

# Conservation and divergence in the frog immunome: pyrosequencing and de novo assembly of immune tissue transcriptomes



Anna E. Savage<sup>a,b,\*</sup>, Karen M. Kiemnec-Tyburczy<sup>b</sup>, Amy R. Ellison<sup>b</sup>, Robert C. Fleischer<sup>a</sup>, Kelly R. Zamudio<sup>b</sup>

<sup>a</sup> Center for Conservation and Evolutionary Genetics, Smithsonian Institution, Washington, DC 20013, United States

<sup>b</sup> Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, United States

## ARTICLE INFO

### Article history:

Received 1 November 2013

Received in revised form 6 March 2014

Accepted 26 March 2014

Available online 27 March 2014

### Keywords:

Anura

Neobatrachia

RNA-seq

Immune system

Functional genomics

## ABSTRACT

**Background:** Frogs are a diverse group of vertebrates for which limited genomic resources are available. Natural frog populations face a multitude of threats, including habitat degradation, infectious disease, and environmental change. Characterizing the functional genomics of anuran tissues in general – and the immune system in particular – will enhance our knowledge of genetic and epigenetic responses to environmental threats and inform conservation and recovery efforts.

**Results:** To increase the number of species with genomic datasets and characterize gene expression in immune-related tissues, we sequenced the transcriptomes of three tissues from two frogs (*Espadarana prosoblepon* and *Lithobates yavapaiensis*) on the Roche 454 GS FLX platform. Our sequencing produced 8881 *E. prosoblepon* and 5428 *L. yavapaiensis* annotated gene products after de novo assembly and Gene Ontology classification. Transcripts of the innate and acquired immune system were expressed in all three tissues. Inflammatory response and acquired immunity transcripts were significantly more diverged between *E. prosoblepon* and *L. yavapaiensis* compared to innate immunity and immune system development transcripts. Immune-related transcripts did not show an overall elevated rate of functional evolution, with the exception of glycosyl proteases, which include lysozymes, central bacterial and fungal-killing enzymes of the innate immune system.

**Conclusions:** The three frog transcriptomes provide more than 600 Mbp of new genomic data, and will serve as a valuable framework for future comparative studies of non-model anurans. Additionally, we show that immune gene divergence varies by functional group and that transcriptome studies can be useful in comparing rates of evolutionary change across gene families.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

Amphibians are a diverse group of vertebrates for which limited genomic resources are available (Calboli et al., 2011). Although over 7200 species of amphibians are recognized (Amphibiaweb 2014), genome-scale sequencing projects are limited by the intrinsic complexity of many amphibian genomes. For example, approximately 43 frog species across 12 families are known to be polyploid, as are numerous salamanders in the Ambystomatidae and Sirenidae families (Mable et al., 2011). Regardless of ploidy levels, amphibian genomes also rank among the largest within animal taxa (Gregory, 2005), thus they require considerably more sequencing to achieve the coverage needed for whole-genome assembly projects. Because of these limitations, most

amphibian genomic data – including EST databases, transcriptomes, and sequenced genomes – come from two model species: the Western clawed frog, *Silurana tropicalis* (Hellsten et al., 2010) and the African clawed frog, *Xenopus laevis* (e.g., Morin et al., 2006). Recently, a draft transcriptome assembly generated using both Sanger and high-throughput sequencing was released for a salamander, the axolotl *Ambystoma mexicanum* (<http://www.ambystoma.org/>). Although these newly developed resources are extremely useful for comparative genomics (e.g., Voss et al., 2011), amphibians are an ancient and diverse vertebrate lineage which has diverged over 350 million years ago (Roelants et al., 2007). Consequently, the genomic architecture and expression profiles of *A. mexicanum*, *S. tropicalis*, and *X. laevis* are likely not representative of amphibians as a whole. *S. tropicalis* and *X. laevis* are members of the family Pipidae, an early diverging group of anurans (Pyron and Wiens, 2011), yet 95% of extant frog species belong to the clade Neobatrachia, a much more recent radiation approximately 300 million years diverged from Pipidae (Igawa et al., 2008) that includes most of the evolutionary diversity within this group (Cannatella et al., 2008). Thus, genomic data from multiple species better representing diversity in the Anuran phylogenetic tree are necessary for comparative inferences about patterns and processes shaping amphibian genomes.

**Abbreviations:** aa, amino acid(s); Asn (N), asparagine; Asp (D), aspartic acid; *Bd*, *Batrachochytrium dendrobatidis*; BLAST, Basic Local Alignment Search Tool; Esp, *Espadarana prosoblepon*; gi, GenBank identification number; Glu (E), glutamic acid; Liya, *Lithobates yavapaiensis*; Lys (K), lysine; MHC, major histocompatibility complex.

\* Corresponding author at: Center for Conservation and Evolutionary Genetics, Smithsonian Institution, Washington, DC 20013, United States.

E-mail addresses: [savagea@si.edu](mailto:savagea@si.edu) (A.E. Savage), [kmk877@humboldt.edu](mailto:kmk877@humboldt.edu) (K.M. Kiemnec-Tyburczy), [are35@cornell.edu](mailto:are35@cornell.edu) (A.R. Ellison), [fleischerr@si.edu](mailto:fleischerr@si.edu) (R.C. Fleischer), [kelly.zamudio@cornell.edu](mailto:kelly.zamudio@cornell.edu) (K.R. Zamudio).

At present, amphibians are severely threatened by diverse factors including emerging infectious diseases, most notably chytridiomycosis caused by the fungus *Batrachochytrium dendrobatidis* (*Bd*) (Kilpatrick et al., 2010) and ranaviral infections caused by at least three DNA viruses in the family Iridoviridae (Chinchar, 2002). Understanding the functional genomics of the immune system, both in terms of conserved and novel features, has become a pressing need in light of these disease threats (Calboli et al., 2011). Although amphibians share many features of cellular immunity with mammals (reviewed in Rollins-Smith et al., 2009) much less is known about molecular immunity in this group of vertebrates. Characterization of immune structure and function in *Xenopus* reveals that the amphibian immune system is fundamentally similar to that of mammals (Du Pasquier et al., 1989). Frogs have a thymus where T cells differentiate and a spleen where B and T cells accumulate; they also have leukocytes such as neutrophils, basophils, eosinophils, monocytes, and macrophage-like cells, which are found in the blood (Robert and Ohta, 2009). However, frogs lack lymph nodes, Peyer's patches, and germinal centers, and consequently, frogs show poor affinity maturation, the process by which B cells make antibodies with increased affinity for antigens during the course of a mammalian immune response (Du Pasquier et al., 2000). Studies of *Silurana* and *Xenopus* provide the best evidence of immune system conservation and differentiation among vertebrate taxa, yet species in these genera are polyploid, making them less than ideal models for a broad understanding of frog immune function.

Immune genes can evolve rapidly under selection when compared to other regions of the genome (Crawford et al., 2010) due to powerful co-evolutionary arms races between molecular mechanisms of host defense and pathogen infectivity. Indeed, immune system genes show higher rates of adaptive evolution than non-immune genes across a range of taxa, including crustaceans (McTaggart et al., 2012), insects (Schlenke and Begun, 2003), birds (Downing et al., 2009), and primates (Fumagalli et al., 2011; Nielsen et al., 2005). However, the evolution of immune genes as a functional group has not been studied in the context of the broader amphibian genome, nor has it been characterized across amphibian taxonomic groups. The genetic diversity and expression of immune genes across amphibian families are of particular interest because they may provide insight into why different species are more resistant or susceptible to diseases (Richmond et al., 2009). In frogs, antimicrobial peptides (Duda et al., 2002) and class I (Kiemnec-Tyburczy et al., 2012; Teacher et al., 2009) and class II Major Histocompatibility Complex (MHC) genes (Hauswaldt et al., 2007; Kiemnec-Tyburczy et al., 2010; Savage and Zamudio, 2011) are some of the few immune genes that have been characterized and shown to be targets of strong diversifying selection. Antimicrobial peptides are well characterized across diverse anuran taxa and in the context of the fungal disease chytridiomycosis (Rollins-Smith et al., 2011). In contrast, fewer studies have focused on MHC genes, but the genome of *S. tropicalis* contains a "primordial" MHC organization (Ohta et al., 2006; Robert and Ohta, 2009), in which class I, II and III genes are linked, but gene order differs from mammals. Aside from these two gene families, however, comparative studies of the molecular underpinnings of amphibian immunity are lacking.

In this study, we focused on total gene expression in tissues with immune function in two New World frog species that have declined in their natural habitats from the fungal disease chytridiomycosis. The objectives of this study were threefold: (1) to increase the existing genomic resources for amphibians by sampling two divergent anuran families, (2) to characterize the full complement of genes expressed in immune tissues in species that are facing disease pressure, and (3) to compare evolutionary rates among different categories of immune genes and between immune and non-immune genes. To accomplish these goals, we sequenced transcriptomes of tissues with immune function in two non-model frog species from the families Ranidae and Centrolenidae and characterized gene expression and polymorphism with respect to each other and to the model frog *Silurana*.

## 2. Methods

### 2.1. Tissue collection

We sequenced three tissues with immune function using opportunistic tissue sampling of two frog species from two Neobatrachian families, the Ranidae and the Centrolenidae. These species are both targets for future work on chytridiomycosis-induced gene expression patterns, and the different tissues and variable number of individuals sampled arise from the uneven availability of field-collected preserved RNA from each species. All animals were euthanized via anesthesia with 10% lidocaine application to the ventral skin surface. Tissues were dissected using forceps and immediately submerged in 500  $\mu$ l of RNeasy Lysis Buffer (Invitrogen, Carlsbad, CA). All tissues were transported to Cornell University for extraction. Euthanasias were approved under Cornell University IACUC protocols for both species.

Intestinal wall tissue was collected from two adult male glass frogs (*Espadarana prosoblepon*, formerly *Centrolene prosoblepon*). This frog is a member of the family Centrolenidae, which contains 149 species distributed throughout Central and South America. This species, in particular, is a common member of frog communities in Costa Rica and Panama, where the specimens for this study were collected. One animal was collected from the province of Coclé (8.6681°N, –80.5920°W) and one from the province of Darien (7.762233°N, –77.72405°W), Panama, in July 2010. Drink patch skin (an area surrounding the vent that is important for osmotic balance) and whole spleen were collected for two male and two female adult lowland leopard frogs (*Lithobates yavapaiensis*, formerly *Rana yavapaiensis*) from four localities in Arizona, USA (UTMs: 270784E, 3894759 N; 554111E, 3640193 N; 344171E, 3754919 N; 570603E, 3570504 N) in December 2010. All sampled frogs were tested for the presence of the chytrid fungal pathogen *Bd* by swabbing the dorsum of each individual with a sterile swab and running *Bd*-specific quantitative real-time PCR assay on extracted swab DNA (Boyle et al., 2004; Hyatt et al., 2007).

### 2.2. RNA isolation and cDNA preparation

We extracted RNA from individuals separately using the Qiagen RNeasy Lipid Tissue Kit (Qiagen, Valencia, CA) according to the manufacturer's instructions. RNA was eluted in 40  $\mu$ l of ddH<sub>2</sub>O and stored at –80 °C until samples were prepared for pyrosequencing. Before library preparation, RNA quality was assessed by running the samples on an Agilent 2100 BioAnalyzer. Samples used for cDNA preparation all had RNA integrity numbers (RIN) higher than 9.0.

First strand cDNA was prepared using Superscript III first strand synthesis system (Invitrogen) with 12  $\mu$ M of two primers (5'-AAGCAGTG GTATCAACGCAGAGTATTTTCTTTTCTTTTCTTTTIVN-3' and 5'--AAGCAGT GGTATCAACGCAGAGTACGrGrG-3'), incubated with 100 ng of RNA at 65 °C for 3 min. cDNA was then amplified for 15–30 cycles using Platinum *Taq* (Invitrogen) following the manufacturer's protocol. Optimum cycle number was assessed by running an initial PCR and removing 3  $\mu$ l aliquots from a single 50  $\mu$ l PCR, sampling every other cycle for cycles 18 through 28 (18, 20, 22, 24, 26, 28) and running the subsamples on an agarose gel. A second PCR was then run with the empirically determined optimal number of cycles. The cDNA samples were then cleaned using a QIAquick PCR Purification Kit (Qiagen) and quantified using a Nanodrop ND-1000 spectrophotometer. The cleaned samples were normalized using a TRIMMER cDNA Normalization Kit (AXXORA, Farmingdale, NY) following standard protocols. Briefly, purified cDNA was incubated with hybridization buffer and a duplex-specific nuclease for 25 min at 68 °C. Normalized cDNA was purified with a QIAquick PCR Purification Kit and sonicated to approximately 700 bp using the appropriate module on a Covaris S2 Sonicator and visualized on an agarose gel. Ends of cDNA were polished by incubating with T4 DNA polymerase and T4 polynucleotide kinase for 25 °C for 45 min in a thermal cycler, then at 75 °C for 30 min. One  $\mu$ l *Taq* polymerase (New England Biolabs,

Ipswich, MA) was then added to the reaction and incubated at 72 °C for 45 min. Normalized cDNA was purified with a QIAquick PCR purification kit. We then ligated a unique 8 base pair barcode sequence and 454 adaptors A and B onto 500 ng of each cDNA sample. We removed small fragments using Ampure beads according to the NEBNext mRNA Sample Prep Reagent Set 2 and the GS FLX Titanium cDNA Rapid Library Preparation protocol.

### 2.3. Pyrosequencing, de novo assembly and data filtering

Normalized cDNA libraries from each of the three tissues were pooled across individuals at equimolar ratios and run on half a plate of 454 GS FLX with Titanium series chemistry (Roche Applied Science) at Cornell University's Life Sciences Core Laboratory Center. Sequences shorter than 50 nucleotides were excluded from analyses.

We used Newbler v. 2.6 (Roche Applied Science) to remove adapter sequences and perform de novo assembly of all *E. prosoblepon* reads combined and all *L. yavapaiensis* reads combined. Additionally, we generated assemblies for the *L. yavapaiensis* spleen-derived sequences and skin-derived sequences independently. We used default cDNA settings in gsAssembler, except that we modified the -icl parameter to only retain isotigs differing by more than 30 bp. We used CD-HIT-EST (Li and Godzik, 2006) to eliminate redundant contigs with 100% similarity. We then used custom perl scripts to trim short reads ( $\leq 30$  bp) and filter low-complexity reads (reads with  $\geq 45\%$  single or di-nucleotide repeats, allowing for single mismatches). We used PRINSEQ (PReprocessing and INformation of SEquences) v. 0.20.3 (Schmieder and Edwards, 2011) to generate summary statistics and assess sequence quality for each dataset.

### 2.4. Functional annotation and gene ontology analysis

BLAST2GO software v.2.3.1 (<http://www.blast2go.org> website) was used to perform several analyses of each de novo assembly. Contig sequences were searched via BLASTx against the NCBI non-redundant (nr) database, retaining up to ten hits with a minimum E-value of  $1 \times 10^{-6}$  and minimum bit score of 55. For gene ontology mapping (GO; <http://www.geneontology.org>), we used BLAST2GO to extract the GO terms associated with homologies identified with NCBI's QBLAST, producing a list of GO annotations represented as hierarchical categories of increasing specificity. We retained annotations with a minimum E-value of  $1 \times 10^{-6}$ , a minimum annotation cut-off of 55, and a GO weight of 5. We enhanced GO annotations using the annotation augmentation tool ANNEX (Myhre et al., 2006). Finally, we performed InterPro (InterProScan, EBI; Quevillon et al., 2005) searches remotely from BLAST2GO via the InterPro EBI web server and merged InterProScan GOs with the original GO annotations.

We identified significantly over- and under-represented biological processes using the Gossip package for statistical assessment of annotation differences between two sets of sequences (Blüthgen et al., 2005). This package uses a Fisher's Exact Test and corrects for multiple testing using the False Discovery Rate (FDR). We performed pairwise comparisons of annotated genes across individuals and tissues. All tests were two-tailed, and we removed duplicate genes ("remove double IDs" option) from both datasets before performing comparisons, thus our results were insensitive to the direction of the comparison.

### 2.5. Protein divergence analysis

Analyses of protein divergence within functional categories were adapted from Crawford et al. (2010). Putative 1:1 orthologs between *E. prosoblepon* and *L. yavapaiensis* annotated contigs were identified using a reciprocal best-hit analysis using BLASTN with an E-value threshold of  $1 \times 10^{-6}$ . For each contig with an ortholog, the open reading frame (ORF) was identified using a custom python script (<http://toolshed.g2.bx>

[psu.edu/view/peterjc/get\\_orfs\\_or\\_cds](http://psu.edu/view/peterjc/get_orfs_or_cds)). ORFs were determined as those giving the longest translation, flanked by methionine and a STOP codon (unless a longer translated region encompassed either the start or end of the sequence or both). Contigs with ORF less than 50 amino acids were discarded from further analysis. We used reciprocal best-hit BLASTP to ensure that equivalent translations were used in each species. Identity from the pairwise alignments of BLASTP searches (i.e., the proportion of amino acids that differed between orthologous pairs) was used as a measure of protein divergence. Automated sequence alignment can be unreliable at high divergence levels, so we excluded orthologous pairs with less than 30% identity to avoid false mismatches introduced by low confidence alignments. Lastly, we assigned each orthologous pair to one of three categories based on its level of divergence (i.e., the proportion of amino acids that differed): High ( $>0.2143$ ), Intermediate (0.0959 and 0.2143) and Low ( $<0.0959$ ) divergence, with bin cutoffs empirically determined so that one-third of the orthologous pairs fell into each category.

Orthologous pairs were assigned to functional categories based on their GO-slim terms. We then used a chi-squared test to identify whether each GO-Slim category was enriched (or depleted) for any of the divergence categories. As each pair could be assigned multiple GO-slim terms and the average number of terms assigned per pair within the divergence categories differed significantly (Low = 13.8, Intermediate = 10.6, High = 8.6; ANOVA  $F = 104.7$ ,  $P < 0.0001$ ), expected proportions used in chi-squared tests were adjusted accordingly (Low = 0.42, Intermediate = 0.32, High = 0.26). We used false discovery rate adjusted *P*-values to assign significance in tests (Benjamini and Hochberg, 1995). As the power of chi-square analyses increases with increasing number of observations, we limited our intra-functional category comparisons to categories populated by at least 15 contigs.

Protein divergence analyses were repeated with *S. tropicalis* orthologs compared to each of our target species individually. *S. tropicalis* protein sequences were downloaded from NCBI RefSeq database.

All orthologous gene pairs falling under the GO category "immune system process" (GO:0002376) were grouped into the following five immune function categories based on each pair's most specific GO term: innate immunity, cell-mediated acquired immunity, humoral acquired immunity, immune system development, or inflammation. GO categories that could not be unambiguously assigned to one of these groups (e.g., "positive regulation of immune response") were excluded from analysis. We performed this grouping separately for each of the pairwise species comparisons: *E. prosoblepon* and *L. yavapaiensis* (EL), *E. prosoblepon* and *S. tropicalis* (ES), and *L. yavapaiensis* and *S. tropicalis* (LS). We assessed statistically significant differences in average protein divergence among immune function groups for each species pair using Mann-Whitney U tests with Bonferroni correction (Shaffer, 1995) for all 30 comparisons (adjusted  $P < 0.0017$ ).

### 2.6. Single nucleotide polymorphisms

Single nucleotide polymorphisms (SNPs) were identified as variable sites across all individuals sampled per species meeting the following criteria: 1) minimum read coverage of 8, 2) average base call quality score of greater than 40, and 3) minor allele frequency of at least 30%. SNP calling was carried out with a customized version of the python pipeline 4Pipe4 (<https://github.com/StuntsPT/4Pipe4>). Only SNPs found in the best (longest) ORF of successfully annotated contigs were used for subsequent analyses. Chi-squared tests were used to identify any significant (after FDR correction for multiple tests) excess/deficit of synonymous or non-synonymous SNPs within each GO-slim category. We assessed statistically significant differences in total SNPs, synonymous SNPs and non-synonymous SNPs for immune system contigs compared to all other contigs using Mann-Whitney U tests with Bonferroni correction.

## 2.7. Lysozyme alignment and genealogy reconstruction

All lysozyme contigs recovered from *E. prosoblepon* and *L. yavapaiensis* and each of their top BLASTx hits in GenBank were translated to amino acids, and coding sequences were aligned using default parameters in CLUSTAL (Thompson et al., 1994), implemented in Geneious version 7.0.6 (available from <http://www.geneious.com/>). Bayesian trees were generated from the alignment with MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001) using no outgroup and parameters selected by hierarchical likelihood ratio tests with ProtTest version 3.8 (Abascal et al., 2005). MrBayes was run for 2,000,000 generations with four simultaneous Metropolis-coupled Monte Carlo Markov chains sampled every 100 generations. The first 25% of trees were discarded as burn-in with the remaining samples used to generate the consensus trees.

## 3. Results and discussion

### 3.1. Sequencing of three frog tissue transcriptomes

We generated a total of 275,490 *L. yavapaiensis* and 638,515 *E. prosoblepon* high quality FLX reads (Table 1). RNA-seq datasets have been accessioned in the EBI-ENA Sequence Read Archive (SRA) under accession numbers SAMN02630764 (*L. yavapaiensis* skin), SAMN02463986 (*L. yavapaiensis* spleen) and SAMN02630904 (*E. prosoblepon* intestine). In total, we generated over 600 Mb of frog transcriptome sequence. Average assembled contig length was 613 bp for the combined *L. yavapaiensis* tissues and 531 bp for *E. prosoblepon* intestine (Table 1); shorter average contig lengths in *E. prosoblepon* are due to the bimodal frequency distribution of contig lengths, with a peak of short (~100 bp) fragments not seen in the *L. yavapaiensis* dataset (Fig. S1). We found a strong positive relationship between contig length and the proportion of contigs showing significant homology to known genes (Fig. S1).

GC content was normally distributed for both species (Fig. S2), but mean GC content was slightly lower for *L. yavapaiensis* (40.8%) compared to *E. prosoblepon* (43.4%). *S. tropicalis* (version 4.1 assembly), which is currently the only annotated frog genome, has a genome-wide average GC content of 40.1%, which is only slightly lower than our transcriptome-derived values. This suggests little relationship between GC content and DNA expression, consistent with patterns recently detected in mammalian genomes (Sémon et al., 2005).

Overall, protein sequences of fewer than half the contigs showed significant homology to known genes at our BLASTx E-value minimum of  $1 \times 10^{-6}$  (Table 1), highlighting the lack of amphibian genomic resources present in genetic databases. *E. prosoblepon* intestine had the highest proportion of annotated contigs (42%), while *L. yavapaiensis* spleen had the lowest proportion of annotated contigs (30%), suggesting a positive relationship between total number of contigs and the proportion of annotated contigs. The majority of contigs had significant BLASTx hits to *Xenopus* or *Silurana* genes, with the remaining contigs showing significant homology to genes from other amphibian species or genes from a variety of sequenced mammalian and reptilian genomes (Fig. S3; Tables S1–S3).

The genome of *S. tropicalis* contains approximately 28,000 genes with an average transcript length of 1300 bp. The number of annotated genes per tissue (ranging from 2516 to 8881 per tissue) and average contig length (500–600 bp) recovered in our study species are somewhat lower than those from *S. tropicalis*. However, not all genes are

expected to be expressed in a given tissue and time point, and while insufficient sequencing depth cannot be ruled out, these results are not unexpected for de novo assembled transcriptomes of non-model organisms lacking a closely related reference genome (e.g., Liao et al., 2013; Looso et al., 2013; Wang et al., 2013; Yang et al., 2012).

GO terms were assigned to a total of 8881 *E. prosoblepon* and 5428 *L. yavapaiensis* contigs (Table 1). The higher number in *E. prosoblepon* likely reflects a higher depth of sequencing and not necessarily a higher number of expressed tissue transcripts, although transcriptome studies in human and mouse demonstrate different levels of transcriptome complexity across tissues, even within the same individual (Ramsköld et al., 2009). For each contig, the specific annotated terms were mapped to hierarchically more general parent terms falling into three ontology vocabularies: biological process, cellular component, and molecular function. Across the most general (level two) biological process categories, all three tissues were nearly identical in the proportion of annotated GO terms falling into each category (Fig. 1). A notable exception to this similarity was the absence in *L. yavapaiensis* spleen of the two biological processes “biological adhesion” and “growth” (Fig. 1B).

### 3.2. The Neobatrachian immunome

In total, 483 of *E. prosoblepon* intestine transcripts (2% of GO terms), 268 of *L. yavapaiensis* skin transcripts (2% of GO terms), and 299 of *L. yavapaiensis* spleen transcripts (3% of GO terms) were ascribed to the biological process parent term “immune system process” (Fig. 1; Table 2). All three tissues expressed transcripts coding for functions in both the innate and acquired branches of the immune system (Table 2).

In pairwise comparison testing for GO category enrichment across tissues and species, skin and spleen were not significantly different from each other for any GO category, but both *L. yavapaiensis* tissues were significantly enriched compared to *E. prosoblepon* for a number of GO terms (Fig. 2). This pattern may be attributed to several non-exclusive variables, including species, *Bd*-infection status, or innate similarities and differences across the sampled organs; however, we cannot discriminate among these possibilities with our current tissue sampling. One noteworthy pattern is the small proportion of immune contigs ascribed to B cell activation and leukocyte activation in *E. prosoblepon* intestine (2% and 6%, respectively) compared to *L. yavapaiensis* skin (9% and 26%) or spleen (9% and 26%; Table 2). In mammals, immune responses in the intestinal tract are suppressed to avoid responses against food antigens or commensal gut-associated bacteria (Dubois et al., 2005). Reduced leukocyte activation in *E. prosoblepon* intestine likely reflects a similar tolerogenic state in anuran gut tissue.

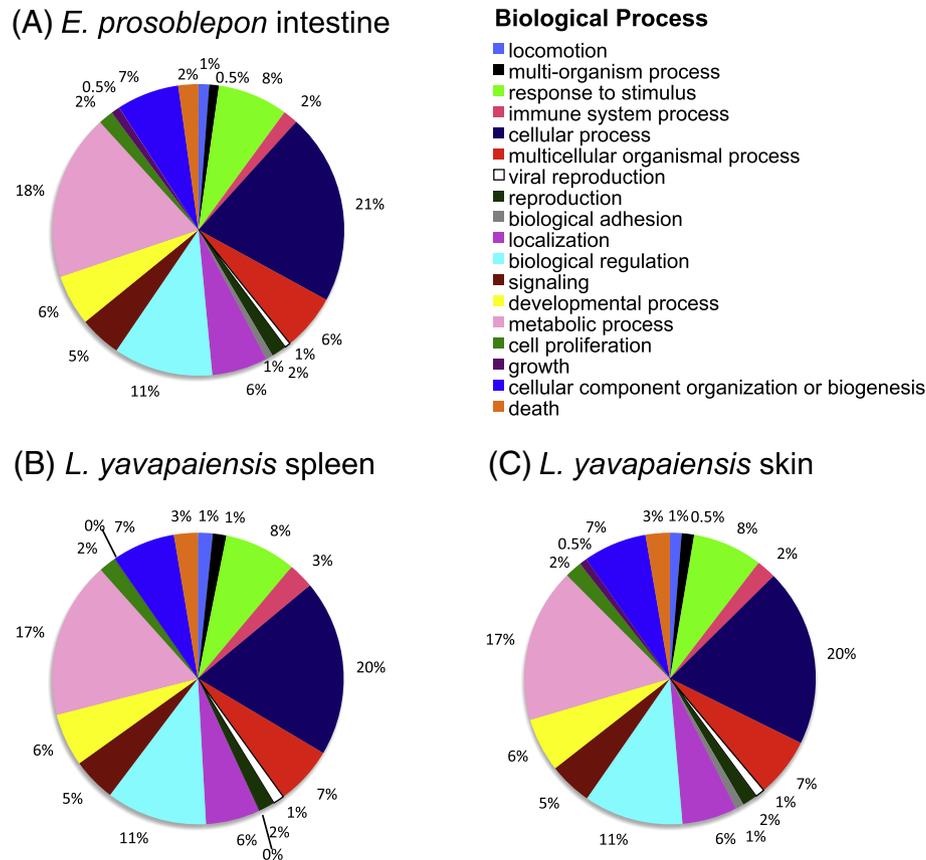
Skin was significantly enriched for more GO terms compared to intestine than spleen versus intestine for all three ontology vocabularies (Fig. 2A). However, for immune-related contigs, skin was significantly enriched for only 4 GO terms compared to intestine, and spleen was significantly enriched for 20 GO terms compared to intestine (Fig. 2), an almost 5-fold difference in GO term enrichment. Thus, while skin was the most differentiated tissue overall, spleen was the most enriched for immune function. The amphibian spleen is the major lymphatic tissue (Du Pasquier et al., 1989), thus enrichment for immune function is consistent with the known functions of this organ and can be viewed as a positive control for our whole-transcriptome characterizations.

As a measure of disease pressure that could skew the complement of expressed transcripts towards a particular type of immune response in

**Table 1**

Summary statistics for 454 pyrosequencing, assembly and annotation of normalized tissue transcriptomes from skin, spleen and intestine of two frog species.

Species	Tissue	No. high quality reads	No. assembled contigs	Average contig length (bp)	No. annotated contigs
<i>E. prosoblepon</i>	Intestine	638,515	21,227	531	8881 (42%)
<i>L. yavapaiensis</i>	Skin & spleen	275,490	16,374	615	5428 (33%)
<i>L. yavapaiensis</i>	Skin	167,176	10,993	513	3638 (33%)
<i>L. yavapaiensis</i>	Spleen	108,314	8349	472	2516 (30%)



**Fig. 1.** Gene Ontology analysis showing level two Biological Processes for A, *E. prosoblepon* intestine; B, *L. yavapaiensis* spleen; and C, *L. yavapaiensis* skin.

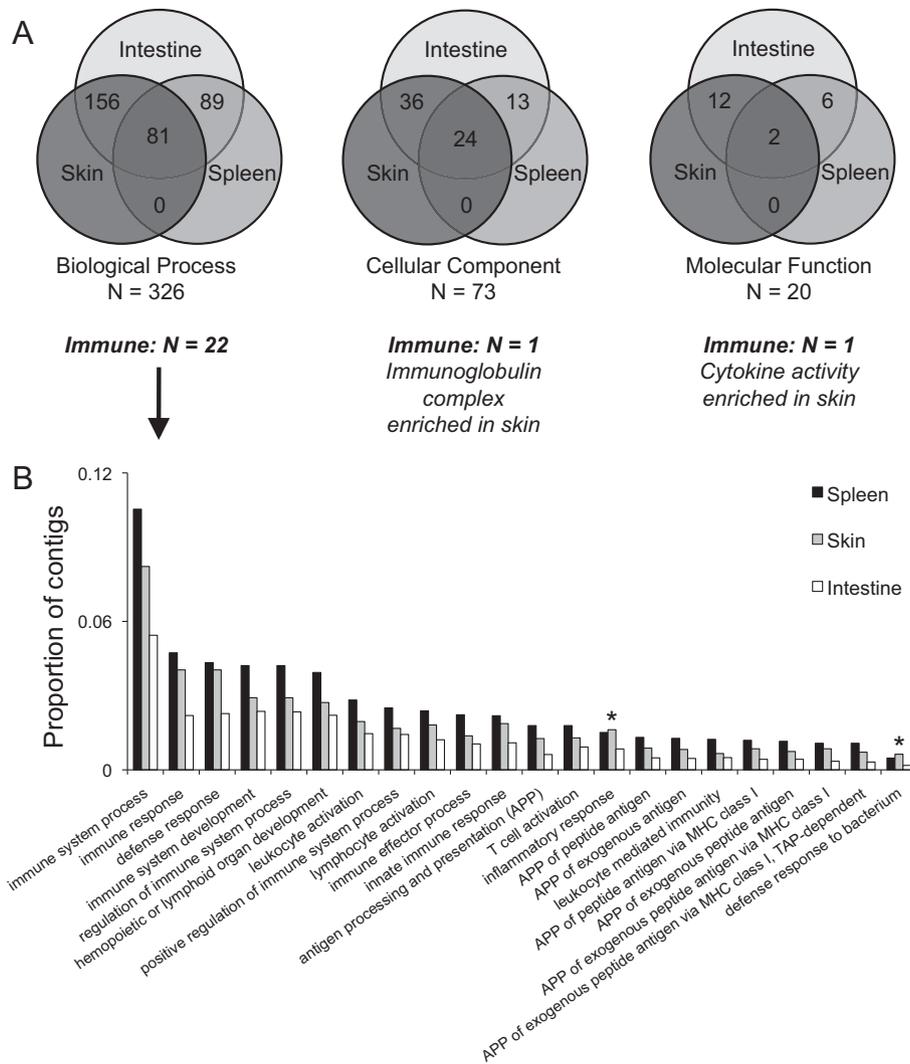
each sampled tissue, we tested for the presence of the chytridiomycosis-causing fungal pathogen *Bd* on the skin of the six sampled frogs. Neither of the *E. prosoblepon* individuals was infected with *Bd*, but three of the four *L. yavapaiensis* individuals were *Bd* positive and one of the frogs showed clinical signs of disease (lethargy, ventral redness, and loss of righting). As the *L. yavapaiensis* transcriptomes are pools of both infected and uninfected individuals, we cannot infer specific responses to *Bd* from our datasets. However, disease-related transcripts in *L. yavapaiensis* but not *E. prosoblepon* may be expressed in response to *Bd*, and we highlight these findings as a framework for future gene expression studies of amphibian disease.

Eight of the enriched GO terms in spleen related to antigen processing and presentation or T-cell activation, suggesting a role for MHC-based (T-cell mediated) immune responses in amphibian pathogen defense, as previously demonstrated for experimental chytridiomycosis infections in *L. yavapaiensis* (Savage and Zamudio, 2011). Of note, one of the two biological function GO terms significantly enriched in skin compared to intestine was “inflammatory response” (Fig. 2B). *Bd* infections are restricted to keratinized epidermal cells of adult frogs (Berger et al., 1998), thus recruitment of inflammatory cells into the skin is expected in animals mounting a robust immune response against *Bd*. To date, transcriptomic studies of chytridiomycosis dynamics have not

**Table 2**

Number of contigs annotated for immune system function across species and tissues. Percentage of immune processes devoted to each function is shown in parentheses. Espr = *Espadarana prosoblepon*; Liya = *Lithobates yavapaiensis*.

Immune system process	Intestine (Espr)	Skin (Liya)	Spleen (Liya)	Total (Espr & Liya)
Immune response	195 (40%)	147 (49%)	123 (46%)	465 (44%)
Immune system development	210 (43%)	113 (38%)	106 (40%)	429 (41%)
Regulation of immune process	208 (43%)	110 (37%)	106 (40%)	424 (40%)
Innate immune response	97 (20%)	68 (23%)	55 (21%)	220 (21%)
Immune effector process	93 (19%)	50 (17%)	56 (21%)	199 (19%)
Leukocyte activation	30 (6%)	79 (26%)	71 (26%)	180 (17%)
T cell activation	82 (17%)	47 (16%)	45 (17%)	174 (17%)
Inflammatory response	75 (16%)	59 (20%)	38 (14%)	172 (16%)
Antigen processing and presentation via MHC class I	38 (8%)	31 (10%)	30 (11%)	99 (9%)
Acquired immune response	34 (7%)	19 (6%)	22 (8%)	75 (7%)
B cell activation	10 (2%)	26 (9%)	23 (9%)	59 (6%)
Defense response to virus	26 (5%)	14 (5%)	15 (6%)	55 (5%)
Defense response to bacterium	16 (3%)	23 (8%)	12 (4%)	51 (5%)
Immunoglobulin production	8 (2%)	11 (4%)	13 (5%)	32 (3%)
Antigen processing and presentation via MHC class II	7 (1%)	3 (1%)	6 (2%)	11 (1%)
Defense response to fungus	3 (1%)	4 (1%)	1 (0.004%)	8 (1%)
All	483	299	268	1050



**Fig. 2.** Significant functional differentiation across tissue transcriptomes. A, Number of GO terms significantly enriched in at least one tissue and grouped by Biological Process, Cellular Component, and Molecular Function. No processes were significantly enriched in intestine; the number of GO terms common to all three tissues represents significant enrichment for both skin and spleen relative to intestine. B, Immune-related biological processes significantly enriched in skin relative to intestine (asterisks) and spleen relative to intestine (all other GO terms).

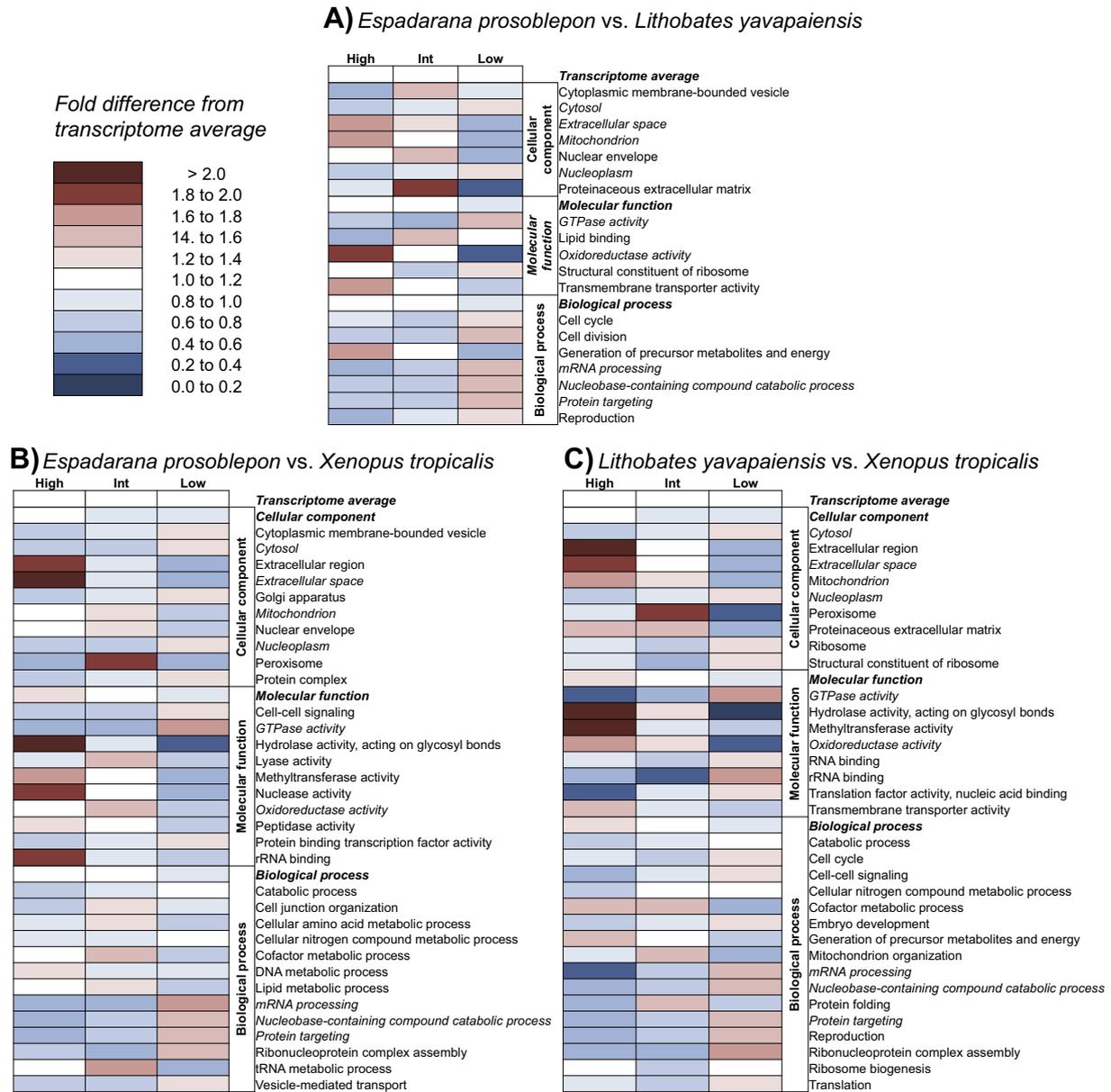
revealed strong epidermal inflammatory responses (Ribas et al., 2009; Rosenblum et al., 2009, 2012); however, only a handful of species have been studied in this context. Further genetic and morphological investigation of inflammatory responses to *Bd* in a range of affected species is necessary to infer the importance of skin immunity in dictating *Bd* resistance and susceptibility.

### 3.3. Protein divergence in anuran transcripts

We found a total of 2651 putative orthologous *E. prosoblepon* and *L. yavapaiensis* contigs (Table S4). Of the functional categories tested, seven cellular component categories, seven biological process categories and five molecular function categories showed significant deviations from expected proportions of high, intermediate and low amino acid (aa) divergence categories. We found 5229 orthologous contigs shared by *E. prosoblepon* and *S. tropicalis* (Table S5). Ten cellular component categories, 13 biological process categories and 10 molecular function categories showed significant deviations from expected proportions of high, intermediate and low divergence categories. Finally, 3570 orthologs were found between *L. yavapaiensis* and *S. tropicalis* protein sequences (Table S6). Nine cellular component categories, 16 biological process categories and 8 molecular function categories showed significant deviations from expected proportions of high, intermediate and low

divergence categories (Fig. 3). Nine GO categories showed significantly deviated proportions of divergence values for all three pairwise species comparisons (Fig. 3, italics). None of these categories is specific for immune system function; however, the cellular component category “extracellular space” (which showed a higher than expected proportion of high-divergence gene pairs in all species comparisons; Fig. 3) may be of particular relevance to disease resistance. The fungal pathogen *Bd* is known to degrade the amphibian extracellular matrix protein elastin (Moss et al., 2010), thus the enhanced divergence in extracellular matrix proteins may reflect functional adaptation in response to chytridiomycosis.

Populations of both *E. prosoblepon* and *L. yavapaiensis* have suffered negative fitness consequences due to chytridiomycosis, whereas natural populations of *Silurana* in Africa are asymptomatic and act as *Bd*-transmitting vectors (Weldon et al., 2004). Interestingly, we found a higher than expected proportion of high-divergence gene pairs for the molecular function category “Hydrolase activity, acting on glycosyl bonds” in both *E. prosoblepon* and *L. yavapaiensis* comparisons with *S. tropicalis* orthologs. Further, divergence values for these ortholog pairs were the highest values among all significant comparisons (2.50 in *L. yavapaiensis* and 2.38 in *E. prosoblepon*; Fig. 3) demonstrating an enhanced evolutionary rate for glycosyl hydrolases in Neobatrachian frogs compared to Mesobatrachian frogs. An important subset of



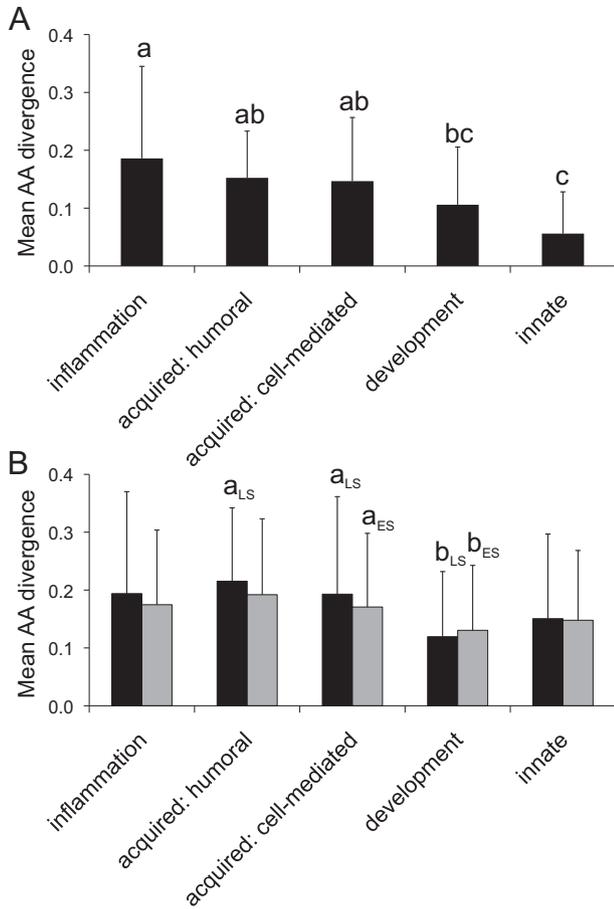
**Fig. 3.** Orthologous protein pairs showing significant deviations from expected proportions of high, intermediate and low amino acid (aa) divergence categories for all pairwise comparisons between *E. prosoblepon*, *L. yavapaiensis* (EL) and *S. tropicalis* orthologs. GO categories showing significant deviations are parsed into the following three functional categories: cellular component, biological process, and molecular function. GO terms showing significant deviations in all three species pairwise comparisons are shown in italics.

glycosyl hydrolases are lysozymes, enzymes of the innate immune system that damage bacterial and fungal cell walls (McKenzie and White, 1991) and provide a major first line of defense against pathogens. Notably, transcriptome-wide gene expression studies in *S. tropicalis* found significant up-regulation of lysozyme expression in *Bd* infected frogs compared to controls (Rosenblum et al., 2009). In the Chinese brown frog, EST sequencing of skin cDNA also showed lysozyme expression in the skin (Zhang et al., 2009).

### 3.4. Lysozyme evolution and anuran immunity

To further explore lysozyme evolution in Neobatrachian frogs, we reconstructed a Bayesian genealogy of *L. yavapaiensis* and *E. prosoblepon* lysozyme contigs and their top BLASTx matches (Fig. 5). *L. yavapaiensis* transcriptomes include at least five goose (g)-type lysozymes expressed in the skin and three in the spleen, as well as one chicken (c)-type lysozyme expressed in the skin and one in the spleen (Fig. 5A). While some

of these contigs group closely with other frog sequences (*X. laevis* and *Lithobates catesbeianus*), several others fall in distinct clades from their closest anuran ortholog. Similarly, the *E. prosoblepon* intestine transcriptome includes at least one expressed g-type lysozyme, three expressed c-type lysozymes, and two highly divergent expressed lysozymes that group most closely with an invertebrate (*Caenorhabditis elegans*) protist-type lysozyme (Fig. 5A). Two possibilities account for this pattern: *E. prosoblepon* may express a novel group of lysozymes not recovered from any other vertebrate to date, or these lysozymes represent genes being expressed by invertebrate parasites in the gut of the sampled *E. prosoblepon* individuals. Amino acid sequence identity of these contigs to the top-hit *C. elegans* sequence is only 55%, thus further characterization of lysozymes in other anurans is necessary to resolve this pattern. Regardless, both Neobatrachian species described here express a high number of phylogenetically distinct lysozymes across tissues. This pattern is concordant with other vertebrate species for which whole genomes have been sequenced and annotated: in mammals the number



**Fig. 4.** Mean protein divergence across orthologous gene pairs falling under five immune function categories. A, Protein divergence for *E. prosoblepon* and *L. yavapaiensis* (EL) across immune function groups. B, Protein divergence for *E. prosoblepon* and *S. tropicalis* (gray bars, ES), and *L. yavapaiensis* and *S. tropicalis* (black bars, LS). All values are shown as mean divergence + standard deviation. Categories with different letters represent significant differences in divergence based on Mann–Whitney U tests with Bonferroni correction.

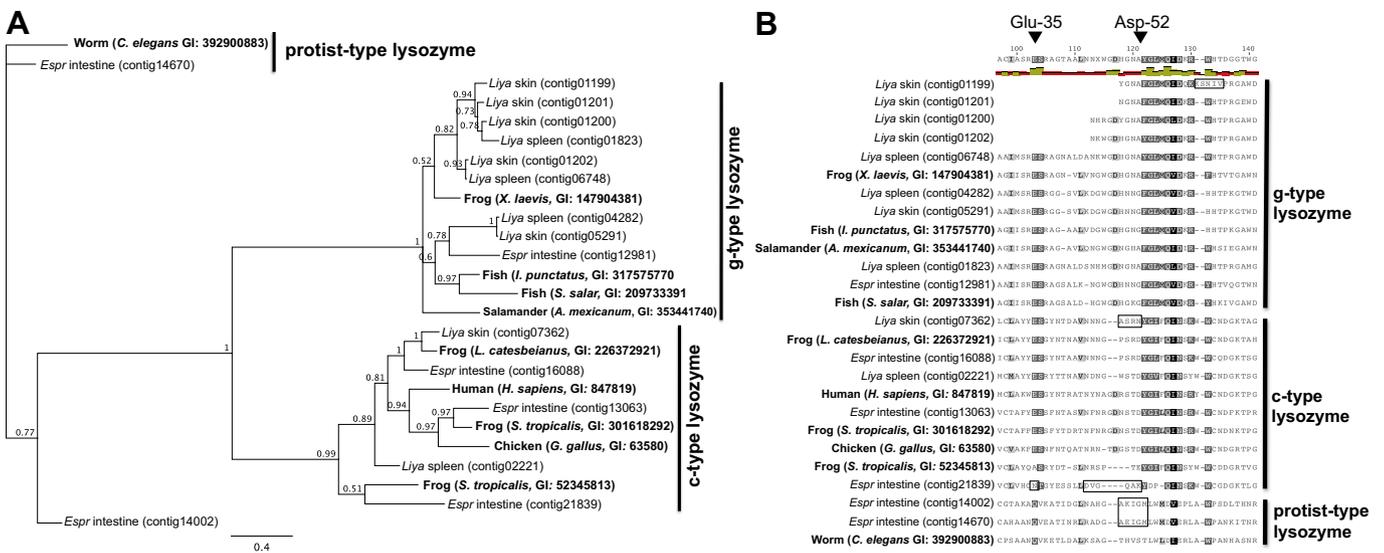
of lysozyme-like genes ranges from five in opossum to 18 in cow, the anole lizard has eight lysozyme-like genes, and *S. tropicalis* has 16 (Irwin et al., 2011).

The key active site residues in c-type lysozymes are Glu35 (glutamic acid) and Asp52 (aspartic acid; Fig. 5B), which both contribute to the catalytic mechanism of cleaving bacterial cell walls (Cheetham et al., 1992). The novel c-type sequences we recovered from *L. yavapaiensis* and *E. prosoblepon* are all conserved at residue Glu35, with the exception of one *E. prosoblepon* sequence that has an asparagine (Asn) substitution (Fig. 5B) and may not have the same catalytic activity. Similarly, our c-type sequences are all conserved at residue Asp52, except that one *L. yavapaiensis* skin has an asparagine substitution and one *E. prosoblepon* has a lysine (Lys) substitution (Fig. 5B). However, Asp52 is not essential: a mutant enzyme with an Asn substitution for Asp52 showed some residual bacteriocidal activity (Malcolm et al., 1989). Thus, the novel sequences we recovered with Asp52 substitutions likely retain at least some catalytic activity. Among g-type lysozyme sequences, the key active site is Glu73 instead of Glu35 (Weaver et al., 1995), which is present in all sequences recovered here (Fig. S4), suggesting that they all function as lysozymes. Finally, for protist-type (or invertebrate) lysozymes, Asp30 is the only key active site conserved across taxa (Goto et al., 2007), and we also recovered Asp30 in our novel sequences (Fig. S4) indicating bacteriocidal activity.

Given the large number of novel, expressed, bacteriocidal lysozyme sequences we recovered across species and tissues, lysozymes are primary candidates for highly diversified targets of positive selection in response to pathogen pressure that contribute to anuran host defenses against chytridiomycosis and other emerging diseases. We recommend that future immunological studies in amphibians focus on lysozyme activity across orthologs and tissues as important mechanisms of pathogen defense.

3.5. Evolutionary rates of immune system genes

Across five functional immune categories, mean amino acid divergence between *L. yavapaiensis* and *E. prosoblepon* orthologous contigs was significantly different across several groups (Fig. 4A). Inflammatory response genes and both categories of acquired immunity genes were significantly more diverged compared to innate immune genes. However,



**Fig. 5.** Lysozyme contigs recovered in our study (regular text) and their top matches from BLASTx searches in GenBank (bold text). A, Unrooted Bayesian phylogeny of the three groups of lysozyme genes recovered from *E. prosoblepon* and *L. yavapaiensis* transcriptomes. Posterior probabilities are shown above each node, and top BLASTx hits are shown in bold. B, Amino acid alignment of the region spanning key active site lysozyme residues showing variation across novel expressed sequences recovered from *E. prosoblepon* and *L. yavapaiensis*. Regions of unique amino acid substitutions compared to previously characterized lysozymes are highlighted in boxes.

only inflammatory response genes were significantly more diverged compared to immune system development genes. Comparison between *S. tropicalis* and *E. prosoblepon* orthologous contigs indicates that only cell-mediated acquired immunity genes were significantly more diverged compared to immune system development genes (Fig. 4B). In contrast, divergence between *S. tropicalis* and *L. yavapaiensis* orthologous contigs was significantly higher for both cell-mediated and humoral acquired immunity genes compared to immune system development genes. No other categories were significantly different in species pairwise comparisons.

At the level of individual genes, the top tier of diverged immune system orthologs reflects the broader pattern of inflammation and acquired immunity genes showing the highest divergence across species. Among the eleven ortholog pairs with more than 40% aa divergence between *L. yavapaiensis* and *E. prosoblepon*, seven are involved in regulation of immune responses, one is an MHC gene, and two code for pro-inflammatory molecules (Table S7). Similarly, of the six ortholog pairs with more than 50% aa divergence between *S. tropicalis* and *E. prosoblepon*, three are pro-inflammatory cytokines, and of the fourteen ortholog pairs with more than 50% aa divergence between *S. tropicalis* and *L. yavapaiensis*, five are cytokines, three are MHC genes (two class II genes and beta-2-microglobulin), and two are T-cell regulation genes (CD86 and CD97; Table S7).

A total of 6011 and 12,665 SNPs were identified in *E. prosoblepon* and *L. yavapaiensis* annotated contigs, respectively. Higher SNP frequency in *L. yavapaiensis* may reflect the higher number of individuals used to construct transcriptomes ( $N = 4$ ) compared to *E. prosoblepon* ( $N = 2$ ), especially given the greater number of reads and contigs for *E. prosoblepon*. No significant excess/deficit of synonymous or non-synonymous SNPs was found in any GO-slim category for either species. The mean number of SNPs per kb, synonymous SNPs per kb, or non-synonymous SNPs per kb was not significantly different for contigs with immune function compared to all other contigs, or for immune-related contigs in *L. yavapaiensis* compared to *E. prosoblepon*.

We anticipated detecting an elevated evolutionary rate in frog immune genes based on the numerous studies in other taxonomic groups showing more adaptive evolution in immune compared to non-immune proteins (e.g. Downing et al., 2009; Fumagalli et al., 2011; Nielsen et al., 2005). However, not all studies detect this pattern; one analysis comparing *Drosophila* species found no evidence for selective maintenance of protein diversity in immune versus non-immune genes, but immune genes showed a significant pattern of directional selection (Schlenke and Begun, 2003). Immune genes with significantly elevated evolutionary rates tend to code for proteins with pathogen recognition function (Sackton et al., 2007), consistent with our finding of an elevated level of protein divergence in glycosyl proteases (Fig. 3), which recognize and destroy bacterial and fungal cell walls. Finally, in addition to protein function, rates of protein evolution are affected by a number of other factors, including gene essentiality (Wall et al., 2005), expression level or tissue specificity (Subramanian and Kumar, 2004), and gene compactness (Liao et al., 2006). Thus, a variety of gene and protein properties that have not yet been explored in anurans may explain the lower evolutionary rates found in this study compared to other investigations of immune genes in animal genomes.

#### 4. Conclusions

Few transcriptome studies of amphibian immune function are available (Zhang et al., 2009). Microarray expression data, designed from either EST or whole transcriptome sequencing, have shown the lack of a robust immune response in *S. tropicalis* and the *Lithobates* (*Rana*) *muscosa*/*Lithobates sierrae* species complex in liver, spleen and skin tissues when individuals were exposed to *Bd* (Rosenblum et al., 2009, 2012). In *S. tropicalis*, only *Bd* infection at non-optimal pathogen temperatures activates a suite of innate immune genes (Ribas et al., 2009). Similarly, in a study of salamander immune responses to pathogens, *A. mexicanum* individuals exposed to ranavirus upregulated a suite of

innate immune system genes in the spleen, including interferons, chemokines and complement proteins (Cotter et al., 2008). Our results are consistent with a large role for innate immunity in amphibian pathogen defense; *L. yavapaiensis* individuals, three of which harbored *Bd* infections, were enriched for inflammatory responses in skin when compared to *E. prosoblepon* intestine. Further, the large number of innate and acquired immune processes represented in spleen transcripts may suggest a role for acquired immunity in pathogen defense. However, the known immune function of spleen may explain this pattern rather than *Bd* infection status, and future studies should examine responses to *Bd* for the same set of immune tissues across a range of host species and *Bd* susceptibilities. More broadly, both species showed a significantly elevated divergence rate compared to *Silurana* for glycosyl proteases, which function in innate immunity, but not for any acquired immunity genes. In contrast, among immune categories, *L. yavapaiensis* and *E. prosoblepon* were significantly more diverged for acquired immunity processes compared to other immune categories. Taken together, these findings suggest a central role for innate immune function in pathogen defense over the evolutionary history of anurans, but a more recent role for selection shaping acquired immunity genes in Neobatrachian frogs.

In summary, we sequenced transcriptomes of three immune-related tissues in two non-model frog species, providing new genomic resources for amphibians and demonstrating innate and acquired immune function in skin, spleen and intestine. These data enhance our current knowledge of tissue-specific immune gene expression in frog species threatened by chytridiomycosis, providing an important genomic framework for future studies of disease resistance as well as conservation and recovery efforts.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.gene.2014.03.051>.

#### Conflict of interests

The authors declare that they have no conflict of interests.

#### Authors' contributions

AS, KKT, and KZ were responsible for the design and coordination of this study. AS and AE carried out most of the bioinformatics analyses; KKT and RF also contributed to data analysis. AS, KKT and AE wrote the manuscript. All the authors read and approved the final manuscript.

#### Acknowledgments

We thank Karen Lips for assistance with specimen collection, Steve Bogdanowicz for library preparation, Jacob Crawford, Francisco Pina-Martins and Dan Ilut for assistance with programming, and Autoridad Nacional del Ambiente, Panama and Arizona Game and Fish Department for scientific collecting permits. This study was supported by the National Science Foundation (DEB-0815315 and DEB-1120249), a grant from the Cornell Center for Vertebrate Genomics, a Cornell Center for Comparative and Population Genomics Fellowship, and a Smithsonian Institution Molecular Evolution Postdoctoral Fellowship.

#### References

- Abascal, F., Zardoya, R., Posada, D., 2005. ProfTest: selection of best-fit models of protein evolution. *Bioinformatics* 21, 2104–2105.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57, 289–300.
- Berger, L., Speare, R., Daszak, P., Green, D.E., Cunningham, A.A., Goggin, C.L., Slocombe, R., Ragan, M.A., Hyatt, A.D., McDonald, K.R., Hines, H.B., Lips, K.R., Marantelli, G., Parkes, H., 1998. Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *Proceedings of the National Academy of Sciences of the United States of America* 95, 9031–9036.
- Blüthgen, N., Brand, K., Cajavec, B., Swat, M., Herzel, H., Beule, D., 2005. Biological profiling of gene groups utilizing Gene Ontology. *Genome Informatics* 16, 106–115.

- Boyle, D.G., Boyle, D.B., Olsen, V., Morgan, J.A.T., Hyatt, A.D., 2004. Rapid quantitative detection of chytridiomycosis (*Batrachochytrium dendrobatidis*) in amphibian samples using real-time Taqman PCR assay. *Diseases of Aquatic Organisms* 60, 141–148.
- Calboli, F., Fisher, M.C., Garner, T.W.J., Jehle, R., 2011. The need for jumpstarting amphibian genome projects. *Trends in Ecology & Evolution* 26, 378–379.
- Cannatella, D.C., Ford, D., Bockstanz, L., 2008. Neobatrachia. "Higher" Frogs. Version 20. The Tree of Life Web Project.
- Cheatham, J.C., Artymiuk, P.J., Phillips, D.C., 1992. Refinement of an enzyme complex with inhibitor bound at partial occupancy. Hen egg-white lysozyme and tri-N-acetylchitotriose at 1.75 Å resolution. *Journal of Molecular Biology* 224, 613–628.
- Chinchar, V.G., 2002. Ranaviruses (family Iridoviridae): emerging cold-blooded killers. *Archives of Virology* 147, 447–470.
- Cotter, J.D., Storfer, A., Page, R.B., Beachy, C.K., Voss, S.R., 2008. Transcriptional response of Mexican axolotls to *Ambystoma tigrinum* virus (ATV) infection. *BMC Genomics* 9, 493.
- Crawford, J.E., Guelbeogo, W.M., Sanou, A., Traoré, A., Vernick, K.D., Sagnon, N., Lazzaro, B. P., 2010. De novo transcriptome sequencing in *Anopheles funestus* using illumina RNA-seq technology. *PLoS ONE* 5, e14202.
- Downing, T., Cormican, P., O'Farrelly, C., Bradley, D.G., Lloyd, A.T., 2009. Evidence of the adaptive evolution of immune genes in chicken. *BMC Research Notes* 2, 254.
- Du Pasquier, L., Schwager, J., Flajnik, M.F., 1989. The immune system of *Xenopus*. *Annual Review of Immunology* 7, 251–275.
- Du Pasquier, L., Robert, J., Courtet, M., Musmann, R., 2000. B-cell development in the amphibian *Xenopus*. *Immunological Reviews* 175, 201–213.
- Dubois, B., Goubier, A., Joubert, G., Kaiserlian, D., 2005. Oral tolerance and regulation of mucosal immunity. *Cellular and Molecular Life Sciences* 62, 1322–1332.
- Duda Jr., T.F., Vanhoye, D., Nicolas, P., 2002. Roles of diversifying selection and coordinated evolution in the evolution of amphibian antimicrobial peptides. *Molecular Biology and Evolution* 19, 858–864.
- Fumagalli, M., Sironi, M., Pozzoli, U., Ferrer-Admetlla, A., Pattini, L., Nielsen, R., 2011. Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genetics* 7, e1002355.
- Goto, T., Abe, Y., Kakuta, Y., Takeshita, K., Imoto, T., Ueda, T., 2007. Crystal structure of *Tapes japonica* Lysozyme with substrate analogue: structural basis of the catalytic mechanism and manifestation of its chitinase activity accompanied by quaternary structural change. *Journal of Biological Chemistry* 282, 27459–27467.
- Gregory, T.R., 2005. Genome size evolution in animals. In: Gregory, T.R. (Ed.), *The Evolution of the Genome*. Elsevier, San Diego, pp. 3–87.
- Hauswaldt, J., Stuckas, H., Pfautsch, S., Tiedemann, R., 2007. Molecular characterization of MHC class II in a nonmodel anuran species, the fire-bellied toad *Bombina bombina*. *Immunogenetics* 59, 479–491.
- Hellsten, U., Harland, R.M., Gilchrist, M.J., Hendrix, D., Jurka, J., Kapitonov, V., Ovcharenko, I., Putnam, N.H., Shu, S., Taher, L., Blitz, I.L., Blumberg, B., Dichmann, D.S., Dubchak, I., Amaya, E., Dettler, J.C., Fletcher, R., Gerhardt, D.S., Goodstein, D., Graves, T., Grigoriev, I. V., Grimwood, J., Kawashima, T., Lindquist, E., Lucas, S.M., Mead, P.E., Mitros, T., Ogino, H., Ohta, Y., Poliakov, A.V., et al., 2010. The genome of the Western clawed frog *Xenopus tropicalis*. *Science* 328, 633–636.
- Huelsenbeck, J., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755.
- Hyatt, A.D., Boyle, D.G., Olsen, V., Boyle, D.B., Berger, L., Obendorf, D., Dalton, A., Kriger, K., Heros, M., Hines, H., Phillott, R., Campbell, R., Marantelli, G., Gleason, F., Coiling, A., 2007. Diagnostic assays and sampling protocols for the detection of *Batrachochytrium dendrobatidis*. *Diseases of Aquatic Organisms* 73, 175–192.
- Igawa, T., Kurabayashi, A., Usuki, C., Fujii, T., Sumida, M., 2008. Complete mitochondrial genomes of three Neobatrachian anurans: a case study of divergence time estimation using different data and calibration settings. *Gene* 407, 116–129.
- Irwin, D.M., Biegel, J.M., Stewart, C.-B., 2011. Evolution of the mammalian lysozyme gene family. *BMC Evolutionary Biology* 11, 166.
- Kiennec-Tyburczy, K.M., Richmond, J.Q., Savage, A.E., Zamudio, K.R., 2010. Selection, trans-species polymorphism and locus identification of major histocompatibility complex class IIβ alleles of New World ranid frogs. *Immunogenetics* 62, 741–752.
- Kiennec-Tyburczy, K.M., Richmond, J.Q., Savage, A.E., Lips, K.R., Zamudio, K.R., 2012. Genetic diversity of MHC class I loci in six non-model frogs is shaped by positive selection and gene duplication. *Heredity* 109, 146–155.
- Kilpatrick, A.M., Briggs, C.J., Daszak, P., 2010. The ecology and impact of chytridiomycosis: an emerging disease of amphibians. *Trends in Ecology & Evolution* 25, 109–118.
- Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659.
- Liao, B.-Y., Scott, N.M., Zhang, J., 2006. Impacts of gene essentiality, expression pattern, and gene compactness on the evolutionary rate of mammalian proteins. *Molecular Biology and Evolution* 23, 2072–2080.
- Liao, X., Cheng, L., Xu, P., Lu, G., Wachholtz, M., Sun, X., Chen, S., 2013. Transcriptome analysis of crucian carp (*Carassius auratus*), an important aquaculture and hypoxia-tolerant species. *PLoS ONE* 8, e62308.
- Looso, M., Preussner, J., Sousounis, K., Bruckschotten, M., Michel, C.S., Lignelli, E., Reinhardt, R., Hoeffner, S., Krueger, M., Tsonis, P.A., Borchardt, T., Braun, T., 2013. A de novo assembly of the new transcriptome combined with proteomic validation identifies new protein families expressed during tissue regeneration. *Genome Biology* 14, R16.
- Mable, B.K., Alexandrou, M.A., Taylor, M.I., 2011. Genome duplication in amphibians and fish: an extended synthesis. *Journal of Zoology* 284, 151–182.
- Malcolm, B.A., Rosenberg, S., Corey, M.J., Allen, J.S., de Baetselier, A., Kirsch, J.F., 1989. Site-directed mutagenesis of the catalytic residues Asp-52 and Glu-35 of chicken egg white lysozyme. *Proceedings of the National Academy of Sciences of the United States of America* 86, 133–137.
- McKenzie, H.A., White Jr., F.H., 1991. Lysozyme and alpha-lactalbumin: structure, function, and interrelationships. *Advances in Protein Chemistry* 41, 173–315.
- McTaggart, S.J., Obbard, D.J., Conlon, C., Little, T.J., 2012. Immune genes undergo more adaptive evolution than non-immune system genes in *Daphnia pulex*. *BMC Evolutionary Biology* 12, 63.
- Morin, R.D., Chang, E., Petrescu, A., Liao, N., Griffith, M., Kirkpatrick, R., Butterfield, Y.S., Young, A.C., Stott, J., Barber, S., Babakaiff, R., Dickson, M.C., Matsuo, C., Wong, D., Yang, G.S., Smailus, D.E., Wetherby, K.D., Kwong, P.N., Grimwood, J., Brinkley, C.P., Brown-John, M., Reddix-Dugue, N.D., Mayo, M., Schmutz, J., Beland, J., Park, M., Gibson, S., Olson, T., Bouffard, G.G., Tsai, M., et al., 2006. Sequencing and analysis of 10,967 full-length cDNA clones from *Xenopus laevis* and *Xenopus tropicalis* reveals post-tetraploidization transcriptome remodeling. *Genome Research* 16, 796–803.
- Moss, A.S., Carty, N., San Francisco, M.J., 2010. Identification and partial characterization of an elastolytic protease in the amphibian pathogen *Batrachochytrium dendrobatidis*. *Diseases of Aquatic Organisms* 92, 149–158.
- Myhre, S., Tveit, H., Mollestad, T., Lægred, A., 2006. Additional Gene Ontology structure for improved biological reasoning. *Bioinformatics* 22, 2020–2027.
- Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civallo, D., White, T.J., Sninsky, J., Adams, M.D., Cargill, M., 2005. A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biology* 3, e170.
- Ohta, Y., Goetz, W., Hossain, M.Z., Nonaka, M., Flajnik, M.F., 2006. Ancestral organization of the MHC revealed in the amphibian *Xenopus*. *Journal of Immunology* 176, 3674–3685.
- Pyron, R.A., Wiens, J.J., 2011. A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Molecular Phylogenetics and Evolution* 61, 543–583.
- Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., Lopez, R., 2005. InterProScan: protein domains identifier. *Nucleic Acids Research* 33 (Web Server issue), W116–120.
- Ramsköld, D., Wang, E.T., Burge, C.B., Sandberg, R., 2009. An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. *PLoS Computational Biology* 5, e1000598.
- Ribas, L., Li, M.-S., Doddington, B.J., Robert, J., Seidel, J.A., Kroll, J.S., Zimmerman, L.B., Grassly, N.C., Garner, T.W.J., Fisher, M.C., 2009. Expression profiling the temperature-dependent amphibian response to infection by *Batrachochytrium dendrobatidis*. *PLoS ONE* 4, e8408.
- Richmond, J.Q., Savage, A.E., Zamudio, K.R., Rosenblum, E.B., 2009. Toward immunogenetic studies of amphibian chytridiomycosis: linking innate and acquired immunity. *BioScience* 59, 311–320.
- Robert, J., Ohta, Y., 2009. Comparative and developmental study of the immune system in *Xenopus*. *Developmental Dynamics* 238, 1249–1270.
- Roelants, K., Gower, D.J., Wilkinson, M., Loader, S.P., Biju, S.D., Guillaume, K., Moriau, L., Bossuyt, F., 2007. Global patterns of diversification in the history of modern amphibians. *Proceedings of the National Academy of Sciences of the United States of America* 104, 887–892.
- Rollins-Smith, L.A., Ramsey, J.P., Reinert, L.K., Woodhams, D.C., Livo, L.J., Carey, C., 2009. Immune defenses of *Xenopus laevis* against *Batrachochytrium dendrobatidis*. *Frontiers in Bioscience* 1, 68–91.
- Rollins-Smith, L.A., Ramsey, J.P., Pask, J.D., Reinert, L.K., Woodhams, D.C., 2011. Amphibian immune defenses against chytridiomycosis: impacts of changing environments. *Integrative and Comparative Biology* 51, 552–562.
- Rosenblum, E.B., Poorten, T.J., Settles, M., Murdoch, G.K., Robert, J., Maddox, N., Eisen, M.B., 2009. Genome-wide transcriptional response of *Silurana (Xenopus) tropicalis* to infection with the deadly chytrid fungus. *PLoS ONE* 4, e6494.
- Rosenblum, E.B., Poorten, T.J., Settles, M., Murdoch, G.K., 2012. Only skin deep: shared genetic response to the deadly chytrid fungus in susceptible frog species. *Molecular Ecology* 21, 3110–3120.
- Sackton, T.B., Lazzaro, B.P., Schlenke, T.A., Evans, J.D., Hultmark, D., Clark, A.G., 2007. Dynamic evolution of the innate immune system in *Drosophila*. *Nature Genetics* 39, 1461–1468.
- Savage, A.E., Zamudio, K.R., 2011. MHC genotypes associate with resistance to a frog-killing fungus. *Proceedings of the National Academy of Sciences* 108, 16705–16710.
- Schlenke, T.A., Begun, D.J., 2003. Natural selection drives *Drosophila* immune system evolution. *Genetics* 164, 1471–1480.
- Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27, 863–864.
- Sémon, M., Mouchiroud, D., Duret, L., 2005. Relationship between gene expression and GC-content in mammals: statistical significance and biological relevance. *Human Molecular Genetics* 14, 421–427.
- Shaffer, J.P., 1995. Multiple hypothesis testing. *Annual Review of Psychology* 46, 563–584.
- Subramanian, S., Kumar, S., 2004. Gene expression intensity shapes evolutionary rates of the proteins encoded by the vertebrate genome. *Genetics* 168, 373–381.
- Teacher, A.G.F., Garner, T.W.J., Nichols, R.A., 2009. Evidence for directional selection at a novel major histocompatibility class I marker in wild common frogs (*Rana temporaria*) exposed to a viral pathogen (ranavirus). *PLoS One* 4, e4616.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22, 4673–4680.
- Voss, S.R., Kump, D.K., Putta, S., Pauly, N., Reynolds, A., Henry, R.J., Basa, S., Walker, J.A., Smith, J.J., 2011. Origin of amphibian and avian chromosomes by fission, fusion, and retention of ancestral chromosomes. *Genome Research* 21, 1306–1312.
- Wall, D.P., Hirsh, A.E., Fraser, H.B., Kumm, J., Giaever, G., Eisen, M.B., Feldman, M.W., 2005. Functional genomic analysis of the rates of protein evolution. *PNAS* 102, 5483–5488.
- Wang, S., Hou, R., Bao, Z., Du, H., He, Y., Su, H., Zhang, Y., Fu, X., Jiao, W., Li, Y., Zhang, L., Wang, S., Hu, X., 2013. Transcriptome sequencing of zhikong scallop (*Chlamys farreri*)

- and comparative transcriptomic analysis with yesso scallop (*Patinopecten yessoensis*). PLoS ONE 8, e63927.
- Weaver, L.H., Grütter, M.G., Matthews, B.W., 1995. The refined structures of goose lysozyme and its complex with a bound trisaccharide show that the "goose-type" lysozymes lack a catalytic aspartate residue. *Journal of Molecular Biology* 245, 54–68.
- Weldon, C., du Preez, L.H., Hyatt, A.D., Müller, R., Speare, R., 2004. Origin of the amphibian chytrid fungus. *Emerging Infectious Diseases* 10, 2100–2105.
- Yang, W., Qi, Y., Bi, K., Fu, J., 2012. Toward understanding the genetic basis of adaptation to high-elevation life in poikilothermic species: a comparative transcriptomic analysis of two ranid frogs, *Rana chensinensis* and *R. kukunoris*. *BMC Genomics* 13, 588.
- Zhang, Z., Zhang, B., Nie, X., Liu, Q., Xie, F., Shang, D., 2009. Transcriptome analysis and identification of genes related to immune function in skin of the Chinese Brown Frog. *Zoological Science* 26, 80–86.