity to address the evolution of male reproductive genes that play a role similar to that of seminal fluid proteins, but is distinctly different in terms of production and derivation.

In this study, we explore the hypothesis that foam gland genes are targets of rapid adaptive divergence, as observed in many genes that encode seminal fluid proteins. We combine an RNA-Seq and comparative genomics approach to compare patterns of molecular evolution for genes expressed in the foam gland with patterns for genes expressed in a nonreproductive, glandular tissue (liver) or a second reproductive tissue (testis), as well as with genes expressed in all three tissues in male Japanese quail (nonspecific). Because genes with high levels of expression in the foam gland may have acquired novel, foam-producing roles in the quail, we used a phylogenetic framework to identify shifts in evolutionary rate along the quail lineage after splitting with the chicken. We report pronounced heterogeneity in evolutionary rate across male reproductive tissues, with foam gland genes showing surprisingly slow rates of protein evolution and no major shifts in selective pressures specific to the quail lineage. We also demonstrate that genes with enriched expression in testis evolve rapidly, with a burst of accelerated divergence along the quail lineage, despite physiological polyspermy in birds. These patterns are consistent across comparisons of quail with three other avian species and are robust to differences in gene expression.

Results

Summary of Transcriptome Assembly and RNA-Seq

We sequenced and assembled a Japanese quail transcriptome representing genes expressed in two reproductive tissues (foam gland and testis) and one nonreproductive glandular tissue (liver) from six sexually mature Japanese quail males. The vast majority of the transcriptome was represented by

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

full-length or nearly full-length transcripts as identified by the ortholog hit ratio, or the length of quail transcripts relative to their chicken orthologs (supplementary fig. S1*a*). Raw data and transcriptome summary statistics are reported in supplementary tables S1 and S2, Supplementary Material online. Samples from the three different tissues displayed distinct gene expression profiles (supplementary fig. S1*b*, Supplementary Material online). Tissue explains the major differences in expression profiles as the biological coefficient of variation (or the coefficient of variation with which the abundance of a gene varies between RNA samples [McCarthy et al. 2012]) is large in between-tissue comparisons, but small within tissue (supplementary fig. S1*b*, Supplementary Material online).

We used an RNA-Seq approach to identify genes that were either significantly enriched in expression in a single tissue (foam gland, testis, and liver; "tissue-enriched") or were only expressed in one tissue ("tissue-restricted"). For each category, we also identified a panel of "nonspecific" genes that were broadly expressed in all three tissues and, for the tissue-enriched classification, additionally required that nonspecific genes were not enriched in any tissue. To confirm that tissue-enriched patterns were not being driven by lowly expressed genes, we also obtained a conservative gene list that used more stringent abundance and differential expression criteria to classify a gene as being enriched in a tissue.

Results were similar whether we considered genes that were tissue-enriched or tissue-restricted, so we focus on patterns of genes enriched in expression and report methods and results for genes with tissue-restricted expression in the supplementary materials (supplementary methods S1 and supplementary figs. S2–S6, Supplementary Material online). The distribution of genes with enriched expression in a particular tissue is shown in figure 1. In all cases, $>99.5\%$ of genes identified as tissue-restricted in a given tissue were also enriched in that same tissue (supplementary fig. S2, Supplementary Material online). The few cases for which a gene was found to be "restricted" but not "enriched" in

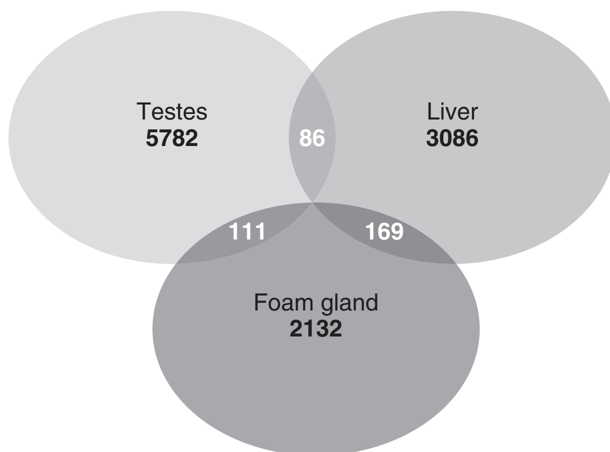

Fig. 1. Venn diagram depicting the number of genes enriched by at least log 2 fold in one tissue versus the other two. There were 2100 genes that were expressed, but not enriched, in any tissue; these were considered "nonspecific." Genes with enrichment in two tissues (white) were removed from the "tissue-enriched" data set.

expression in a given tissue were because low expression levels gave little power to assign significant enrichment. Tissues varied in the number of genes with enriched expression, with testis having the most and the foam gland the least (fig. 1 and supplementary fig. S2, Supplementary Material online). These differences were not due to differences in library size as our analyses accounted for that source of variation. For the subset of genes that we tested, quantitative Polymerase Chain Reaction (qPCR) data confirmed significant upregulation in the expected tissue (supplementary S3, Supplementary Material online).

Pairwise Estimation of Evolutionary Rates

We used our transcriptome to identify 1:1 orthologs between quail and either chicken, turkey, or zebra finch, calculated evolutionary rates (d_N/d_S ratios) for each quail-species ortholog pair, and compared average rates among tissue classes. Genes with enriched expression in the foam gland evolved more slowly, whereas genes with enriched expression in testis evolved more rapidly, than genes from other single tissues (nonoverlapping confidence interval [CIs]; table 1 and supplementary table S4, Supplementary Material online). These trends were consistent across all three species pairs, with the exception that foam gland-enriched genes evolved at the same rate as liver-enriched genes in the *Coturnix–Gallus* comparison. Examining the list of genes assigned using more stringent requirements for determining enrichment confirmed our findings, as major trends are the same (supplementary table S5, Supplementary Material online). In the *Coturnix–Gallus* and *Coturnix–Meleagris* comparisons, genes more narrowly expressed in the testis (i.e., testis-restricted) averaged faster rates of protein evolution (ω) than those only enriched in the testis (table 1 and supplementary table S4, Supplementary Material online), suggesting genes specifically expressed in the testis evolve particularly quickly. Evolutionary rate did not differ with expression specificity for genes from the other two tissues or nonspecific genes (table 1 and supplementary table S4, Supplementary Material online). In almost all cases, the testis revealed the highest number of genes exhibiting a signature of positive selection ($\omega > 1$), even when outnumbered by nonspecific genes in the *Coturnix–Meleagris* and *Coturnix–Gallus* comparisons (table 1 and supplementary table S4, Supplementary Material online).

Patterns of variation in ω could be produced by differences in either d_N or d_S. Although we observed significant differences in d_N among tissue classes consistent with trends in ω, d_S was significantly higher in genes enriched in the liver versus at least one other single tissue for all three species comparisons (table 1). To confirm that the observed ω patterns were due to differences in d_N, we focused on the quail–chicken comparison and regressed d_N on d_S, evaluating whether the residuals varied according to tissue class. This analysis suggests that the ω values are robust to differences in d_S, as average residual differences across tissue classes replicate major patterns from pairwise ω estimates (supplementary fig. S4, Supplementary Material online). Moreover, the

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

Table 1. Patterns of Protein Evolution Derived from 1:1 Orthologs for Tissue-Enriched Genes.

Comparison	Tissue	N	d_N (95% CI)[a]	d_S (95% CI)	ω (95% CI)	$\omega > 1$[b]	ω (95% CI) Secreted[c]
Gallus–Coturnix							
	Foam gland	705	0.024 (0.022–0.026)	0.153 (0.147–0.160)	0.154 (0.142–0.167)	2	0.274 (0.220–0.334) (77)
	Testis	1858	0.035 (0.033–0.037)	0.150 (0.147–0.153)	0.226 (0.216–0.236)	16	0.216 (0.154–0.300) (58)
	Liver	848	0.030 (0.028–0.033)	0.166 (0.160–0.172)	0.187 (0.144–0.200)	5	0.262 (0.223–0.309) (131)
	Nonspecific	506	0.024 (0.022–0.027)	0.156 (0.149–0.162)	0.150 (0.137–0.167)	1	0.235 (0.140–0.352) (23)
Meleagris–Coturnix							
	Foam gland	539	0.027 (0.025–0.030)	0.158 (0.153–0.163)	0.171 (0.157–0.184)	2	0.269 (0.221–0.325) (65)
	Testis	1484	0.039 (0.037–0.041)	0.157 (0.154–0.160)	0.242 (0.230–0.253)	13	0.254 (0.172–0.368) (44)
	Liver	665	0.033 (0.030–0.036)	0.168 (0.164–0.173)	0.199 (0.187–0.212)	3	0.244 (0.211–0.280) (112)
	Nonspecific	393	0.026 (0.023–0.029)	0.157 (0.151–0.163)	0.157 (0.141–0.172)	1	0.257 (0.134–0.414) (19)
Taeniopygia–Coturnix							
	Foam gland	585	0.067 (0.062–0.073)	0.511 (0.496–0.530)	0.134 (0.124–0.146)	1	0.176 (0.147–0.210) (64)
	Testis	1579	0.093 (0.088–0.097)	0.510 (0.501–0520)	0.180 (0.172–0.188)	4	0.152 (0.120–0.186) (43)
	Liver	711	0.081 (0.076–0.086)	0.534 (0.521–0.550)	0.155 (0.147–0.164)	0	0.198 (0.175–0.224) (101)
	Nonspecific	418	0.057 (0.052–0.063)	0.505 (0.486–0.523)	0.117 (0.106–0.129)	0	0.148 (0.102–0.196) (21)

[a]All values represent means and 95% CIs. CIs were generated from 10,000 bootstrap resamplings of the mean without assuming normality.
[b]Number of genes with $\omega > 1$.
[c]Secretion status determined by the presence of a predicted signal peptide. These analyses were restricted to contigs with ortholog hit ratios $>= 0.8$. Sample sizes for secretion analysis are given in parentheses.

tissue-restricted comparisons and the list of conservatively assigned enriched genes (i.e., those with higher expression levels and greater differential expression) revealed no differences in d_S across tissues, suggesting that the d_S patterns initially observed are driven by lowly expressed genes that are not specifically expressed in the liver (supplementary tables S4 and S5, Supplementary Material online).

We also explored d_N/d_S ratios of putatively secreted proteins, as identified by the presence of a signal peptide. For tissue-enriched genes, we observed the highest number of tissue-enriched genes with a signal peptide from the liver and foam gland, despite the testis having far more expressed genes than the other tissues (table 1). When comparing the evolutionary rates of secreted proteins across tissues, the striking heterogeneity observed from analyses of the complete gene sets was no longer observed. In nearly all cases, evolutionary rates of genes with signal peptides did not differ across tissue classes within each species comparisons (table 1 and supplementary table S4, Supplementary Material online). The only exceptions were that nonspecific genes evolved slower than genes with expression restricted to the liver for the *Coturnix–Meleagris* and *Coturnix–Taeniopygia* comparisons. Because pairwise patterns of molecular variation were largely consistent across species, we focused all downstream analyses on the *Coturnix–Gallus* comparison, as the chicken is the most closely related species to the Japanese quail and has the best genomic resources and annotation available (Kimball and Braun 2008).

Tissue-Specific Trends in ω after Correcting for Differences in Expression

Expression level is a strong negative predictor of the rate of protein evolution (Pál et al. 2001; Drummond et al. 2005; Lemos et al. 2005; Larracuente et al. 2008), yet most studies examining patterns of molecular evolution for reproductive proteins do not correct for this (but see Meisel 2011). To evaluate whether the trends in evolutionary rates were sensitive to differences in expression, we performed partial correlations between evolutionary rate, expression level, and tissue enrichment. We recovered the expected negative correlation between expression level (measured as reads per kilobase per million mapped reads; RPKM) and evolutionary rate (ω) in all three tissues, and found that the observed

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

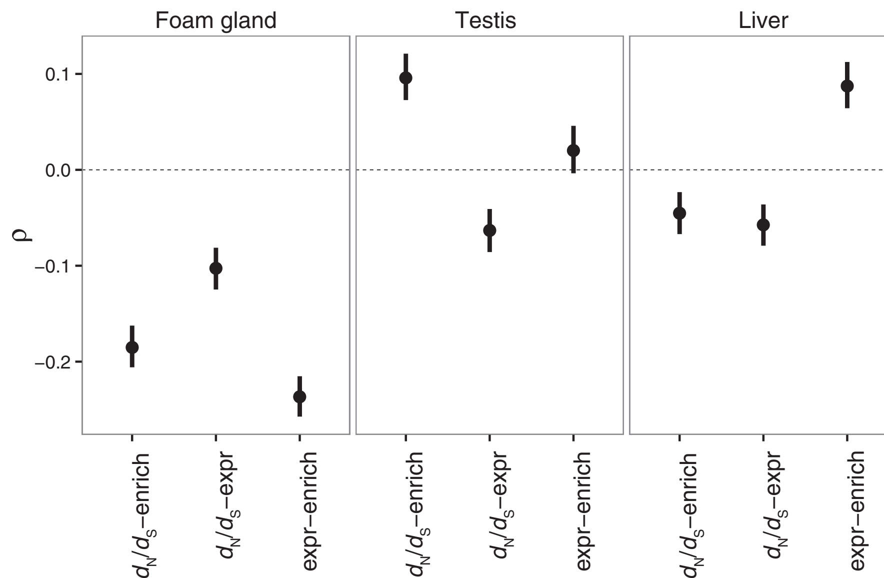

Fig. 2. Point estimates and 95% CIs (bars) of the partial correlation coefficient (ρ) between pairwise chicken:quail evolutionary rate (d_N/d_S), gene expression level (expr) estimated by RPKM, and tissue-enrichment (enrich) measured by log fold change of a gene in a particular tissue (foam gland, testis, or liver) when compared with the average of the other two. Bars that cross zero (dashed line) indicate no correlation, whereas estimates above and below the line suggest positive and negative correlations, respectively, between the two focal parameters after correcting for correlations with the third parameter. CIs were constructed from 1,000 bootstrapping resamplings of the mean without assuming normality. Sample size is 8,238 for all classes.

variation in rates of amino acid substitution across tissues is robust to differences in expression level (fig. 2). Enrichment in both the foam gland and liver was negatively correlated ($\rho < 0$) with the rate of protein evolution (ω or d_N/d_S), but the correlation was significantly stronger for the foam gland than the liver. Enrichment in testis was positively correlated ($\rho > 0$) with evolutionary rate after correcting for expression levels. Enrichment and expression level were also negatively correlated in the foam gland (and only the foam gland), suggesting that many of the genes specific to the foam gland are lowly expressed. This finding makes the slow evolution of foam gland-expressed genes all the more prominent, as lowly expressed genes generally evolve more rapidly (Pál et al. 2001; Drummond et al. 2005; Lemos et al. 2005; Larracuente et al. 2008).

Lineage-Specific Genes

Genes involved in foam-gland function may be specific to the quail lineage as a result of gene duplication or some other mechanism, both because genes from male reproductive tissues tend to have high levels of lineage-specific genes and because the foam gland represents a novel phenotype (e.g., Baker et al. 2012; Bailey et al. 2013). To address this topic, we examined the proportion of genes from each tissue class that 1) had no Basic Local Alignment Search Tool (BLAST) hits to chicken, 2) had no clear 1:1 orthologs with chicken, and 3) were putative quail-specific duplicates as identified by OrthoMCL (Li 2003; Chen et al. 2006). Overall, tissue class significantly explained variation in proportions of genes without BLAST hits (tissue-enriched: $P << 0.0001$, Pearson's $\chi^2 = 128.91$, df = 3; tissue-restricted: $P << 0.0001$, Pearson's $\chi^2 = 1644.55$, df = 3) or 1:1 orthologs in chicken (tissue-enriched: $P << 0.0001$, Pearson's $\chi^2 = 89.90$, df = 3; tissue-restricted: $P << 0.0001$, Pearson's $\chi^2 = 1683.23$, df = 3). After correcting for multiple pairwise comparisons in the tissue-enriched gene designation, we found that liver, not foam gland or testis, had the highest proportion of lineage-specific genes in terms of genes without BLAST hits or 1:1 orthologs in the chicken (table 2). When comparing proportions of lineage-specific genes across the tissue-enriched and tissue-restricted designations, nonspecific genes showed contrasting patterns (high in tissue-enriched, low in tissue-restricted). The difference in the definition of nonspecific genes likely explains the reversed outcomes. For tissue-enriched genes, nonspecific genes are those that do not exhibit any biased expression in any tissue. This definition is more exclusive than the nonspecific classification for tissue-restricted genes, which simply requires that a gene must be expressed above some level in all three tissues. The latter classification is more similar to what are generally considered broadly expressed genes, and match better expectations of widely expressed genes from other systems (e.g., Bailey et al. 2013). Tissue also explained significant variation in the proportion of genes with potential duplicates (tissue-enriched: $P = 0.007$, Pearson's $\chi^2 = 12.12$, df = 3; tissue-restricted: $P = 0.006$, Pearson's $\chi^2 = 12.53$, df = 3). After correcting for multiple pairwise comparisons, genes with enriched expression in testis showed significantly higher proportions of potential duplicates than nonspecific genes ($P = 0.022$; table 2). No other pairwise comparison was significant ($P > 0.05$ in all cases).

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

Table 2. Potential Lineage-Specific Genes across Tissue Classes.

Tissue	*N* Genes in Tissue	% Genes without Chicken BLAST Hit	% Genes without 1:1 Chicken Orthologs	% Potential Duplicates
Tissue-enriched:				
Foam gland	2132	42.87[a]	61.02[a]	3.90[ab]
Testis	5782	44.79[a]	61.73[a]	3.16[a]
Liver	3086	53.56[b]	66.11[b]	2.37[ab]
Nonspecific	2100	55.29[b]	72.33[c]	1.95[b]
Tissue-restricted:				
Foam gland	500	64.80[a]	77.60[a]	1.80[a]
Testis	4360	59.29[b]	77.45[a]	2.50[a]
Liver	1282	74.49[c]	83.39[b]	1.95[a]
Nonspecific	12980	32.77[d]	47.43[c]	3.19[a]

Unique letters indicate the proportions within each tissue designation are significantly different from each other at $P < 0.05$ after correcting for multiple tests.

Lineage-Specific Shifts in Evolutionary Rates

Pairwise estimates average ω over the entire history separating two species, but genes from the foam gland may have experienced shifts in selective pressure specific to the quail lineage. The evolutionary rate of foam genes may have accelerated along the quail lineage ($\omega_{lineage} > \omega_{tree}$) due to positive selection acting on foam gland genes as they acquired novel, foam-producing functions. Alternatively, the rate of evolution for foam gland genes may have decelerated along the quail branch ($\omega_{quail} < \omega_{tree}$), if acquisition of foam-producing function is associated with increased purifying selection due to additional constraints. To determine whether classwide shifts in evolutionary rate occurred, we placed the quail genes in a phylogenetic framework by aligning orthologs across turkey, chicken, and zebra finch. We then estimated whether ω for each gene differed significantly along the quail (ω_{quail}) or chicken ($\omega_{chicken}$) lineages when compared with the rest of the tree (ω_{tree}). Specifically, we predicted that the proportion of genes with significant accelerations ($\omega_{lineage} > \omega_{tree}$) or decelerations ($\omega_{lineage} < \omega_{tree}$) would be greatest in foam gland genes along the quail lineage. In fact, after correcting for multiple comparisons, genes with enriched expression in the foam gland and liver showed no significant differences from expectations in any case (fig. 3 and supplementary fig. S5, Supplementary Material online). In contrast, genes with enriched expression in the testis experienced proportionally more accelerations in evolutionary rate along the quail lineage ($\omega_{quail} > \omega_{tree}$) than expectations based on transcriptome-wide values from quail (Pearson's $\chi^2 = 17.214$, $P = 6.01 \times 10^{-4}$ after correcting for multiple tests, df = 1; fig. 3). Although it appears that there are strong trends toward classwide decelerations of foam gland-enriched genes along the quail lineage and foam gland-restricted genes along the chicken lineage (fig. 2 and supplementary fig. S5, Supplementary Material online), these trends were not significant after correcting for multiple comparisons. For these analyses, it is important to recognize that 1) the tissue-restricted gene classes had low sample sizes for the foam gland ($N = 49$) and 2) the dashed lines in figure 3 represent expectations generated from the complete transcriptome, whereas each tissue was actually compared with transcriptome-wide values minus those from the focal tissue. Testis-enriched genes also experienced larger than expected shifts in selective pressure along the quail lineage (CIs > 0; fig. 4 and supplementary fig. S6, Supplementary Material online). These shifts were also significantly larger than for any tissue examined along the quail lineage and greater than testis-expressed genes along the chicken branch. It should be noted that, because we are only examining a small number of taxa with fairly long branches, our power to detect significant shifts in selection is low.

Five quail genes with significant accelerations in evolutionary rate along the quail lineage also showed signs of positive selection along the quail lineage ($\omega_{quail} > 1$; table 3). Four of

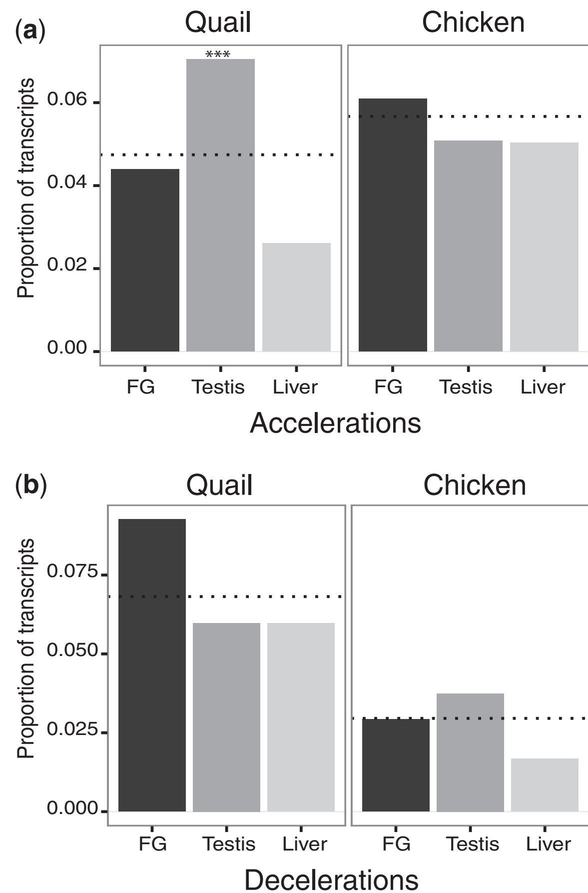

FIG. 3. The proportion of genes enriched in a particular tissue experiencing (*a*) accelerations ($\omega_{lineage} > \omega_{tree}$) or (*b*) decelerations ($\omega_{lineage} < \omega_{tree}$) in ω estimates along either the quail or chicken lineage when compared with the rest of the tree. Estimates are based on tissue-enriched genes with 1:1:1:1 orthologs between quail, chicken, turkey, and zebra finch. Tissue designations are based on expression patterns in quail. Each subset was compared with the species-specific transcriptome wide expectation for a particular designation (dotted lines) minus the subset under investigation. Asterisks indicate a significant difference in a χ^2 test after correcting for multiple comparisons (*** $P < 0.001$). Sample sizes: FG = 410, Testis = 1,121, Liver = 536.

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

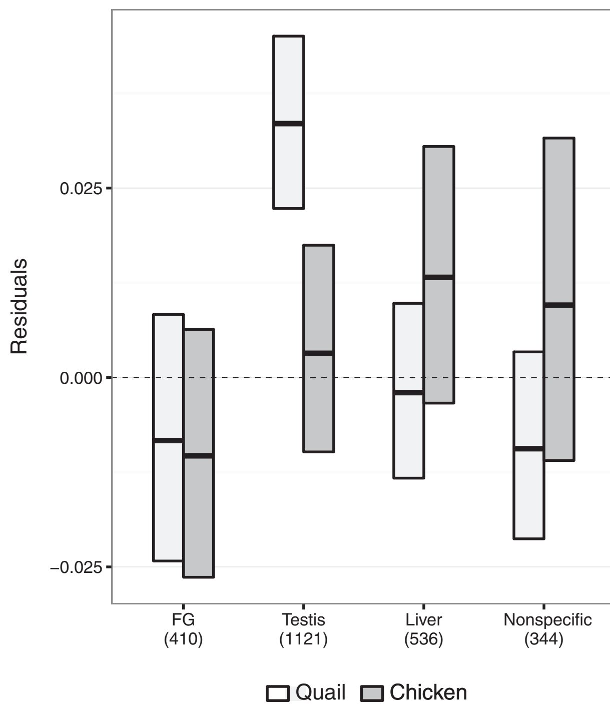

Fig. 4. Average residuals (bar) and 95% CIs (boxes) from models regressing ω estimates along a single lineage (quail or chicken) on ω estimated from the rest of the tree for tissue-enriched genes. These values are interpreted as the magnitude of class-wide shifts in selective pressure along either the quail or chicken lineage for a particular set of genes. Boxes that do not overlap with the dotted line (0) are significantly different from expectation. Lineage-specific estimates from quail are light gray and those from chicken are dark gray. Tissue designations were based on expression pattern in quail. Estimates are based on the subset of all tissue-enriched genes with 1:1:1:1 orthologs between quail, chicken, turkey, and zebra finch. CIs derived from bootstrap resamplings of the mean without assuming normality. Sample sizes are in parentheses. FG = foam gland.

these five genes exhibited testis-specific expression and one revealed foam gland-specific expression in the quail. All testis-enriched genes under positive selection were also expressed in chicken testis in at least one of the chicken expression data sets. Many genes evolving adaptively may not be identified by our analysis, because estimates of ω over entire genes are quite conservative (i.e., $\omega > 1$ may hold for only a small part of a protein-coding gene).

Discussion

Foam-Gland Expressed Genes Diverge Surprisingly Slowly

The phenomenon of rapid divergence of genes encoding reproductive proteins, and seminal fluid proteins in particular, is one of the most striking and consistent patterns across molecular evolution (Swanson and Vacquier 2002a, 2002b; Clark et al. 2006; Panhuis et al. 2006; Turner and Hoekstra 2008). Here we have examined whether this pattern extends to genes expressed in the foam gland of male Japanese quail. Evolution of the protein constituents of foam from this novel structure provides an interesting comparison to seminal fluid protein evolution, because foam functions in sexual reproduction independent of semen. We computed rates of protein evolution (ω) between quail and three different avian species and compared patterns across genes with biased expression in three tissues (foam gland, testis, or liver) and genes with unbiased expression. Our study highlights the value of examining reproductive protein evolution in diverse taxa and tissue types, as we find an unexpected and dominant role for selective constraint shaping the evolution of foam gland-expressed genes.

Given the nearly ubiquitous finding of elevated divergence of seminal fluid proteins as a class across diverse taxa (e.g., Swanson et al. 2001; Clark and Swanson 2005; Dean et al. 2009; Walters and Harrison 2010), we expected that genes expressed in the foam gland would likewise evolve rapidly. Surprisingly, we find that foam gland genes evolved more slowly than genes expressed in testis and (in most cases) liver, even after correcting for differences in gene expression level (fig. 2 and table 1 and supplementary table S4, Supplementary Material online). In fact, the more that expression of a gene is enriched in the foam gland, the more slowly it evolves (fig. 2). The extent of purifying selection acting on foam gland-expressed genes is similar to that for nonspecific genes (those that are broadly expressed), which are generally considered to exhibit the slowest rate of evolution in the genome (Duret and Mouchiroud 2000; Zhang and Li 2004; Larracuente et al. 2008). In genes with expression restricted to the foam gland, we also see a nonsignificant trend toward a slowdown in evolutionary rate, specifically along the quail lineage (fig. 3). Taken together, our results provide convincing evidence that genes with biased expression in the foam gland do not display the classwide rapid divergence dynamics typical of many seminal fluid proteins; instead, they appear to evolve under strong selective constraint. In other systems, many groups of genes from tissues of the male reproductive tract also evolve relatively slowly as a class and even a majority of individual seminal fluid proteins evolve under purifying selection (Swanson et al. 2001; Dean et al. 2009).

Why do genes expressed primarily in the foam gland evolve so slowly? The signature of increased purifying selection on foam gland genes appears to be driven by genes responsible for the structure of the foam gland, rather than those genes encoding secreted proteins likely to comprise foam (table 1). Because the foam–foam gland complex represents a novel phenotype, the proteins forming the structural components of the foam gland may be a result of co-option followed by conservation due to their other functions (Klemm et al. 1973). Consistent with this hypothesis, the foam gland is comprised of secretory cells housed in a series of individual glandular units similar to the “simple glands” forming the digestive tract, with the exception that the foam gland is bound by connective tissue (Klemm et al. 1973). Additionally, proteins that give foam its characteristic physical properties may also be co-opted; the foam is similar in texture and form to whipped egg white and any shared

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

Table 3. Evolutionary Patterns and Descriptions of Accelerated Genes along the Quail Lineage, Where $\omega_{quail} > 1$.

Quail Transcript	Chicken Ortholog Name	d_N[a]	d_S[a]	ω[a]	Length[a]	Divergence[a]	ω_{tree}[b]	ω_{quail}[b]	P[c]	Enriched[d]	Log FC[e]	Ensembl Gene ID[f]
Coja22788_c0_seq1	Radial spoke head 1 homolog (*RSPH1*)	0.151	0.204	0.743	774	0.163	0.247	1.249	0.026	Testis	8.836	ENSGALG00000016175
Coja15527_c0_seq1	Cancer susceptibility candidate 1 (*CASC1*)	0.120	0.133	0.905	2139	0.123	0.443	1.167	0.002	Testis	4.336	ENSGALG00000014030
Coja19413_c0_seq1	Dynein, axonemal, assembly factor 1 (*DNAAF1*)	0.021	0.138	0.153	753	0.049	0.184	1.017	0.001	Testis	7.712	ENSGALG00000003258
Coja35683_c0_seq1	Pyridoxal-dependent decarboxylase domain containing 1 (*PXDC 1*)	0.041	0.093	0.441	1482	0.054	0.177	1.011	0.000	Foam gland	2.355	ENSGALG00000002796
Coja15477_c0_seq1	K(lysine) acetyltransferase 6A (*KAT6*)	0.013	0.122	0.110	1974	0.039	0.115	1.008	0.000	Testis	2.474	ENSGALG00000003641

[a]Pairwise estimates from 1:1 chicken orthologs.
[b]Estimated from quail:chicken lineage specific analyses.
[c]P value from LRT of one versus two-ratio branch models after applying a FDR cutoff of 0.05.
[d]Tissue expression in quail designated in this study.
[e]Log fold change of transcript in enriched tissue versus average of other two.
[f]Based on chicken ortholog

proteins may be under constraint due to their functions in females.

Alternatively, for those genes encoding individual foam proteins that play a role in sperm competition, purifying selection may be acting more efficiently on them than genes with enriched expression in testis or liver. For example, foam plays a role in sperm competition and if a single "type" of foam is always favored in sperm competition, any deviations from this type may be disfavored (Finseth et al. 2013). One piece of evidence supporting this idea is that foam enhances sperm motility whether self-derived or from a different male (Cheng, McIntyre, et al. 1989). Intense sperm competition could quickly weed out any deleterious mutations in genes encoding certain foam proteins and decrease the rate of amino acid substitution. Pollen competition appears to be driving a similar effect in pollen-specific genes, though this process is amplified in plants as deleterious recessive alleles are unmasked due to haploid gene expression in pollen tubes (Arunkumar et al. 2013). If sperm competition does indeed intensify the strength of purifying selection acting on foam, expression variation in the abundance of individual components of the foam or the size of the foam complement may be more important to an individual's reproductive success.

Finally, although foam mediates the outcome of sperm competition in the lab (Cheng, McIntyre, et al. 1989; Finseth et al. 2013), it may not be the primary selective force shaping foam gland genes. Circumstantial evidence suggests that Japanese quail males experience high levels of sperm competition, but we still lack a comprehensive study of their mating system in the field (Nichols 1991; Adkins-Regan 1995; Teijeiro et al. 2003). Basic functions of foam in reproduction (improving sperm transfer, transport, and/or motility) may place constraints on foam gland-expressed genes outside of the context of sperm competition.

Some genes important in the evolution of the foam gland may not have been identified by our analyses of evolutionary rates. These would include genes that evolve so rapidly that we could not identify orthologs, genes specific to the quail lineage resulting from gene duplication events since divergence from the chicken lineage (our 1:1 reciprocal best blast approach does not detect paralogs), or genes arising de novo in the quail (Long et al. 2003). To address this issue, we compared the incidence of lineage-specific genes across tissue classes. Based on what we know about the evolution of male reproductive genes in other systems and because the foam gland represents a novel structure, we expected that genes expressed in the male reproductive tissues would show an abundance of lineage-specific genes (i.e., few BLAST hits and orthologs, many duplicates; e.g., Baker et al. 2012; Bailey et al. 2013). Instead, we find no evidence that genes with enriched expression in the foam gland are represented by relatively high numbers of lineage-specific genes (table 2). As de novo transcriptomes may contain alternatively spliced isoforms, collapsed paralogs, or allelic variants, follow-up analyses with a well-annotated genome will clarify the role of lineage-specific genes during the evolution of the foam gland.

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

Heterogeneity of Evolutionary Rate across Male Reproductive Tissues Disappears when Considering Secreted Genes

Our data contribute to the growing body of evidence that genes expressed in particular tissues exhibit heterogeneity in mean evolutionary rates (Dean et al. 2008; Grassa and Kulathinal 2011; Arunkumar et al. 2013). We find contrasting evolutionary rates for genes derived from two different avian male reproductive tissues. Genes from the foam gland diverge relatively slowly, whereas those from the testis evolve rapidly (fig. 2 and table 1 and supplementary table S4, Supplementary Material online). This pattern is driven by differences in d_N across tissue classes and is robust to variation in expression level (fig. 2 and supplementary fig. S4, Supplementary Material online).

One model of rate heterogeneity (Dorus et al. 2006; Dean et al. 2009; Dorus et al. 2010) suggests that reproductive proteins that interact with the environment (either the male or female reproductive tract or other sperm), and therefore have the potential to be involved in sexual selection, are expected to evolve more rapidly than those that do not. Consistent with this hypothesis, proteins from the sperm cell membrane, sperm acrosome, and seminal vesicle (which produce the copulatory plug) diversify faster than other kinds of male reproductive proteins (Dorus et al. 2006; Dean et al. 2009; Dorus et al. 2010; Dean 2013). Likewise, sperm proteins involved in postmating, but not premating, interactions show elevated rates of evolution in species with higher levels of sperm competition (Schumacher et al. 2013). Because the foam complement is passed on to females and interacts with sperm to mediate the outcome of sperm competition, it would seem that the observation of slowly evolving foam genes contradicts this hypothesis (Cheng, McIntyre, et al. 1989; Finseth et al. 2013). However, to properly test this model of rate heterogeneity, the subset of foam gland-expressed genes that produce proteins that interact with the environment (i.e., those encoding foam proteins) need to be distinguished from genes with expression enriched in the foam gland, but not producing a component of foam.

One first step is to examine the group of foam gland-expressed genes that encode proteins that are likely secreted (presumably components of foam). Secreted proteins are more likely to interact with "foreign" proteins, including those from the female reproductive tract and sperm/seminal fluid from other males. Indeed, secreted proteins from seminal fluid often exhibit elevated d_N/d_S ratios when compared with nonsecreted proteins (Swanson et al. 2001; Clark and Swanson 2005; Dean et al. 2008; Turner et al. 2008). Here, we focused on whether evolutionary rates of secreted proteins varied across tissue class. Intriguingly, the dramatic heterogeneity in pairwise ω values across tissues from the complete data set is largely erased when considering the tissue-specific genes that produce products that are putatively secreted (identified by the presence of a signal peptide; table 1 and supplementary table S4, Supplementary Material online). Therefore, proteins found in foam may not be diverging slowly but at rates similar to secreted proteins from testis and liver. It should be noted that many seminal fluid proteins do not contain canonical signal peptides, so many (potentially slowly evolving) genes encoding foam proteins may not be represented by the subset of genes with a signal peptide (Swanson et al. 2001; Clark and Swanson 2005). Additionally, incomplete contigs may have contained signal peptides that were not characterized in the de novo transcriptome assembly.

The Foam Gland, a Novel Phenotype, Arose without Major Changes in Selective Pressure

Because of divergence of the quail and chicken lineages, evolution along the *Coturnix* branch has resulted in the appearance of a qualitatively new male reproductive organ. The foam gland of Japanese quail represents major phenotypic divergence in both morphology and physiology. The gland is an enlarged, reddened protuberance consisting of many glandular units interdigitated with the cloacal sphincter muscle, the likes of which are not found in other genera (McFarland et al. 1968; Klemm et al. 1973). The gland also produces a chemically and physically unique foam (Klemm et al. 1973; Fujihara 1992). Thus, the foam gland represents a novel phenotype and our study can begin to shed light on the nature of genetic changes responsible for novel structures.

Our results provide evidence that phenotypic change in the foam gland arose without pervasive change in the sequences of protein-coding genes (table 1 and supplementary table S4, Supplementary Material online). However, the intensity or direction of selection can change over time, resulting in variation in ω along specific branches of a phylogenetic tree. This is particularly relevant for genes that encode proteins incorporated into novel structures such as the foam gland, because their new functional roles may alter selective constraints. We hypothesized that, as a class, genes expressed in the foam gland should experience changes in selective pressure along the *Coturnix* branch with the acquisition of the novel functions associated with foam gland development and foam production. We also expected that shifts in ω would be relatively larger for foam gland-expressed genes along the quail lineage than for genes from other categories. Contrary to our predictions, the foam gland evolved without major, classwide shifts in selective pressures along the quail lineage. We find no evidence of either greater than expected accelerations (i.e., relaxation of selective constraint or elevated adaptive evolution) or decelerations (i.e., increased purifying selection) of foam gland genes along either chicken or quail lineages (fig. 3, supplementary fig. S5, Supplementary Material online). Additionally, the magnitude of shifts in selective pressure on foam gland genes did not vary from transcriptome-wide expectations (fig. 4 and supplementary fig. S6, Supplementary Material online).

How did the foam gland evolve without major protein-coding changes in foam gland genes? One likely scenario is that divergence in gene expression levels, not in protein-coding sequences, is primarily responsible for the evolution of the foam gland. Novel phenotypes can arise from mutations affecting *cis*-regulatory elements, altering the spatial and

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

temporal patterns of gene expression (Carroll 2005; Wray 2007; Carroll 2008). In general, postmating sexual selection may target regulation of reproductive proteins as expression of male-derived reproductive proteins evolves rapidly (Nuzhdin et al. 2004; Khaitovich 2005; Ellegren and Parsch 2007) and rate of production of sperm proteins is a consequence of sperm competition (Ramm and Stockley 2010). Adaptive changes in traits driven by postmating sexual selection could therefore accumulate in regulatory, rather than protein-coding, regions (Martin-Coello et al. 2009).

Testis-Enriched Proteins Evolve Rapidly in a System with Physiological Polyspermy

In many species, females evolve mechanisms to prevent polyspermy, because multiple sperm entering the egg results in embryo mortality. Polyspermy avoidance has been suggested to promote the rapid diversification of many reproductive proteins (Swanson and Vacquier 2002a, 2002b). In the avian fertilization system, there is no selective disadvantage to polyspermy; females tolerate physiological polyspermy and always permit multiple sperm to enter the egg (Perry 1987; Snook et al. 2011). Thus, examinations of molecular evolutionary patterns of male reproductive proteins from avian systems like Japanese quail allow an important contrast with other taxa, given that sexual conflict over polyspermy avoidance is not a potential driver of observed elevated evolutionary rates (Swanson and Vacquier 2002a). Earlier studies in avian taxa showed that 1) a few, but not all, individual gamete-recognition genes evolve adaptively (Berlin and Smith 2005; Calkins et al. 2007; Berlin et al. 2008) and 2) genes expressed in chicken testis have lower than expected rates of nonsynonymous change, but those expressed exclusively in the testis ($N = 12$) exhibit elevated rates of nonsynonymous change (Grassa and Kulathinal 2011).

Despite physiological polyspermy in birds, we find strong evidence for the rapid evolution of testis genes. Genes with increased expression in testis exhibit 1) elevated rates of protein evolution (table 1 and supplementary table S4, Supplementary Material online), 2) the greatest number of genes under positive selection ($\omega > 1$; table 1 and supplementary table S4, Supplementary Material online), 3) a burst of accelerated evolution along the quail lineage (fig. 3), 4) relatively large accelerations in evolutionary rate (fig. 4 and supplementary fig. S6, Supplementary Material online), and 5) slightly higher levels of lineage-specific duplication than nonspecific genes (table 2). Avian systems provide a rich source of comparative reproductive data for further work that could discriminate among other hypotheses for rapid evolutionary change of reproductive genes, e.g., sperm competition, cryptic female choice, other types of sexual conflict, or host-pathogen avoidance.

The more that gene expression is enriched in the testis, the more rapidly the gene evolves, even after correcting for expression differences (fig. 2 and table 1 and supplementary table S4, Supplementary Material online: restricted > enriched for *Coturnix*–*Gallus* and *Coturnix*–*Meleagris* comparisons). Similar patterns are found in genes specialized in the mouse epididymis (Dean et al. 2008) or exclusively expressed in male reproductive tissues (Dean et al. 2009). This is not surprising given that genes with tissue-specific expression tend to evolve more rapidly than genes expressed in many tissues (Duret and Mouchiroud 2000; Zhang and Li 2004; Liao 2006; Larracuente et al. 2008). However, genes with enriched expression in foam gland and liver do not show the same trend (fig. 2 and table 1 and supplementary table S4, Supplementary Material online). In numerous taxa, sex-biased genes in sex-limited tissues evolve faster than other narrowly expressed genes (Meisel 2011; Parsch and Ellegren 2013). Although the testis is sex limited, the liver and foam gland are also found in females, and this may place additional constraints on them. For example, because the foam gland is closely associated with the cloacal sphincter muscle, the presence of a rudimentary, nonfoam-producing foam gland in females may be important for proper cloacal function (McFarland et al. 1968; Klemm et al. 1973; King 1981). Positive selection may therefore act more efficiently on testis-specific genes, as these are free from possible pleiotropic constraints associated with shared expression in female tissues.

To understand what may be driving the rapid evolution of genes in quail, we explored the function of genes with significant accelerations along the quail lineage and $\omega_{\text{quail}} > 1$, because these may have evolved under positive selection. Functions are based on annotations of the chicken orthologs. Strikingly, four out of five genes with these characteristics exhibit highly enriched expression in quail testis and are also expressed in (and sometimes restricted to) chicken testis (table 3). One of this subset, *CASC1* (*Coja15527_c0_seq1*), is found in mouse whole sperm, whereas two others encode proteins related to axonemes, *RSPH1* (*Coja22788_c0_seq1*) and *DNAAF1* (*Coja19413_c0_seq1*), which are the molecular motors of sperm flagella. Although sperm flagella-associated genes tend to evolve slowly in other systems (Dorus et al. 2010), this class of genes may be a general target of selection due to sperm competition in birds (Rands et al. 2013).

It is notable that we observe accelerated and adaptive evolution of sperm flagellar proteins along the *Coturnix*, but not the chicken, branch. Among Galliformes, Japanese quail sperm have several unique features including: 1) flagella that are over twice as long as other nonpasserine birds (Korn et al. 2000), 2) exceptionally long midpieces covering 64–74% of total sperm flagellum length (Woolley 1995; Korn et al. 2000), and 3) midpieces that contain numerous mitochondria (1,400–2,500) per cell (Woolley 1995; Korn et al. 2000). Taken together, the co-occurrence of quail-specific sperm attributes, rapidly and adaptively evolving sperm-flagella associated genes (table 3), and accelerated evolution of testis-derived genes (figs. 3 and 4) suggests sperm have experienced an exceptional amount of phenotypic and genetic change along the quail lineage when compared with other Galliformes. The rapid evolution of sperm phenotype in *Coturnix* quail may not have been independent of foam. Previously, foam and sperm midpiece length have been hypothesized to coevolve. For example, sperm may require

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

more energy to move through the foam or, alternatively, foam may provide substrates such as lactate that are necessary for sperm motility (Korn et al. 2000; Singh, Sastry, Pandey, et al. 2011).

Conclusions

Although it is true that many male reproductive genes evolve quickly, classwide elevations of evolutionary rate are driven by a subset of reproductive genes and the majority of male accessory gland proteins actually evolve under constraint. In sexual selection and fertilization, foam presumably plays a role similar to that of seminal fluid proteins, yet genes expressed in the foam gland evolve slowly. In contrast to the foam gland, genes with biased expression in the testis evolve rapidly and experienced a burst of accelerated evolution along the quail lineage, despite polyspermy tolerance in birds. Our data suggest that the rapid evolution of reproductive tract genes may not always be a byproduct of sexual selection and reinforces the role of selective constraint as an important force shaping male reproductive genes.

Materials and Methods

Subjects

Japanese quail were lab-reared and housed on a 16:8 light:dark cycle to simulate breeding conditions. Three tissue samples (foam gland, testis, and liver) were collected from six sexually mature Japanese quail males with phenotypically normal foam glands. Two males were approximately 1-year old and euthanized in February 2011. The remaining four males were approximately 5 months old and euthanized in November 2011. After euthanizing with CO_2, we immediately dissected out foam gland, testis, and liver tissues from each male and froze samples on liquid nitrogen. Tissues were later frozen at $-80\,^{\circ}C$ until RNA extraction. All animal procedures were approved by Cornell University's Institutional Animal Care and Use Committee under permit 2002-0117.

Library Preparation and Sequencing

We extracted RNA from 18 samples (3 tissues $\times$ 6 males) using the Agencourt RNAdvance Tissue Kit (Beckman Coulter) following the manufacturer's instructions, except that we performed half reactions. RNA quality and concentration was assessed by agarose gel electrophoresis and NanoDrop spectrophotometry. We confirmed RNA purity and integrity using an Agilent 2100 BioAnalyzer. In January 2012, Illumina libraries were prepared from 1.2 μg total RNA using the TruSeq RNA Sample Preparation Kit (Illumina) following manufacturer's instructions. All samples were tagged with a unique adapter index, pooled, and single-end sequenced on the equivalent of two lanes of an Illumina HiSeq 2000, with a target read length of 100 bp. Sequencing was performed by the Cornell University Institute of Biotechnology's Genomics Facility.

Transcriptome Assembly and Characterization

Initial quality filtering and barcode removal were conducted by the Cornell University Institute of Biotechnology's Genomics Facility. We used fastq-mcf (http://code.google.com/p/ea-utils/wiki/FastqMcf, last accessed September 16, 2014) to remove Illumina adaptors, trim low-quality terminal ends, discard short sequences, and filter reads with phred scores < 20. We merged the 18 libraries into a single file and assembled a transcriptome with the Trinity pipeline with default parameters (Grabherr et al. 2011). Trinity previously produced a high-quality foam gland transcriptome for this species (Finseth and Harrison forthcoming). The Trinity assembly was executed on a Linux, Dell PowerEdge R710 with 16 cores, 64-GB RAM, and 1-TB hard disk drive (HDD) housed at the Cornell University Institute of Biotechnology's Genomics Facility. The resulting library had 85,900 contigs. We reduced all contigs with putative isoforms identified by Trinity (i.e., multiple contigs with same component number) to the single longest isoform (removed 2,677 contigs). In 27 instances, there was not a single longest isoform. If this was the case, we retained the isoform with the best BLASTp to the chicken, or, if that did not break the tie, we kept the first sequence generated by Trinity. We identified open reading frames with OrfPredictor (Min et al. 2005) and kept only those sequences with at least one open reading frame. The transcriptome had 81,868 contigs and this version is referred to as the "exhaustive" transcriptome.

Filtered reads from each sample were aligned to the transcriptome using the aln method in the Burrows-Wheeler Aligner (Li and Durbin 2009). The number of reads per sample uniquely mapped to each contig was tabulated with samtools as implemented in custom python scripts (Li et al. 2009). To generate a set of high-quality contigs that represent real transcripts, we filtered out lowly expressed contigs by retaining only those with at least one read aligned per every million reads for at least six samples (i.e., the number of biological replicates per tissue). We refer to the remaining 24,035 transcripts as the "filtered" transcriptome throughout and used this version of the transcriptome for most downstream analyses. We calculated standard metrics of transcriptome quality for the filtered transcriptome including: average transcript length, median transcript length, N50 (median transcript size, weighted by length), and summed transcript lengths (Kumar and Blaxter 2010; Hornett and Wheat 2012).

Tissue-Specific Gene Expression

We employed an RNA-Seq approach with 18 samples to assign tissue specificity for a given gene. EdgeR version 3.0.8 (Robinson et al. 2010) was used to normalize our RNA-Seq data by the trimmed mean of M values approach (Robinson and Oshlack 2010) and visualize overall sample relationships using the multidimensional scaling feature (similar to a principal components analysis). We designated tissue specificity of gene expression in two ways, "tissue-enriched" and "tissue-restricted." A gene was considered tissue-enriched when itsexpression levels were significantly higher in the focal tissue compared with the other two. A gene exhibited tissue-restricted expression when it was expressed in only one tissue. Major results were similar for both designations, so we report methodological details and results of

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

tissue-restricted analyses in the supplementary materials (supplementary methods S1, Supplementary Material online).

For the tissue-enriched designation, we tested for differential expression of genes using the multifactor generalized linear models (glms) approach in EdgeR. We fit negative binomial glms with Cox-Reid tagwise dispersion estimates to models that included tissue and male ID as factors, as our experiment was paired by subject. Our design matrix specified contrasts to analyze expression level in one tissue versus the average of the other two. To determine differential expression, we performed likelihood ratio tests by dropping one coefficient from the design matrix (i.e., the "null" model) and comparing it to the full model (McCarthy et al. 2012). The list of genes significantly upregulated by more than log 2-fold change in each tissue were considered "enriched" in a particular tissue based on a false discovery rate (FDR) of 5%. Any gene that was classified as enriched in two tissues was excluded from further analyses ($N = 366$ out of 11,366 enriched genes). Genes not enriched for any of the three tissues comprised a fourth category termed "nonspecific."

Orthology Assignment

To identify orthologs of quail genes in chicken (*Gallus gallus*), turkey (*Meleagris gallopavo*), and zebra finch (*Taeniopygia guttata*), we obtained all Ensembl protein and coding sequences from each species (Ensembl version 69: *G. gallus* assembly WASHUC2; *M. gallopavo* assembly UMD2; *T. guttata* assembly taeGut3.2.4) via BioMart (www.biomart.org, last accessed September 16, 2014). We filtered the downloaded protein sets to remove redundant entries (i.e., alternative splice variants) by self-BLAST following Hornett and Wheat (2012). We identified the longest open reading frames and associated protein translations of our exhaustive transcriptome with OrfPredictor (Min et al. 2005). We then identified quail:species orthologs via the reciprocal best blast method (Tatusov 1997; Bork and Koonin 1998; Koonin 2005). In short, using BLASTp, we compared the proteins predicted from the translation of our quail transcriptome with each species' protein sequences, with a cutoff e-value of 1×10^{-6} and vice versa. For chicken, orthologs were called when the top chicken hit (based on bit score) from the quail to chicken BLAST returned the original quail query in the chicken to quail BLAST (10,129 orthologs from the exhaustive transcriptome, 9,620 from the filtered transcriptome). We restricted all downstream analyses to the subset of ortholog pairs (8,668 orthologs) without multiple stop codons in the chicken coding sequences as identified by custom Perl scripts. We repeated this for the turkey (8,717 orthologs) and zebra finch (9,746 orthologs).

To evaluate whether our assembly returned full-length transcripts, we computed the "ortholog hit ratio" for each transcript from the filtered transcriptome, as described by O'Neil et al. (2010). This ratio represents the length of a putative coding region of a quail transcript divided by the length of the coding region of the orthologous chicken transcript. For calculation of ortholog hit ratios, the putative coding region was estimated from the alignment length of the best BLASTp result between a transcript and its chicken ortholog. An ortholog completely represented by a transcript would have a ratio of "1." Ratios less than 1 indicate instances where transcripts only partially covered orthologs, whereas ratios greater than 1 indicate insertions in transcripts.

Pairwise Estimation of Evolutionary Rates

Protein sequences from the 1:1 quail:species (chicken, turkey, or zebra finch) orthologs were aligned with Clustal W version 2.1 (Larkin et al. 2007), which guided alignments of their corresponding DNA sequences with PAL2NAL (Suyama et al. 2006) as implemented in the Parallel Alignment and Translation (ParaAT) version 1.0 tool (Zhang et al. 2012). From these alignments, we estimated the ratio of the nonsynonymous substitution rate (d_N) to the synonymous substitution rate (d_S). The d_N/d_S ratio, or ω, is the normalized amino acid substitution rate and we interpret this as the "evolutionary rate" of proteins. We calculated pairwise ω values using the model averaging approach in the software package KaKs_Calculator version 1.2 (Zhang et al. 2006). Model averaging in KaKs _Calculator implements a set of 14 candidate models in a maximum likelihood framework and computes the weighted average of d_N, d_S, and ω for each gene. We removed orthologs for which $d_S \geq 2$ times the mean d_S estimated from all orthologs for a particular species comparison, as these might reflect poor alignment. We also removed ortholog pairs for which d_S estimates approached 0, resulting in an artificially inflated ω (e.g., $\omega \sim 50$; three for *Coturnix–Gallus*, two for *Coturnix–Taeniopygia*). As further quality control, we confirmed that all alignments were at least 100 amino acids long and that no d_N values were greater than 1 (indicating more than 1 nonsynonymous substitution per nonsynonymous site). For d_N, d_S, and ω of the remaining orthologs, we constructed 95% CIs by performing 10,000 bootstrap resamplings of tissue-specific means without assuming normality. We tabulated the number of genes under positive selection ($\omega > 1$) between tissues classes. Additionally, we compared evolutionary rates among putatively secreted proteins for each tissue by predicting the presence of a signal peptide sequence in SignalP 4.1 using default parameters (Petersen et al. 2011). Because the signal peptide could have been truncated in genes that were only partially recovered in our transcriptome assembly, only those genes with ortholog hit ratios > 0.80 were used for the "secreted" analysis.

To confirm that results were not driven by lowly expressed genes, we also generated a conservative list of differentially expressed genes that restricted the tissue-enriched designation to those genes with an average of at least log 2-counts per million and log 3-fold change in the tissue of interest. We repeated the quail–chicken pairwise evolutionary rate analysis on this group of genes.

Lineage-Specific Genes

We investigated the incidence of lineage-specific genes by calculating the proportion of genes from each tissue class 1) without BLAST hits to chicken, 2) with no clear orthologs

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

in chicken, and 3) that are potential quail-specific duplicates. Ortholog assignments were performed as mentioned earlier, restricting the gene lists to those from the filtered transcriptome with ortholog hit ratio $\geq$0.50 (calculated from the chicken ortholog). To identify putative duplicates, we determined paralogs from the exhaustive transcriptome with OrthoMCL using default parameters (Li 2003; Chen et al. 2006). OrthoMCL combines a reciprocal best BLAST approach with a graph-clustering algorithm to identify homologous proteins and distinguish potential orthologs from paralogs. We used the "exhaustive" transcriptome in OrthoMCL to ensure that the best BLAST hit was identified. We restricted the list of genes considered a possible lineage-specific duplicate to the "co-orthologs" and "inparalogs" identified by OrthoMCL that are also found in the filtered transcriptome and, for the co-orthologs, that have the same best BLAST hit to chicken in our analyses. Because we are designating duplicates based on a de novo assembled transcriptome, some alternative transcripts and alleles may be included in this data set despite the fact that our transcriptome assembly and processing attempted to remove them.

Partial Correlation of Evolutionary Rate, Expression Level, and Enrichment

Spearman's ρ was used to estimate partial correlations between quail:chicken pairwise evolutionary rates (ω), enrichment, and expression level for each tissue class. Enrichment is the log fold change of a gene expression level in one tissue compared with the average of the other as described under "Tissue-Specific Gene Expression". For each gene, expression levels were estimated as RPKM (Mortazavi et al. 2008). Partial correlations were computed from the pairwise correlation coefficients in the R packages corpcor version 1.6.5 (Schäfer and Strimmer 2005). The 95% CIs of each partial correlation were approximated by bootstrapping the data 1,000 times with the R package boot version 1.3.7.

Lineage-Specific Shifts in Evolutionary Rates

To test for variation in selective pressure along a specific lineage, we compared "branch" models in a maximum likelihood framework implemented in the CODEML program of the software package PAML version 4.7 (Yang 2007). First, we identified 1:1:1:1 orthologs between quail: chicken: turkey: and zebra finch. Ortholog determination and sequence alignment were performed as described earlier (N = 5,281 orthologous groups). CODEML was used to compare the likelihood of null models where the intensity of selection (ω) was the same along all branches in the tree (i.e., one-ratio models; M0) to two alternative models (M2): one where the quail terminal lineage has one ω (ω_{quail}) and the rest of the tree has another (ω_{tree}), and one where the chicken terminal lineage has one ω ($\omega_{chicken}$) and the rest of the tree has another (ω_{tree}). We estimated ω along the phylogeny (((quail, chicken), turkey), zebra finch) with Galliformes relationships deduced from the well-supported, multilocus phylogeny of Kimball and Braun (2008). Significance was determined by likelihood ratio tests fitted to χ^2 distributions. In 88 cases where the same gene was identified as having a significant shift in selective pressure along both the quail and chicken branches, we chose the most likely scenario according to log likelihood scores. We labeled genes according to their expression pattern in quail.

We compared the proportion of genes with significant shifts in ω for a particular tissue's gene set to the transcriptome-wide expectation from either quail or chicken. Our comparisons were to species-specific transcriptome-wide expectations, rather than between species directly, because differences in effective population sizes between species can distort the rate of adaptive evolution in reproductive proteins (e.g., Good et al. 2013). The subset of genes from the tissue in question was subtracted from the transcriptome-wide numbers prior to comparison. To examine whether the magnitude of shifts varied across tissues, we regressed lineage-specific ω estimates on ω values generated for the rest of the tree and tested whether the residuals from this model differed significantly according to tissue designation or species. We constructed 95% CIs by 10,000 bootstrap resamplings of tissue-specific means without assuming normality.

To explore the function of genes possibly under positive selection in the quail, we annotated genes with significant accelerations in ω_{quail} versus ω_{tree} and with $\omega_{quail} > 1$. For these genes, we determined spatial expression of the *G. gallus* ortholog in the chicken based on either Unigene expressed sequence tag data (http://www.ncbi.nlm.nih.gov/unigene, last accessed September 16, 2014) or RNA-Seq data from Ensembl (http://useast.ensembl.org/index.html, last accessed September 16, 2014).

Real-Time Quantitative PCR Validation of RNA-Seq

We performed RT-qPCR on the 18 RNA samples (six males $\times$ three tissues) used for RNA-Seq. RNA was treated with Turbo DNase (Ambion) and confirmed to be free of genomic DNA by attempting to PCR amplify a panel of three housekeeping genes (see supplementary table S3, Supplementary Material online) directly from RNA. Two hundred nanograms of RNA was reverse transcribed into cDNA with SuperScript III First Strand cDNA Synthesis Kit (Invitrogen) following the manufacturer's instructions. The primers pairs were designed from three genes significantly enriched in each tissue (nine genes in total). We verified that primers amplified the intended target by Sanger sequencing. β-Actin served as an internal control and was confirmed to be stable across treatments using Normfinder (Andersen et al. 2004). RT-qPCR reactions (25 μl) were performed in duplicate with 33 ng of cDNA template and 200 nM of each primer using the Power SYBR Green Master Mix (Applied Biosystems). Samples were run on a ViiA 7 (Applied Biosystems) thermocycler with the following parameters: 95 °C for 10 min, 40 cycles of 95 °C for 15 s and 60 °C for 60s. Primer efficiencies were calculated with real-time PCRMiner and ranged from 95 to 100% (Zhao and Fernald 2005). To compare expression levels among tissues, we followed the $2^{-\Delta CT}$ method as described by Schmittgen and Livak (2008). We generated 95% CIs around these values for each tissue to ascertain whether genes were

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

significantly enriched in the expected tissues. Primer sequences, target sizes, efficiencies, and log fold changes are given in supplementary table S3, Supplementary Material online.

Data Analysis

All analyses were performed in R version 2.15.2 (R Core Team 2012, last accessed July 20, 2014) unless otherwise stated. CIs were generated in the Hmisc package of R (Harrell 2014, last accessed July 20, 2014) unless otherwise stated. Significantly different proportions were determined with Pearson's χ^2 test without Yates' continuity correction. Where appropriate, the false discovery rate was applied at a cutoff of 0.05 to correct for multiple tests (Benjamini and Hochberg 1995). BLAST steps were performed in parallel via Cornell University's Computational Biology Application Suite for High Performance Computing (biohpc.org, last accessed September 16, 2014). Venn diagrams were made in the VennDiagram version 1.5.1 package in R (Chen and Boutros 2011). Unless otherwise stated, bioinformatics analyses were performed on a Linux, Dell Precision T3500n with four cores, 24-GB RAM, and 4 TB HDD housed at the Cornell University's Institute of Biotechnology's Bioinformatics Facility.

Supplementary Material

Supplementary figures S1–S6, method S1 and tables S1–S5 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors thank R. Meisel, B. Lazzaro, N. Chen, and the Harrison lab for helpful discussions and advice; E. Adkins-Regan for quail samples and housing; M. Mouchka, R. Bukowski and Q. Sun for assistance with transcriptome assembly and RNA-Seq data analysis; S. Bogdanowicz, J. Mosher, and P. Schweitzer for aiding with cDNA library preparation and Illumina sequencing; S. Iacovelli and N. Baran for quail handling; T. van Deusen, S. Martin, L. Vann, and P. Smith for quail care; and two anonymous reviewers for comments on a previous version of this article. This work was supported by the National Science Foundation (DDIG DEB-1010757 to F.R.F. and R.G.H.), and P.E.O. Scholar, Cornell Sigma Xi, Andrew W. Mellon, and Paul. F. Feeny awards to F.R.F.

References

Adkins-Regan E. 1995. Predictors of fertilization in the Japanese quail, *Coturnix japonica*. *Anim Behav.* 50:1405–1415.

Adkins-Regan E. 1999. Foam produced by male *Coturnix* quail: what is its function? *Auk* 116:184–193.

Almeida F, DeSalle R. 2008. Evidence of adaptive evolution of accessory gland proteins in closely related species of the *Drosophila repleta* group. *Mol Biol Evol.* 25:2043–2053.

Andersen CL, Jensen JL, Ørntoft TF. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 64:5245–5250.

Anderson MJ, Dixson AF. 2002. Sperm competition: motility and the midpiece in primates. *Nature* 416:496.

Andrés JA, Maroja LS, Bogdanowicz SM, Swanson WJ, Harrison RG. 2006. Molecular evolution of seminal proteins in field crickets. *Mol Biol Evol.* 23:1574–1584.

Andrés JA, Maroja LS, Harrison RG. 2008. Searching for candidate speciation genes using a proteomic approach: seminal proteins in field crickets. *Proc R Soc Lond B Biol Sci.* 275:1975–1983.

Arunkumar R, Josephs EB, Williamson RJ, Wright SI. 2013. Pollen-specific, but not sperm-specific, genes show stronger purifying selection and higher rates of positive selection than sporophytic genes in *Capsella grandiflora*. *Mol Biol Evol.* 30:2475–2486.

Bailey NW, Veltsos P, Tan Y-F, Millar AH, Ritchie MG, Simmons LW. 2013. Tissue-specific transcriptomics in the field cricket *Teleogryllus oceanicus*. *G3* 3:225–230.

Baker RH, Narechania A, Johns PM, Wilkinson GS. 2012. Gene duplication, tissue-specific gene expression and sexual conflict in stalk-eyed flies (Diopsidae). *Philos Trans R Soc B.* 367:2357–2375.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol.* 57:289–300.

Berlin S, Qu L, Ellegren H. 2008. Adaptive evolution of gamete-recognition proteins in birds. *J Mol Evol.* 67:488–496.

Berlin S, Smith NGC. 2005. Testing for adaptive evolution of the female reproductive protein ZPC in mammals, birds and fishes reveals problems with the M7-M8 likelihood ratio test. *BMC Evol Biol.* 5:65.

Bork P, Koonin EV. 1998. Predicting functions from protein sequences—where are the bottlenecks? *Nat Genet.* 18:313–318.

Calkins JD, El-Hinn D, Swanson WJ. 2007. Adaptive evolution in an avian reproductive protein: ZP3. *J Mol Evol.* 65:555–563.

Carnahan-Craig SJ, Jensen-Seaman MI. 2013. Rates of evolution of hominoid seminal proteins are correlated with function and expression, rather than mating system. *J Mol Evol.* 78:87–99.

Carroll S. 2005. Evolution at two levels: on genes and form. *Plos Biol.* 3:1159–1166.

Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134:25–36.

Chen F, Mackey AJ, Stoeckert CJ, Roos DS. 2006. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res.* 34:D363–D368.

Chen H, Boutros PC. 2011. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* 12:35.

Cheng KM, Hickman AR, Nichols CR. 1989. Role of the proctodeal gland foam of male Japanese quail in natural copulations. *Auk* 106:279–285.

Cheng KM, McIntyre RF, Hickman AR. 1989. Proctodeal gland foam enhances competitive fertilization in domestic Japanese quail. *Auk* 106:286–291.

Clark NL, Aagaard JE, Swanson WJ. 2006. Evolution of reproductive proteins from animals and plants. *J Reprod Fertil.* 131:11–22.

Clark NL, Swanson W. 2005. Pervasive adaptive evolution in primate seminal proteins. *PLoS Genet.* 1:335–342.

Dean MD, Clark NL, Findlay GD, Karn RC, Yi X, Swanson WJ, MacCoss MJ, Nachman MW. 2009. Proteomics and comparative genomic investigations reveal heterogeneity in evolutionary rate of male reproductive proteins in mice (*Mus domesticus*). *Mol Biol Evol.* 26:1733–1743.

Dean MD, Good JM, Nachman MW. 2008. Adaptive evolution of proteins secreted during sperm maturation: an analysis of the mouse epididymal transcriptome. *Mol Biol Evol.* 25:383–392.

Dean MD, Nachman MW. 2009. Faster fertilization rate in conspecific versus heterospecific matings in house mice. *Evolution* 63:20–28.

Dean MD. 2013. Genetic disruption of the copulatory plug in mice leads to severely reduced fertility. *PLoS Genet.* 9:e1003185.

Dorus S, Busby SA, Gerike U, Shabanowitz J, Hunt DF, Karr TL. 2006. Genomic and functional evolution of the *Drosophila melanogaster* sperm proteome. *Nat Genet.* 38:1440–1445.

Dorus S, Evans PD, Wyckoff GJ, Choi SS, Lahn BT. 2004. Rate of molecular evolution of the seminal protein gene SEMG2 correlates with levels of female promiscuity. *Nat Genet.* 36:1326–1329.

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

Dorus S, Wasbrough ER, Busby J, Wilkin EC, Karr TL. 2010. Sperm proteomics reveals intensified selection on mouse sperm membrane and acrosome genes. *Mol Biol Evol.* 27: 1235–1246.

Drummond DA, Bloom JD, Adami C, Wilke CO, Arnold FH. 2005. Why highly expressed proteins evolve slowly. *Proc Natl Acad Sci U S A.* 102:14338–14343.

Duret L, Mouchiroud D. 2000. Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. *Mol Biol Evol.* 17:68–70.

Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet.* 8:689–698.

Findlay GD, Yi X, MacCoss MJ, Swanson WJ. 2008. Proteomics reveals novel *Drosophila* seminal fluid proteins transferred at mating. *Plos Biol.* 6:e178.

Finn S, Civetta A. 2010. Sexual selection and the molecular evolution of ADAM proteins. *J Mol Evol.* 71:231–240.

Finseth FR, Iacovelli SR, Harrison RG, Adkins-Regan EK. 2013. A nonsemen copulatory fluid influences the outcome of sperm competition in Japanese quail. *J Evol Biol.* 26:1875–1889.

Finseth FR, Harrison RG. Forthcoming. A comparison of next generation sequencing technologies for transcriptome assembly and utility of RNA-Seq in a non-model bird. *Plos ONE*. Doi: 10.1371/journal.pone.0108550.

Fujihara N. 1992. Accessory reproductive fluids and organs in male domestic birds. *Worlds Poult Sci J.* 48:39–56.

Good JM, Nachman MW. 2005. Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis. *Mol Biol Evol.* 22:1044–1052.

Good JM, Wiebe V, Albert FW, Burbano HA, Kircher M, Green RE, Halbwax M, Andre C, Atencia R, Fisher A, et al. 2013. Comparative population genomics of the ejaculate in humans and the great apes. *Mol Biol Evol.* 30:964–976.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 29:644–652.

Grassa CJ, Kulathinal RJ. 2011. Elevated evolutionary rates among functionally diverged reproductive genes across deep vertebrate lineages. *Int J Evol Biol.* 2011:1509–1514.

Harrell FE Jr., with contributions from C Dupont and many others. 2014. Hmisc: Harrell Miscellaneous. R package version 3.14-4. Available from: http://CRAN.R-project.org/package=Hmisc.

Herlyn H, Zischler H. 2007. Sequence evolution of the sperm ligand zonadhesin correlates negatively with body weight dimorphism in primates. *Evolution* 61:289–298.

Hornett EA, Wheat CW. 2012. Quantitative RNA-Seq analysis in non-model species: assessing transcriptome assemblies as a scaffold and the utility of evolutionary divergent genomic reference species. *BMC Genomics* 13:361.

Hosken DJ. 1998. Testes mass in megachiropteran bats varies in accordance with sperm competition theory. *Behav Ecol Sociobiol.* 44: 169–177.

Hurle B, Swanson W, NISC Comparative Sequencing Program,Green ED. 2007. Comparative sequence analyses reveal rapid and divergent evolutionary changes of the WFDC locus in the primate lineage. *Genome Res.* 17:276–286.

Ikeda K, Taji K. 1954. On the foamy ejaculate of Japanese quail, *Coturnix coturnix japonica*. *Scientif. Rep. Matsuyama Agric Coll.* 13:1–4.

Kelleher ES, Swanson WJ, Markow TA. 2007. Gene duplication and adaptive evolution of digestive proteases in *Drosophila arizonae* female reproductive tracts. *PLoS Genet.* 3:e148.

Khaitovich P. 2005. Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* 309: 1850–1854.

Kimball RT, Braun EL. 2008. A multigene phylogeny of Galliformes supports a single origin of erectile ability in non-feathered facial traits. *J Avian Biol.* 39:438–445.

King AS. 1981. Cloaca. In: King AS, McLelland J, editors. Form and function in birds, Vol. 2. New York: Academic Press. p. 63–105.

Klemm R, Knight CE, Stein S. 1973. Gross and microscopic morphology of glandula-proctodealis (foam gland) of *Coturnix c. japonica* (aves). *J Morphol.* 141:171–184.

Koonin EV. 2005. Orthologs, paralogs, and evolutionary genomics. *Annu Rev Genet.* 39:309–338.

Korn N, Thurston R, Pooser B, Scott T. 2000. Ultrastructure of spermatozoa from Japanese quail. *Poultry Sci.* 79:407–414.

Kumar S, Blaxter ML. 2010. Comparing de novo assemblers for 454 transcriptome data. *BMC Genomics* 11:571.

Lake P. 1981. Male genital organs. In: King AS, McLelland J, editors. Form and function in birds, Vol. 2. London: Academic Press. p. 1–61.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.

Larracuente AM, Sackton TB, Greenberg AJ, Wong A, Singh ND, Sturgill D, Zhang Y, Oliver B, Clark AG. 2008. Evolution of protein-coding genes in *Drosophila*. *Trends Genet.* 24:114–123.

Lemos B, Bettencourt BR, Meiklejohn CD, Hartl DL. 2005. Evolution of proteins and gene expression levels are coupled in *Drosophila* and are independently associated with mRNA abundance, protein length, and number of protein-protein interactions. *Mol Biol Evol.* 22:1345–1354.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079.

Li L. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13:2178–2189.

Liao BY. 2006. Impacts of gene essentiality, expression pattern, and gene compactness on the evolutionary rate of mammalian proteins. *Mol Biol Evol.* 23:2072–2080.

Long M, Betrán E, Thornton K, Wang W. 2003. The origin of new genes: glimpses from the young and old. *Nat Rev Genet.* 4:865–875.

Marks H, Lepore P. 1965. A procedure for artificial insemination of Japanese quail. *Poultry Sci.* 44:1001–1003.

Marshall JL, Huestis DL, Garcia C, Hiromasa Y, Wheeler S, Noh S, Tomich JM, Howard DJ. 2010. Comparative proteomics uncovers the signature of natural selection acting on the ejaculate proteomes of two cricket species isolated by postmating, prezygotic phenotypes. *Mol Biol Evol.* 28:423–435.

Martin-Coello J, Dopazo H, Arbiza L, Ausio J, Roldan ERS, Gomendio M. 2009. Sexual selection drives weak positive selection in protamine genes and high promoter divergence, enhancing sperm competitiveness. *Proc R Soc Lond B Biol Sci.* 276:2427–2436.

McCarthy DJ, Chen Y, Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40:4288–4297.

McFarland LZ, Mather FB, Warner RL, Wilson WO. 1968. Cloacal gland complex of Japanese quail. *Experientia* 24:941–943.

Meisel RP. 2011. Towards a more nuanced understanding of the relationship between sex-biased gene expression and rates of protein-coding sequence evolution. *Mol Biol Evol.* 28:1893–1900.

Min XJ, Butler G, Storms R, Tsang A. 2005. OrfPredictor: predicting protein-coding regions in EST-derived sequences. *Nucleic Acids Res.* 33:W677–W680.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods.* 5:621–628.

Nichols CR. 1991. A comparison of the reproductive and behavioural differences of feral and domestic Japanese quail [masters thesis]. Vancouver (BC): University of British Columbia.

Nuzhdin SV, Wayne ML, Harmon KL, McIntyre LM. 2004. Common pattern of evolution of gene expression level and protein sequence in *Drosophila*. *Mol Biol Evol.* 21:1308–1317.

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022

O'Connor TD, Mundy NI. 2009. Genotype–phenotype associations: substitution models to detect evolutionary associations between phenotypic variables and genotypic evolutionary rate. *Bioinformatics* 25: i94–i100.

O'Neil ST, Dzurisin JD, Carmichael RD, Lobo NF, Emrich SJ, Hellman JJ. 2010. Population-level transcriptome sequencing of nonmodel organisms *Erynnis propertius* and *Papilio zelicaon*. *BMC Genomics* 11:310.

Panhuis TM, Clark NL, Swanson WJ. 2006. Rapid evolution of reproductive proteins in abalone and *Drosophila*. *Philos Trans R Soc B.* 361: 261–268.

Parsch J, Ellegren H. 2013. The evolutionary causes and consequences of sex-biased gene expression. *Nat Rev Genet.* 14:83–87.

Pál C, Papp B, Hurst LD. 2001. Highly expressed genes in yeast evolve slowly. *Genetics* 158:927–931.

Perry MM. 1987. Nuclear events from fertilisation to the early cleavage stages in the domestic fowl (*Gallus domesticus*). *J Anat.* 150: 99–109.

Petersen TN, Brunak S, Heijne von G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods.* 8:785–786.

Pitnick S, Wolfner MF, Suarez SS. 2008. Ejaculate–female and sperm–female interactions. In: Birkhead TR, Hosken DJ, Pitnick S, editors. Sperm biology. London: Academic Press. p. 247–304.

Poiani A. 2006. Complexity of seminal fluid: a review. *Behav Ecol Sociobiol.* 60:289–310.

R Core Team. 2012. R: A language and environment for statistical computing R Foundation for Statistical Computing. (Vienna) Austria. [cited 2014 Sep 16]. Available from: http://www.R-project.org/.

Ramm SA, Oliver P, Ponting C, Stockley P, Emes R. 2008. Sexual selection and the adaptive evolution of mammalian ejaculate proteins. *Mol Biol Evol.* 25:207–219.

Ramm SA, Stockley P. 2010. Sperm competition and sperm length influence the rate of mammalian spermatogenesis. *Biol Lett.* 6: 219–221.

Rands CM, Darling A, Fujita M, Kong L, Webster MT, Clabaut C, Emes RD, Heger A, Meader S, Hawkins MB, et al. 2013. Insights into the evolution of Darwin's finches from comparative analysis of the *Geospiza magnirostris* genome sequences. *BMC Genomics* 14:95.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140.

Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11:R25.

Schäfer J, Strimmer K. 2005. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat Appl Genet Mol Biol.* 4:32.

Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative C_T method. *Nat Prot.* 3:1101–1108.

Schumacher J, Rosenkranz D, Herlyn H. 2013. Mating systems and protein-protein interactions determine evolutionary rates of primate sperm proteins. *Proc R Soc Lond B Biol Sci.* 281:20132607–20132607.

Seiwert C, Adkins-Regan E. 1998. The foam production system of the male Japanese quail: characterization of structure and function. *Brain Behav Evol.* 52:61–80.

Singh RP, Sastry KVH, Shit N, Pandey NK, Singh KB, Mohan J, Moudgal RP. 2011. Cloacal gland foam enhances motility and disaggregation of spermatozoa in Japanese quail (*Coturnix japonica*). *Theriogenology* 75:563–569.

Singh RP, Sastry KVH, Pandey NK, Shit N, Agrawal R, Singh KB, Mohan J, Saxena VK, Moudgal RP. 2011. Characterization of lactate dehydrogenase enzyme in seminal plasma of Japanese quail (*Coturnix coturnix japonica*). *Theriogenology* 75:555–562.

Singh RP, Sastry KVH, Pandey NK, Singh KB, Malecki IA, Farooq U, Mohan J, Saxena VK, Moudgal RP. 2012. The role of the male cloacal gland in reproductive success in Japanese quail (*Coturnix japonica*). *Reprod Fertil Dev.* 24:405.

Sirot LK, Laflamme BA, Sitnik JL, Rubinstein CD, Avila FW, Chow CY, Wolfner MF. 2009. Molecular social interactions: *Drosophila melanogaster* seminal fluid proteins as a case study. *Adv Genet.* 68:23–56.

Snook RR, Hosken DJ, Karr TL. 2011. The biology and evolution of polyspermy: insights from cellular and functional studies of sperm and centrosomal behavior in the fertilized egg. *Reproduction* 142: 779–792.

Swanson W, Clark AG, Waldrip-Dail H, Wolfner MF, Aquadro C. 2001. Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proc Natl Acad Sci U S A.* 98:7375–7379.

Swanson W, Vacquier V. 2002a. Reproductive protein evolution. *Annu Rev Ecol Syst.* 33:161–179.

Swanson W, Vacquier V. 2002b. The rapid evolution of reproductive proteins. *Nat Rev Genet.* 3:137–144.

Suyama M, Torrents D, Bork P. 2006. PAL2NAL:robust conversion of protein sequence alignments into the corresponding codon alignments. *Nuc Acids Res.* 34:W609–W612.

Tatusov RL. 1997. A genomic perspective on protein families. *Science* 278:631–637.

Teijeiro JR, Puigcerver M, Gallego S. 2003. Pair bonding and multiple paternity in the polygamous common quail *Coturnix coturnix*. *Ethology* 109:291–302.

Turner L, Chuong E, Hoekstra H. 2008. Comparative analysis of testis protein evolution in rodents. *Genetics* 179:2075–2089.

Turner LM, Hoekstra HE. 2008. Causes and consequences of the evolution of reproductive proteins. *Int J Dev Biol.* 52:769–780.

Wagstaff BJ. 2005. Molecular population genetics of accessory gland protein genes and testis-expressed genes in *Drosophila mojavensis* and *D. arizonae*. *Genetics* 171:1083–1101.

Walters JR, Harrison RG. 2010. Combined EST and proteomic analysis identifies rapidly evolving seminal fluid proteins in *Heliconius* butterflies. *Mol Biol Evol.* 27:2000–2013.

Walters JR, Harrison RG. 2011. Decoupling of rapid and adaptive evolution among seminal fluid proteins in *Heliconius* butterflies with divergent mating systems. *Evolution* 65:2855–2871.

Wong A. 2010. Testing the effects of mating system variation on rates of molecular evolution in primates. *Evolution* 64: 2779–2785.

Wong A. 2011. The molecular evolution of animal reproductive tract proteins: what have we learned from mating-system comparisons? *Int J Evol Biol.* 2011:1–9.

Woolley DM. 1995. The structure of the spermatozoon of the Japanese quail, *Coturnix coturnix* L., var. *japonica*. *Acta Zool.* 76:45–50.

Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet.* 8:206–216.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 24:1586–1591.

Zhang L, Li W-H. 2004. Mammalian housekeeping genes evolve more slowly than tissue-specific genes. *Mol Biol Evol.* 21:236–239.

Zhang Z, Li J, Zhao X-Q, Wang J, Wong GK-S, Yu J. 2006. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. *Genom Proteom Bioinformatics.* 4:259–263.

Zhang Z, Xiao J, Wu J, Zhang H, Liu G, Wang X, Dai L. 2012. ParaAT: a parallel tool for constructing multiple protein-coding DNA alignments. *Biochem Biophys Res Commun.* 419:779–781.

Zhao S, Fernald RD. 2005. Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J Comput Biol.* 12:1047–1064.

Downloaded from https://academic.oup.com/mbe/article/31/12/3266/2925697 by guest on 16 August 2022