Diversity of genes conferring resistance to toxins inferred from transcriptomic data in frogs and toads.

María del Pilar Rodríguez-Ordóñez[a]*, Peter Andolfatto[b]and Andrew J. Crawford[acd].

[a]Departamento de CienciasBiológicas, Universidad de los Andes, A.A. 4976 Bogotá, Colombia

[b] Department of Ecology and Evolutionary Biology& The Lewis-Sigler Institute or Integrative Genomics , Princeton University, New Jersey, USA.

[c]Smithsonian Tropical Research Institute, Apartado 0843-03092, Panamá, Republic of Panama

[d] Círculo Herpetológico de Panamá, Apartado 0824-00122, Panamá, Republic of Panama

*Corresponding author:

Fax: +57-1-332-4069

Tel.: +57-1-339-4949 ext. 2754

Instituto de Genética, Edif. M1-311

Departamento de Ciencias Biológicas

Universidad de los Andes

Carrera 1 No. 18A – 10

A.A. 4976, Bogotá

Colombia, South America

E-mail address: md.rodriguez39uniandes.edu.co

Abstract

Specific amino acidic substitutions in the Na^+-K^+ ATPase α1 isozyme confer resistance to ouabain in species that confront cardiotonic steroids. As occurs in insects that feed from plants that produce cardenolides, in species such as *Leptodactylus ocellatus*, a toad-eating frog, gene duplication has been suggested as a mechanism of evolution that allows, through the acquisition of a toxin-resistant Na^+-K^+ ATPase α1 copy, adaptation to endogenous steroid resistance. In this study, we investigated, through transcriptomic analyses, whether toxic bufonid species (*Atelopus zeteki*, *A.*sp, *Rhinella marina* and *R. granulosa*) and *Leptodactylus ocellatus* have achieved resistance through amino acid point substitutions of an ancestral sensitive form or through gene duplication and divergence. In *L. ocellatus*, we found two paralogs copies of the Na^+-K^+ ATPase α1 isozyme differing in their toxin sensitivity and differently expressed in stomach and brain. In bufonid toads, in contrast, we found only a single resistant copy of the Na^+-K^+ ATPase α1expressed in all tissues, suggesting that a paralogous copy of the Na^+-K^+ ATPase α1gene exhibiting toxin sensitivity would presumably not be beneficial in such species producing high levels of endogenous toxins in nearly all the tissues. We conclude that in *L. ocellatus* that confront endogenous toxins in their diet, gene duplication has facilitated the neo-functionalization of paralogs copies of the Na^+-K^+ ATPase α1gene differing in their ouabain-sensitivity to copy with toxin resistance. Thus, anurans have achieved resistance through both pathways, adaptive in species with endogenous toxins, versus gene duplication and divergence.

Keywords: Cardiotonic steroids, Na^+-K^+ ATPase α isozymes, gene duplication, toads, *Leptodactylus ocellatus*, toxin resistance.

Introduction

Mechanisms of adaptive molecular evolution include substitutions at the nucleotide and amino acid levels, as well as whole gene duplications followed by adaptive divergence (Zhang 2003; Mable *et al.* 2011). Adaptive substitutions may occur at regulatory sites as well as within coding regions (Pál *et al.* 2006; Odom *et al.* 2007). Adaptive substitutions in genes with negative pleiotropic effects, however, may be limited by constraints imposed by multiple functions of gene products. One possible mechanism to escape such constraints is by gene duplication that facilitates the evolution of new functional adaptations in the new gene copy while maintaining vital functions in the original gene (Prince & Pickett 2002; Hoekstra & Coyne 2007). Studies in species with genome duplication has suggested that indeed the persistence of individual gene duplications promotes adaptation (Chain *et al.* 2008). Therefore, gene duplication could contribute to phenotypic diversification via functional redundancy and increased rates of adaptation (Kondrashov *et al.* 2002; Crow & Wagner 2006; Hittinger & Carroll 2007).

Both convergence and parallelism are as well mechanism of molecular evolution related with adaptation. In convergent evolution, independent lineages acquire the same amino acid changes from different ancestral amino acids whereas in parallel evolution, lineages acquire the same amino acid changes independently from the same ancestral amino acids (Zhang & Kumar 1997). Frequent parallelism and convergence at the molecular level strongly suggest that under similar selective pressures, there are a limited number of pathways to evolve new functions (Christin *et al.* 2010; Conte *et al.* 2012; Martin & Orgogozo 2013).

50 The resistance of animals to their own toxins and those of their preys is an
51 adaptive trait that may involve mechanisms of gene duplication and convergent
52 evolution. Snakes from different lineages have evolved resistance to tetrodotoxin
53 showing convergence in the resistance-conferring amino acidic substitutions in their
54 sodium channels (Feldman *et al.* 2012). Species of pufferfishes have evolved
55 independently resistance to tetrodotoxin and saxitoxin by parallel amino acid
56 replacements across all eight sodium channels present in teleost fish genomes (Jost *et*
57 *al.* 2008). Again, similar selective pressures promote similar patterns of protein
58 evolution.
59 Cardiotonic steroids are widespread in nature, providing an opportunity to
60 study the evolution of toxin resistance in diverse taxa. Cardiotonic steroids include
61 bufodienolides that are synthesized directly by animals, and cardenolides that are
62 derived from plants (Barrueto *et al.* 2006; Bagrov *et al.* 2009; Agrawal *et al.*
63 2012).The main toxic effect of these compounds, of which ouabain is the most
64 common, is to inhibit the Na^+-K^+ ATPase activity, which impedes cation transport
65 and nerve impulse transmission (Yotsu-Yamashita *et al.* 2004; Garraffo *et al.* 2012;
66 Grant *et al.* 2012).
67 Na^+-K^+ ATPase is a transmembrane protein responsible for neural function
68 and muscle contraction through ion transport, specially sodium and potassium, into
69 and out of the cells of almost all organisms. The Na^+-K^+ ATPase protein has three
70 polypeptide subunits codified by different genes (Blanco & Mercer 1998; Pierre &
71 Xie 2006). Specifically the α subunit, composed of ten transmembrane segments that
72 form extracellular and intracellular loops, has the binding site for ATP, K^+ as well as
73 for ouabain, and therefore plays a catalytic function (Blanco & Mercer 1998). In most
74 vertebrates, the α subunit has four isozymes each one coded by distinct paralogs

whose common ancestor presumably dates back to the origin of vertebrates (Takeyasu *et al.* 1990).

The Na^+-K^+ ATPase α1 isozyme is expressed in all tissues (Blanco & Mercer 1998), and in some species, the ouabain-binding site of this isozyme has accumulated mutations that confer resistance to the effects of ouabain (Jaisser *et al.* 1992; Croyle *et al.* 1997; Moore *et al.* 2009). For this reason, any change in its amino acid composition is pivotal to understanding animal resistance to endogenous and exogenous toxins. On the other hand, the ouabain-binding site of the Na^+-K^+ ATPase α1 is highly conserved in *Drosophila,* sheep (*Ovis orientalis aries),* rodents, insects, and humans, which demonstrate its functional importance (Emery *et al.* 1995; Moore *et al.* 2009; Lingrel 2010). The binding of ouabain regulates Na^+-K^+ ATPase activity through the increase of intracellular content of calcium that promotes muscle contraction and activates signaling pathways that facilitate cell proliferation (Schoner & Scheiner-Bobis 2007; Nguyen *et al.* 2011). Therefore, the affinity or insensitivity to cardionotinic steroids has profound impacts in the Na^+-K^+ ATPase activity regulation.

Recent studies (Aardema *et al.* 2012; Dobler *et al.* 2012; Zhen *et al.* 2012) focused on the evolution of the resistance to cardenolides in insects that feed of plants of the family Apocynaceae, have demonstrated through mutagenesis and molecular docking, the parallelism in adaptive substitutions at the ouabain-binding site of the Na^+-K^+ ATPase α1 among unrelated species. In addition, through the use of transcriptomic analyses, Zhen *et al* 2012, demonstrated that species that feed on Apocynaceae present more than one paralogous copy of the Na^+-K^+ ATPase α1 gene, each one differing in its resistance to ouabain. In insects such as dogbane beetles (*Chrysochus auratus*) and milkweed stem weevils (*Rhyssomatus lineaticollis*) with two copies of Na^+-K^+ ATPase α1, the ouabain-sensitive copy shows a unique

100 substitution in contrast to the duplicated ouabain-resistant copy. This finding supports
101 the prediction that in genes with restriction for multiples changes, novel functions and
102 sub-functionalization evolved through duplications of genes (Hughes 1994;
103 Kondrashov *et al.* 2002; Prince & Pickett 2002)

104 Gene duplication also facilitates the differential expression of gene copies
105 among tissues. In the dogbane beetle (*Chrysochus auratus*), for example, the stomach
106 where the toxin first enters the organism, expresses the Na^+-K^+ ATPase α1 copy that
107 is resistant to ouabain, whereas the brain, that receives protection via its cerebral
108 membranes, expresses aNa^+-K^+ ATPase α1 copy that is sensitive to the toxins.
109 Therefore, in addition to changes in *cis* and *trans* regulatory elements, gen duplication
110 arises as an alternative mechanism to mediate differentially expression patters in
111 animal proteins (Zhen *et al.* 2012).

112 Many species of toads use toxins as defense that are produced by internal
113 synthesis and are secreted by skin glands (Yotsu-Yamashita *et al.* 2004; Moore *et al.*
114 2009; Garraffo *et al.* 2012; Grant *et al.* 2012). Specifically, toxic species of the family
115 Bufonidae, which produce bufadienolides, show evidence for positive selection in
116 Na^+-K^+ ATPase α1 amino acid substitutions conferring bufanolide resistance (Moore
117 *et al.* 2009). *Leptodactylus ocellatus,* a predator of the toxic species *Rhinella marina*
118 possesses a resistant copy of the Na^+-K^+ ATPase α1 whose amino acidic substitutions
119 are different from those present in the ouabain-resistant Na^+-K^+ ATPase α1 ofbufonid
120 toads. Additionally, the sensitive copy of *L. ocellatus* gave rise to the resistant copy
121 likely by gene duplication (Moore *et al.* 2009). Therefore, in amphibians confronted
122 with cardiotonic steroids, gene duplication is suggested as the molecular evolutionary
123 processes conferring toxin resistance.

124 The present study aimed to elucidate the molecular processes that took place in

the evolution of toxin resistance in the Na^+-K^+ ATPase α1 protein of toads and frogs. Specifically, we investigated through the use of transcriptomic analysis, whether in bufonid species (*Atelopus zeteki, A.*sp.,*Rhinella marina* and *R.granulosa*) and *Leptodactylus ocellatus* (eating-toad) the duplication of the gene coding for Na^+-K^+ ATPase α1 protein has promoted the adaptation to toxin exposure.

<u>*Hypotheses and predictions*</u>

Hypothesis 1:

Species of toads (Bufonidae) and the toad-eating species *Leptodactylus ocellatus* that confront endogenous toxins, exhibit gene duplication as the mechanism of evolution to toxin resistance.

Prediction 1: Bufonids and *Leptodactylus ocellatus* converge in the mechanism to endure toxicity and therefore possess at least two paralogs genes of the Na^+-K^+ ATPase α1 differing in its ouabain-sensitivity.

Prediction 2: Engystomops pustulosus, a frog species not know for being toxic to predators and not known to eat toads, expresses one copy of the Na^+-K^+ ATPase α1 gene that codes for a protein that is ouabain-sensitive.

Hypothesis 2:

Gen duplication has facilitated tissue-specific expression of the paralogs copies of the Na^+-K^+ ATPase α1 in toads (Bufonidae) and the toad-eating species *Leptodactylus ocellatus.*

Prediction 1:

Both the toad species (Bufonidae) and the toad-eating species *Leptodactylus ocellatus,* express the ouabain-resistant copy of the Na^+-K^+ ATPase α1 in their skin and stomach whereas in their brain they express the ouabain-sensitive Na^+-K^+ ATPase α1 copy.

Prediction 2:

Bufonids, *Leptodactylus ocellatus* and *Engystomops pustulosus* express the Na^+-K^+ ATPase isozymes α 2 and α 3in their brains but not in their stomachs nor in their skin.

Methodology

In this study, we used a transcriptomic analysis to find Na^+-K^+ ATPase α genes implicated in resistance to amphibian toxins. We sampled expressed genes in three different tissues sampled from six anurans species that either i.) produce and metabolize their own toxins, i.e.,the true toads, ii.) species that are toads-eaters, i.e., *Leptodactylus ocellatus*, and control species that are not known to produce or ingest cardenolides. Using RNA-seq (see below) allowed us to characterize multiple Na^+-K^+ ATPase α isozymes.

Ethics statement

All tissues and specimens (except for *Atelopus zeteki*) were obtained in Colombia under permissions granted by the *Ministerio de ambiente y desarrollo territorial* (“*permiso de investigación científica de la diversidad biológica” No 15 del 26 de Julio de 2010* and *“permiso de acceso a recursos genéticos”, Resolución 2408 del 1ro de Diciembre de 2010*). The extracted RNA was exported for sequencing at

The Lewis-Sigler Institute for Integrative Genomics, New Jersey, USA, under export permit ("*exportación de productos de la diversidadbiológica no contemplados en los apéndices cites"*) granted by *Autoridad Nacional de Licencias Ambientales, ANLA,* and the *secretaría distrital de ambiente.* Anna Savage and Brian Gratwicke from the Smithsonian National Zoological provided the tissues of *Atelopus zeteki.*

Sample collection, cDNA library construction and Illumina sequencing

Gene expression data was recorded from 10 tissues of six amphibian species through RNA isolation and next-generation sequencing (Table 1). RNA was preserved in the field using RNAstable (Biomatrica, Inc. San Diego, CA, USA). RNA was extracted with either the RNAeasyMiniKit (Qiagen, Hilden. Germany) or standard TRI Reagent Solution (Ambion Inc., Austin, Texas, USA). Quality and integrity of the isolated RNA were assessed before library preparation using an Agilent 2100 Bioanalyzer and RNA 6000 Nanokit. Complementary DNA (cDNA) libraries were prepared with the IlluminaTrueSeq version 2 kit, using half reactions. The ten libraries were barcoded and run together in a multiplex assay on one lane of an Illumina Hiseq 2000. Thus, each sample provided 8–11 million non-strand specific, single-end reads of 140 base pairs (bp) each, with a mean of 2.29 Gb ± 0.24 SD.

Bioinformatics analysis

Sequences trimming and de novo assembly

Read size and quality was assessed with the software Prinseq online version (Schmieder & Edwards 2011) and FastQC v.0.10.1 (Simon 2011) which confirmed

199 that all sequences had a length of 140 bp and the majority had a Phred quality score
200 greater than 20 (a sequencing error rate of <1 in every 100 base calls, or 99% base call
201 accuracy) (FigureS1). We cleaned the raw reads by removing any base with a Phred
202 quality score less than 20 and proceeding with only those sequences with ≥ 25
203 contiguous bp. This filter was made with a Python script available at
204 (http://code.google.com/p/ngopt/source/browse/trunk/SSPACE/tools/TQSfastq.py),
205 and removed a mean of 8.35% ± 6.68% (SD) of the reads of each sample (Table 2),
206 typical for such sequencing platforms (Singhal 2013).
207 As all the species analyzed lack a reference genome to which we could map
208 the reads, we used the software Trinity v. r2013-02-25 (Grabherr *et al.* 2011) to
209 perform *de novo* assembly of the reads and generate contigs of each transcript. This
210 software is widely recognized for its accuracy in predicting transcript sequences
211 (Duan *et al.* 2012; Singhal 2013; Vijay *et al.* 2013). Briefly, Trinity re-arranges the
212 reads in sequences of overlapping k-mers to reconstruct the most plausible contig
213 sequence of a transcribed gene based on *De Bruijin* graphs. K-mer length for
214 assembly was set to 25 bp. Trinity can generate contigs from different genes or
215 transcripts originating from distinct isoforms produced by alternative splicing. Trinity
216 was run with the default parameters except the "min_contig _length" was changed in
217 two test assemblies of 100 and 200 (default) minimum contig lengths. The transcripts
218 assembled by Trinity with the paramater "min_contig_length" of 200 were used in
219 downstream analyses owing to the metrics obtained in the quality assembly (see
220 results). We also assembled reads using Velvet/Oases (Schulz *et al.* 2012), an
221 assembler that uses different k-mer lengths to construct transcripts, thereby
222 optimizing the assembly in the context of differential levels of transcript expression.
223 Velvet/Oases were run to k-mer length of 31.

Accessing assembly quality

To evaluate the quality of each of the assemblies we used Bowtie (Langmead & Salzberg 2012), with a script implemented in Trinity, to align the reads back to the transcripts in order to check the fraction of reads used in the assembly. We calculated the N50 (the size of the smallest contig such that 50% of the length of the total sequence data is contained in contigs of size N50 or greater) as a metric of assembly contiguity. Higher N50 score indicates the transcripts are less fragmented (Vijay *et al.* 2013). We determined the number of CEGMA genes (Core Eukaryotic Genes Mapping Approach) (Parra *et al.* 2007) recovered by our assembly as a measure of assembly quality and sequencing depth, given that a good assembly should recover a high proportion of the 248 ultra-conserved CEGMA genes, regardless of the organism. To examine the number of transcripts that appeared to be full-length, we performed an alignment of the assembled transcripts against known proteins (Uniprot-sprot database) and obtained the number of proteins that match with Trinity transcript by more than 80% of their lengths.

Transcript annotation

Transdecoder was used to identify transcripts that had a high probability of being a protein-coding gene based on predicted open reading frames (ORF). Trinity implements Trinotate to annotate transcripts by searching databases to identify homologous genes, protein domains and signal peptides. In this analysis, we used the software Hmmer v.3.0, NCBI-Blast v. 2.2.25+, SignalP v.4.0 and Tmhmm v.2.0.6, and the databases Pfam and Uniprot-swissprot. Specifically, to annotate our transcripts we performed blastx and blastp searches of the transcripts and their best

ORF, respectively, against Uniprot-swissprot database. Finally, Trinity generated a report of the annotated genes and their inferred protein sequences, the percentage identity with homologous proteins from other vertebrates and their Gene Ontology (GO), a bioinformatics tool that provides a controlled vocabulary to describe genes and their products (Ashburner *et al.* 2000).

Searching for Na^+-K^+ ATPase α genes

In addition to the above characterization and quality evaluation of the transcriptome, a manual analysis was performed in order to identify transcripts from genes of interest. From the blastx searches against Uniprot-swissprot database, we identified transcripts derived from Na^+-K^+ ATPase α genes and assessed if they corresponded to Na^+-K^+ ATPase α1, paralogs of α1, Na^+-K^+ ATPase α2 or α3 genes. Each of the assemblers (Trinity and Velvet/Oases) generated transcripts clustered in components (Trinity) and loci (Oases/Velvet). In theory, genes clustered in the same component/locus belonged to isoforms originating by alternative splicing or to paralogs loci. However, both assemblers tended to cluster transcripts generated from different Na^+-K^+ ATPase α genes as being the same locus, making it difficult to identify paralogs genes. Those transcripts presumed to be derived from different Na^+-K^+ ATPase α genes but grouped in the same component (Trinity) or locus (Oases), were aligned through blastx against RefSeq database (proteins), in order to identify the Na^+-K^+ ATPase α gene from which they originated.

To identify paralogs copies of Na^+-K^+ ATPase α1, we aligned the transcripts generated from the same component/loci and created a consensus transcript. Raw reads were then aligned back to this consensus transcript to search for possible polymorphic sites or sequencing errors. We identified the sequence motif within the

transcript that corresponded to the ouabain-binding site and verified whether the variable sites had the substitution of the ouabain-resistant or ouabain-sensitive sequences. Consensus transcripts with an excess of variables nucleotide sites were assumed to represent multiple loci from different genes rather than polymorphic sites of the same locus. The alignments described above were performed with Mira v.3.4.0 and visualized with Tablet v.1.13.08.05 (Milne *et al.* 2010).

To identify sensitive versus resistant Na^+-K^+ATPases, we performed Blast alignments of the transcripts and their predicted best ORF against Na^+-K^+ ATPase α1, α2, and α3 protein sequences found in Uniprot-swissprot database. The transcripts that obtained best hits for Na^+-K^+ ATPase α1 were aligned individually against Na^+-K^+ ATPase α1 of *R. marina* (spP30714, AT1A1*)* from Uniprot-swissprot database, for which the amino acid sequence of the ouabain-binding site as well as ouabain-resistant and ouabain-sensitive motifs have been previously identified. The transcripts with best hits for Na^+-K^+ ATPase α2 and Na^+-K^+ ATPase α3 were aligned to the same homologous sequences of *Gallus gallus* (Uniprot-swissprot: sp|P24797, AT1A3), and the nucleotide sequences of *Xenopus tropicalis* (RefSeq:NM_001102969, AT1A2 and NM_001126894, AT1A3).

As a final evaluation of potential gene copies, we used an unpublished and partially assembled genome sequence of *R. marina* (provided kindly by Grant Morahan from the University of Western Australia).We identified all scaffolds with best hits against ATPases proteins, performed a tblastx alignment against our transcripts of *R. marina* (brain and skin) and examined the scaffolds that were the best hits against our transcripts and those with high identity to Na^+-K^+ ATPase α proteins. With the scaffolds aligned to our predicted Na^+-K^+ ATPase α 2 and α3 proteins we

confirmed that those scaffolds has the same amino acid sequences of our predicted proteins (not previously reported in any database).

Phylogenetic relationships between Na^+-K^+ ATPase α genes

To establish the phylogenetic relationships between our transcripts of Na^+-K^+ ATPase α genes relative to other vertebrate sequences in GenBank, we estimated a phylogenetic tree with the software SATé-II v. 2.2.7 (Liu *et al.* 2012) that uses maximum likelihood to iteratively estimate alignments and trees from large datasets. For the Na^+-K^+ ATPase α1 gene the dataset included fishes (4), mammals (29), birds (9), reptiles (2) and amphibians (31). For the Na^+-K^+ ATPase α2 gene the dataset included sequences of mammals (11), birds (1) and amphibians (5). For the Na^+-K^+ ATPase α3 gene we obtained sequences of bird species (3), mammals (6) and amphibians (4). For the Na^+-K^+ ATPase α 4 gene we obtained only sequences from 14 species of mammals.

Results

Assessing De novo assembly quality

The assemblies obtained by using different contig length parameters yielded different percentage of aligned reads. When minimum contig length parameter was set to 200bp, the resulted assembly showed a less percentage of aligned reads than the assembly generated by contig length of 100bp (not significant differences t = 0.674, df = 17, p-value = 0.5093; Figure 1). Even though the assembly obtained by 200bp contigs (min_contig_length parameter set to 200) generated fewer transcripts, the assembly was more contiguous with higher N50 (not significant differences t = -

0.9864, df = 18, p-value = 0.337; Figure3), and recovered a large number of proteins to nearly full-length than the assembly with the min_contig_length parameter set to 100 bp (not significant differences t = -0.6275, df = 18, p-value = 0.5382)(Figure 2). Hence, the contigs of 200bp were used in protein prediction and annotation analyses (Table 3). The assemblies generated by Velvet/Oases with k-mer length of 31 had higher N50 than the assemblies performed by Trinity. Notably, the Velvet/Oases assembly recovered those transcripts, which were not detected by Trinity in the stomach tissue library of *Leptodactylus ocellatus*. The basic quality metrics of the assemblies generated by Velvet/Oases are presented in Table 4. Assemblies generated by Trinity with min_contig_length parameter of 200 bp recovered on average 171.88 ± 31.56 (SD) of the 248 ultra-conserved CEGMA genes, whereas the assemblies with min_contig_length parameter of 100 bp recovered on average172.66 ± 20.31 (SD) of the 248 ultra-conserved CEGMA genes.

Searching for Na^+-K^+ ATPase α genes

When applied to brain transcriptomes, Trinity and Oases failed to distinguish transcripts generated by different Na^+-K^+ ATPase α isozymes genes (e.g., α 2 and α 3) typically expressed in brain and neurons of vertebrates. Na^+-K^+ ATPase α isozymes genes showed high similarity in such a way that the assemblers were not able to separate them based on nucleotide differences. In skin, stomach and brain tissues of bufonids we found only one copy of the ouabain-resistant Na^+-K^+ ATPase gene (Figure 4 and Table 4). The alignment of the reads with the consensus transcript of the Na^+-K^+ ATPase α 1 gene showed that there are no polymorphic sites at the H1-H2 domain (positions 111-122 where ouabain binds to the protein). Therefore, the unique conserved sequence at this position could be responsible for resistance to ouabain

(Figure S2). The alignment also showed that the reads generated from Na^+-K^+ ATPase α 1 genes from tissues of bufonids do not have so much variable sites in comparison with the reads generated from Na^+-K^+ ATPase α 1 gene of the stomach tissue of *L. ocellatus* (Figure S3 and Figure S5). From the brain tissue of *R. marina,* we recovered Na^+-K^+ ATPase α2 and Na^+-K^+ ATPase α3 transcripts and these were homologous to *Xenopus tropicalis* Na^+-K^+ ATPase α2 and α3 genes (NP_001096439.1 and NP_001120366.1, respectively). For the predicted proteins, the amino acid sequence of the ouabain-binding site was confirmed with the scaffolds of *R. marina* (from its partially assembled genome) as these sequences have not been previously reported.

From the brain tissue of *E. pustulosus* and *L. ocellatus,* we also recovered genes of the Na^+-K^+ ATPase α isozymes (α1, α2 and α3) that are more similar in their sequences at the ouabain-binding site (positions 111–122) between each other than with those of *R. marina* (Figure 4, Figure 5 and Figure 6). From the stomach tissue of *L. ocellatus,* we recovered two paralogs copies of the Na^+-K^+ ATPase α1 gene, one ouabain-resistant and the other ouabain-sensitive as reported by Moore et al, 2009 (Figure S4). The alignment of the reads back with the consensus transcript of the Na^+-K^+ ATPase α1 gene showed that there were more polymorphic sites across the alignment. (Figure S5).

Phylogenetic relationships between Na^+-K^+ ATPase α genes

According to the phylogenetic inference made with SatéII (Figure 7), the Na^+-K^+ ATPase α genes were clustered by the class of isozyme they coded for. The inference shows that the ouabain-resistant copy of the Na^+-K^+ ATPase α1 expressed in the stomach of *L. ocellatus* is more recent than the ouabain-sensitive one (Figure

8). This inference agrees with the idea that these two copies originated by gene duplication and diverged by the adoption of different functions.

374 **Discussion**

375 The fact that all bufonids studied here express only an ouabain-resistant copy
376 of the Na^+-K^+ ATPase α1 gene without expressing an ouabain-susceptible copy,
377 suggests that gene duplication and divergence is not the exclusive mechanism by
378 which the Na^+-K^+ ATPase α1 protein has evolved toxin resistance from a susceptible
379 ancestor. In bufonids, the production of potent endogenous toxins as a mechanism of
380 defense against predators is an adaptation that probably evolved in parallel with the
381 ability to tolerate these toxins. Since very low levels of ouabain regulate the ouabain-
382 sensitive Na+-K+ ATPase α1 protein of most animals (Dostanic-Larson *et al.* 2006;
383 Lingrel 2010), in toads that produce five dominant types of potent bufadienolides and
384 at least 90 minor bufadienolides (Hayes *et al.* 2009), the existence of an ouabain-
385 sensitive Na^+-K^+ ATPase α1 gene would not be adaptive. Therefore, the point
386 mutations in the ancestral ouabain-sensitive Na+-K+ ATPase α1 that gave rise to the
387 derived ouabain-resistant Na+-K+ ATPase α1 were likely fixed in a stepwise fashion
388 by positive selection (Moore *et al.* 2009).

389 *Leptodactylus ocellatus*, which possess two paralogs of the Na^+-K^+ ATPase α1
390 gene differing in their ouabain sensitivity, contrasts with bufonids not only in the way
391 it confronts cardiotonic toxins (through their diet but not by intern synthesis) but also
392 in the mechanism of molecular evolution that gave rise to this adaptation. Gene
393 duplication in this toad-eating species has promoted neo-functionalization of the
394 duplicated gene, as predicted when positive selection acts over the novel duplicated
395 genes (Hughes 1994; Kondrashov *et al.* 2002; Prince & Pickett 2002; Zhang 2003).
396 Also, gene duplication has facilitated that this species exhibit tissue-specific
397 expression of these paralogs genes, as occurs in insects specialized in eating ouabain-
398 producing plants. This finding demonstrates that both *L.ocellatus* and cardenolide-

399 eating insects have converged in the mechanism of molecular evolution to handle
400 toxin exposure, whereas in toxic bufonids a different strategy to coping with toxins
401 took place.

402 Diverse molecular and physiological mechanisms to endure toxins are not
403 exclusive to amphibians. Recent studies suggest that tiger moths of the family
404 Arctiidae (Lepidoptera) that store high amounts of cardenolides, and Danaini
405 butterflies of the genus *Euploea*, posses a Na+-K+ ATPase α with high affinity for
406 ouabain (Petschenka *et al.* 2012). The midgut and the musculature of this lepidopteran
407 species lack the Na+-K+ ATPase α, which is only expressed in the brain. Whereas in
408 Danaini species the impermeable midgut prevents the cardenolides from entering the
409 body cavity, in the Arctiid species the neural tissue is well protected by the
410 perineurium layer, which function as blood-brain barrier (Petschenka *et al.* 2013b). In
411 these cases, therefore, insensitivity to cardenolides is not a prerequisite for storing
412 these compounds. On the other hand, in bufonid frog species, with high production of
413 toxic compounds, the optimal strategy to cope with them has been to forgo an
414 ouabain-sensitive Na^+-K^+ ATPase α1 and express the resistant form in the gut, skin
415 and brain tissue (and probably in all tissues). Conversely, *L. ocellatus* expresses the
416 ouabain-resistant copy only in the gut and the sensitive copy in the brain. Evidently,
417 the adaptation for tolerating high levels of cardiotonic steroids has taken diverse
418 evolutionary mechanisms associated with the modes each organism faces the toxin
419 (storing, sequestration, feeding or synthesis).

420 Most non toxin-producing organisms have a single copy of the Na^+-K^+
421 ATPase α1 gene whose protein has high affinity for ouabain (Pierre & Xie 2006;
422 Schoner & Scheiner-Bobis 2007; Li & Xie 2009; Nguyen *et al.* 2011). Apart from its
423 transport function, Na^+-K^+ ATPase α isozymes, through their binding with ouabain,

424 display a series of signaling pathways that promotes cell growth, cell motility and
425 other metabolic pathways (Li & Xie 2009; Nguyen *et al.* 2011). Ouabain also
426 regulates blood pressure and heart rate through the central nervous system (Dostanic
427 *et al.* 2004). But what happens in species such as bufonid toads or rat (*Rattus*
428 *norvegicus*) that lack an ouabain-sensitive Na^+-K^+ ATPase α1? Studies have
429 demonstrated that toads of the species *Bufo viridis* synthetize high amounts of
430 cardioactive steroids in dorsal and ventral skin, and in lesser quantities in brain,
431 spleen, liver, testis, heart, kidney, lung, plasma, small and large intestine (Lichtstein *et*
432 *al.* 1992). Also, this compound circulates naturally in the plasma of *R. marina* (Flier
433 *et al.* 1979). Whether this compounds functions exclusively in defense against
434 predators is a matter of investigation but possibly with this high amount of toxin
435 production in the tissues, particularly in the skin, the best strategy to avoiding fatal
436 Na^+-K^+ ATPase α1 inhibition is having only one Na^+-K^+ ATPase α1 resistant to
437 ouabain effects. In the brain of *R. marina* we found the ouabain-resistant copyof Na^+-
438 K^+ ATPase α1, which suggests that even in this well protected tissue ouabain-resistant
439 Na^+-K^+ ATPase α1 is needed.
440 The types of amino acid substitutions conferring ouabain-resistance in *L.*
441 *ocellatus* and toads are also different. Although there is an evident convergence in the
442 phenotypic adaptive trait defined by an ouabain-resistant Na^+-K^+ ATPase α1, the
443 phenotypic convergence is achieved by distinctive amino acid substitutions. This fact
444 supports the idea of independent origin of resistance among amphibian species, and
445 contrasts with the pattern of parallelism in amino acidic substitutions occurring in
446 cardenolide-adapted insects. Specifically, in milkweed butterflies of the subfamily
447 Danainae (Lepidoptera), the sensitivity to cardenolides was reduced in a stepwise
448 manner during the evolution of this clade (Petschenka *et al.* 2013a). The ancestral

449 amino acid leucine at position 111, present in the subtribes Euploenia and Amaurina,
450 was replaced by valine in the subtribe Danaina. This change was followed by the
451 substitution N122H, exclusive to the most ouabain-resistant species, *Danaus*
452 *plexippus* and *D.erippus*. In the milkweed beetles *Chrysochus cobaltinus* and *C.*
453 *auratus* (Coleoptera),the same order of substitutions occurred (Aardema *et al.* 2012).
454 Despite the fact that several amino acid substitutions conferring ouabain insensitivity
455 have been tested by mutagenesis and structural assays at the H1-H2 domain of the
456 Na^+-K^+ ATPase α1 (Croyle *et al.* 1997; Dobler *et al.* 2012), the convergence in the
457 same amino acidic positions demonstrate the molecular constraints on the molecule.
458 In toads and frogs on the contrary, there are independent molecular pathways to gain
459 the phenotypic adaptation.
460 In vertebrates, ouabain binding to Na^+-K^+ ATPase α1 is functionally important
461 in the inotropic (increase contraction) effect of cardiac muscle, among other
462 functions. In rats that have a ouabain-resistant Na^+-K^+ ATPase α1 isozyme, this
463 function is accomplished by ouabain-sensitive Na^+-K^+ ATPase α2 expressed in the
464 cardiac muscle along with Na^+-K^+ ATPase α1 (Fricke 1984). In this manner the Na^+-
465 K^+ ATPase α2 mediates the inotropic effect of ouabain whereas the Na^+-K^+ ATPase
466 α1 modulates the toxic effect of ouabain (Dostanic *et al.* 2004; Bagrov *et al.* 2009). In
467 the brain of *R. marina* we found expressed genes whose coding protein had high
468 similarity with Na^+-K^+ ATPase α2 from *Xenopus tropicalis* andNa^+-K^+ ATPase α3 of
469 *Gallus gallus*. These Na^+-K^+ ATPase α isozymes are ouabain-sensitive and the union
470 with ouabain modulates their functioning (Schaefer *et al.* 2011). Additionally, in
471 studies of Na^+-K^+ ATPase α1 from *R. marina* and *Bufo viridis* experimental assays
472 with brain membranes and ouabain concentrations demonstrated the existence of a
473 ouabain-sensitive Na^+-K^+ ATPase α in the brain of this species (Morris *et al.* 1997).

Therefore, as occurs in the rat brain, we hypothesize that the Na^+-K^+ ATPase α2 in the brain of *R. marina* plays an important physiological function related with its binding to ouabain.

Mice (*Mus musculus*), as with bufonid species, also have an ouabain-resistant Na^+-K^+ ATPase α1 protein and have been used in genetic experiments investigating the ouabain-resistant Na^+-K^+ ATPase α1 by making mice with an ouabain-sensitive Na^+-K^+ ATPase α1 protein (Loreaux *et al.* 2008). These experiments demonstrated that mice with the sensitive Na^+-K^+ ATPase α1 protein responded to salt load with a greater level of natriuresis (Na^+ elimination) than mice with the wild type (ouabain-resistant) Na^+-K^+ ATPase α1 protein. This confirms the physiological role of the ouabain-binding site. In organisms such as rats, mice and toads, which lack ouabain-sensitive Na^+-K^+ ATPase α1, other mechanisms of sodium excretion in conditions of salt load may act. Examining the transcriptomic of other tissues from bufonid toads such as bladder and kidney, would provide us with a more complete view of the expression of the ouabain-resistant Na^+-K^+ ATPase α1.

Finally, to understand the complete association between Na^+-K^+ ATPase α1 and toxin resistance, structural assays over the entire sequence of the molecule will allow us to elucidate if other sites in the enzyme are responsible for the ouabain resistance. Future studies on the physiological aspects that act synergistically with amphibian resistance to cardioactive compounds, such us toxin excretion and synthesis, will broaden our knowledge about this interesting adaptation.

Acknowledgments

We are very grateful with Lucas Barrientos and Carolina Ortiz, who collaborated in the specimen's collection and field assistance. *Ministerio de Ambiente y DesarrolloSostenible,* Colombia, granted all the permissions and licenses needed to carry on this research. This study was possible for the funding granted by *Facultad de Ciencias* of *Universidad de los Andes*, Bogotá, Colombia through its "*ProyectoSemilla*" graduate grant. We want to thank Anna Savage and Brian Gratwicke from the Smithsonian National Zoo for providing the tissues of *Atelopus zeteki*. We want thank to Ying Zhen for training in cDNA library preparation and especially for her useful advice on bioinformatics analysis. We also want to thank Grant Morahan from the University of Western Australia for providing scaffold of the partially assembled genome from *Rhinella marina*. Also, we want to thank to Alejandro Reyes Muñoz from *Universidad de los Andes* for his useful comments in the bioinformatics analysis. We are very grateful with Adriana Bernal from Universidad de los Andes and with John Malone from University of Connecticut for their useful comments on the manuscript.

514 **Tables**

515 **Table 1**

516 Specimen information showing the location of the samples, the tissues used by RNAseq assay and the number of reads obtained from each

517 tissue.

518

Especie	**Familia**	**Field Collection number**	**Country**	**Department**	**Municipality**	**Latitude**	**Longitude**	**Tissue**	**# reads**
***Atelopus* sp.**	Bufonidae		Colombia	Chocó	Nuquí	W.77.5026	N.5.5705	Skin	10,680,187
Atelopus zeteki	Bufonidae	PGF8	Panamá	Chame	Sorá	W.80.0092	N.8.63443	Skin	11,812,195
Engystomops pustulosus	Leiuperidae	AJC 3734	Colombia	Tolima	Mariquita	W.74.8894	N.5.2627	Brain	10,516,797
Engystomops pustulosus	Leiuperidae	AJC 3734	Colombia	Tolima	Mariquita	W.74.8894	N.5.2627	Stomach	10,800,157
Leptodactylus	Leptodactylidae	AJC 3653	Colombia	Vichada	Puerto Carreño	W.67.3881	N.6.02074	Brain	11,249,796

ocelatus									
Leptodactylus	Stomach	AJC 3653	Colombia	Vichada	Puerto Carreño	W.67.3881	N.6.02074	Stomach	8,633,215
ocelatus									
Rhinella	Bufonidae	AJC 3733	Colombia	Meta	San Juan de	W.73.85179	N.3.22821	Skin	8,418,241
granulosa					Arama				
Rhinella	Bufonidae	AJC 3733	Colombia	Meta	San Juan de	W.73.85179	N.3.22821	Stomach	11,225,142
granulosa					Arama				
Rhinella	Bufonidae	AJC 3732	Colombia	Tolima	Mariquita	W.74.8894	N.5.2627	Brain	8,302,698
marina									
Rhinella	Bufonidae	AJC 3732	Colombia	Tolima	Mariquita	W.74.8894	N.5.2627	Skin	9,397,727
marina									

Table 2

Number of reads before and after the trimming process.

Specie	Tissue	#Raw reads	#Reads after trimming
Atelopus sp.	Skin	10,680,187	9,788,931 (91,65%)
Atelopus zeteki	Skin	11,812,195	10,649,259 (90,155%)
Engystomops pustulosus	Brain	10,516,797	9,651,095 (91,76%)
Engystomops pustulosus	Stomach	10,800,157	9,930,036 (91,94%)
Leptodactylus ocellatus	Brain	11,249,796	10,364,439 (92,13%)
Leptodactylus ocellatus	Stomach	8,633,215	7,995,214 (92,60%)
Rhinella marina	Brain	8,302,698	7,608,887 (91,64%)
Rhinella marina	Skin	9,397,727	8, 573,954 (92,23%)
Rhinella granulosa	Stomach	11,225,142	10,227,412 (91,11)
Rhinella granulosa	Skin	8,418,241	7,689,652 (91,34)

Table 3

Transcripts statistics obtained with the parameter min_contig_length modified between 100 and 200. Note that transcriptomes assembled with a minimum length of 200bp have a higher N50 and therefore, are less fragmented than the transcripts generated with the parameter 100 bp as minimum contig length.

Specie/tissue	**Minimun contig length (bp)**	**Number of Transcripts**	**Maximum contig length (bp)**	**N50**
Atelopus sp./Skin	100	94,506	12,423	885
	200	56,653	12,423	1,090
A. Zeteki/Skin	100	95,638	29,877	457
	200	49,481	29,877	649
E. Pustulosus/Brain	100	126,476	8,557	1,355
	200	76,038	8,557	1,587
E. Pustulosus/Stomach	100	83,210	14,715	691
	200	49,604	14,715	901
*L. ocellatus/*Brain	100	124,549	16,023	1,695
	200	78,246	16,023	1,943
L. ocellatus/Stomach	100	79,231	10,119	670
	200	47,933	10,119	871
R. marina/Brain	100	123,719	14,242	1,085
	200	72,912	14,042	1,362
R. marina/Skin	100	89,154	9,523	714

	200	49,123	9,498	944
R. granulosa/Stomach	100	97,341	7,877	873
	200	63,813	6,919	577
R. granulosa/Skin	100	97,370	7,877	871
	200	56,475	8,167	1,113

552 **Table 4**. **Table 4**. Basic quality metrics obtained for the transcripts assembled with Velvet/Oases. Note that for all the samples, except for *R.*

553 *granulosa*/Stomach, the N50 values obtained with Oases assemblies were higher than those obtained with Trinity.

554

555

Specie/Tissue	**#Transcripts**	**Used reads**	**# Loci**	**Minimum contig length (bp)**	**Maximum contig Length (bp)**	**N50 (bp)**
***A. zeteki*/Skin**	33,793	7,886,659/10,649,259	25,992	100bp	8,108	1,065
***A. sp.* /Skin**	40,912	7,907,593/9,788,931	32,996	94pb	12,448	1,244
***R. marina* /Brain**	52,028	5,702,594/7,608,887	41,460	87pb	14,077	1,512
***R. marina* /Skin**	33,793	6,949,932/8,573,954	29,943	100bp	8,108	1,065
***R. granulosa* /Skin**	34,151	5,678,354/7,689,652	34,151	93bp	7,616	1,114
***R. granulosa* /Stomach**	89,010	8,429,563/10,227,413	50,704	88bp	5,765	371
***L. ocellatus* /Stomach**	29,243	6,511,853/7,995,214	23,931	100bp	12,107	1,055

***L. Ocellatus* /Brain**	48,240	8,364,688/10,364,439	43,265	89pb	14,813	2,004
***E. pustulosus* /Brain**	59,605	7,484,005/9,651,095	43,249	94bp	7,735	1,651
***E. pustulosus* /Stomach**	30,523	8,075,678/9,930,036	25,920	94bp	12,485	1,078

556

557

Table 5. Na+-K+ ATPase α isozyme genes. Length of the best open reading frame (ORF) of each transcript and percentage of identity with homologous proteins of *R. marina* from Uniprot-swissprot database and *Xenopus tropicalis* and *Xenopus laevis* from RefSeq database.

Specie	Tissue	Na+-K+ ATPase α isozyme	ouabain-resistant/ sensitive	Best ORF Length	% IdentityNa+-K+ ATPase α1 proteins
R. granulosa	Skin	α1	Resistant	3075	98.54%*R. marina* P30714.2/Na+-K+ ATPase α1 Swiss-Prot
R. marina	Skin	α1	Resistant	3069	99.32% *R. marina* P30714.2/Na+-K+ ATPase α1 Swiss-Prot
A. zeteki	Skin	α1	Resistant	3075	96% *R. marina* P30714.2/Na+-K+ ATPase α1 Swiss-Prot
A. sp.	Skin	α1	Resistant	3075	97% *R. marina* P30714.2/Na+-K+ ATPase α1 Swiss-Prot
L. ocelltus	Stomach	α1	Resistant	3075	94% *X. laevis* NM_001090595.1/Na+-K+ ATPase α1 RefSeq
L. ocelltus	Stomach	α1	Sensitive	3075	95% *X. laevis*NM_001090595.1/Na+-K+ ATPase α1 RefSeq
E. pustulosus	Stomach	α1	Sensitive	3072	95% *R. marina* P30714.2/Na+-K+ ATPase α1 Swiss-Prot
R. granulosa	Stomach	α1	Resistant	3075	99% *R. marina* P30714.2/Na+-K+ ATPase α1 Swiss-Prot

R. marina	Brain	α1	Resistant	3063	99% *R. marina* P30714.2/Na+-K+ ATPase α1 Swiss-Prot
R. marina	Brain	α2	Sensitive	3063	95% *X. tropicalis* NM_001102969.1/Na+-K+ ATPase α2 RefSeq
R. marina	Brain	α3	Sensitive	3075	95% *X. tropicalis*NM_001126894.1/Na+-K+ ATPase α3 RefSeq
L. ocellatus	Brain	α1	Sensitive	3069	98% *X. tropicalis*NM_001126894.1/Na+-K+ ATPase α3 RefSeq
L. ocellatus	Brain	α2	Sensitive	3063	96%*X. tropicalis*NM_001102969.1/Na+-K+ ATPase α2 RefSeq
L. ocellatus	Brain	α3	Sensitive	3075	97% *X. tropicalis*NM_001126894.1/Na+-K+ ATPase α3 RefSeq
E. pustulosus	Brain	α1	Sensitive	3063	95% *R. marina* P30714/Na+-K+ ATPase α1 Swiss-Prot
E. pustulosus	Brain	α2	Sensitive	3063	95% *X. tropicalis*NM_001102969.1/Na+-K+ ATPase α2 RefSeq
E. pustulosus	Brain	α3	Sensitive	3078	97% *X. tropicalis*NM_001126894.1/Na+-K+ ATPase α3 RefSeq

561

562 **Figures legends**

563 **Figure 1**. Percentage of reads aligned to transcripts produced by either minimum
564 contig length of 100 bp or 200 bp.

565 **Figure 2**. Number of proteins recovered to nearly full-length (>80% of their length)
566 by the transcripts generated by the two types of assemblies (parameter
567 "min_contig_length" 100bp versus 200bp).

568 **Figure 3**. Comparison between the N50 values obtained for the assembly generated
569 by the parameter with minimum contig length of 100bp versus the parameter with
570 minimum contig length of 200bp.

571 **Figure 4**. Amino acid sequences of the ouabain-binding site of the Na^+-K^+ ATPase
572 α 1 isozyme (reference: Na^+-K^+ ATPase α 1 protein of the specie *Xenopus*
573 *tropicalis*(NP_989407.1 from Refseq). Note that all amino acid sequences of bufonids
574 confer resistance to ouabain (sequence: RKASDLEPDNDN). Note that the
575 paralogous copy of Na^+-K^+ ATPase α 1 of *L. ocellatus* that is ouabain-resistant is
576 only expressed in the stomach (sequence: RTATEDEPQNDN). Dots represent
577 identity with the reference sequence.

578 **Figure 5.** Amino acid sequences of Na^+-K^+ ATPase α2 isozyme (reference: Na^+-K^+
579 ATPase α 2 of *Xenopus tropicalis* (NP_001096439.1 from Refseq). Dots represent
580 identity with the reference sequence.

581 **Figure 6**.Amino acid sequences of Na^+-K^+ ATPase α3 isoform (reference: Na^+-K^+
582 ATPase α3 protein of the specie *Xenopus tropicalis (*NP_001120366.1 from Refseq).
583 Dots represent identity with the reference sequence.

584 **Figure 7**. Phylogenetic relationships between Na^+-K^+ ATPase α isozyme genes of
585 fishes, mammals, amphibians, reptiles and birds.

586 **Figure 8**. Phylogenetic relationships between Na^+-K^+ ATPase α 1genes of
587 amphibians. The green chart denotes the duplication event of the Na^+-K^+ ATPase α1
588 gene expressed in the stomach of *Leptodactylus ocellatus*. The red chart denote the
589 divergence of the ouabain-resistant Na^+-K^+ ATPase α 1 gen from Bufonids from the
590 ouabain-sensitive Na^+-K^+ ATPase α 1 genes of the others amphibians.
591
592
593

Figure 1.

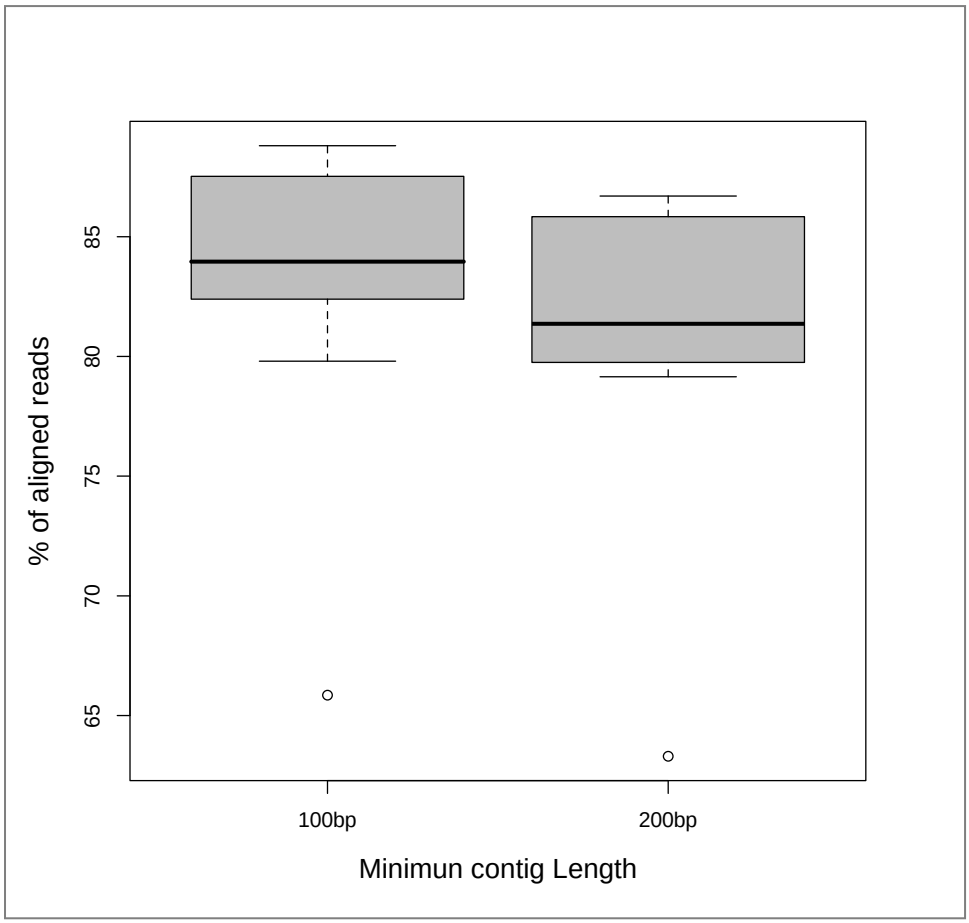

% of aligned reads
65
70
75
80
85
100bp
200bp
Minimun contig Length

Figure 2

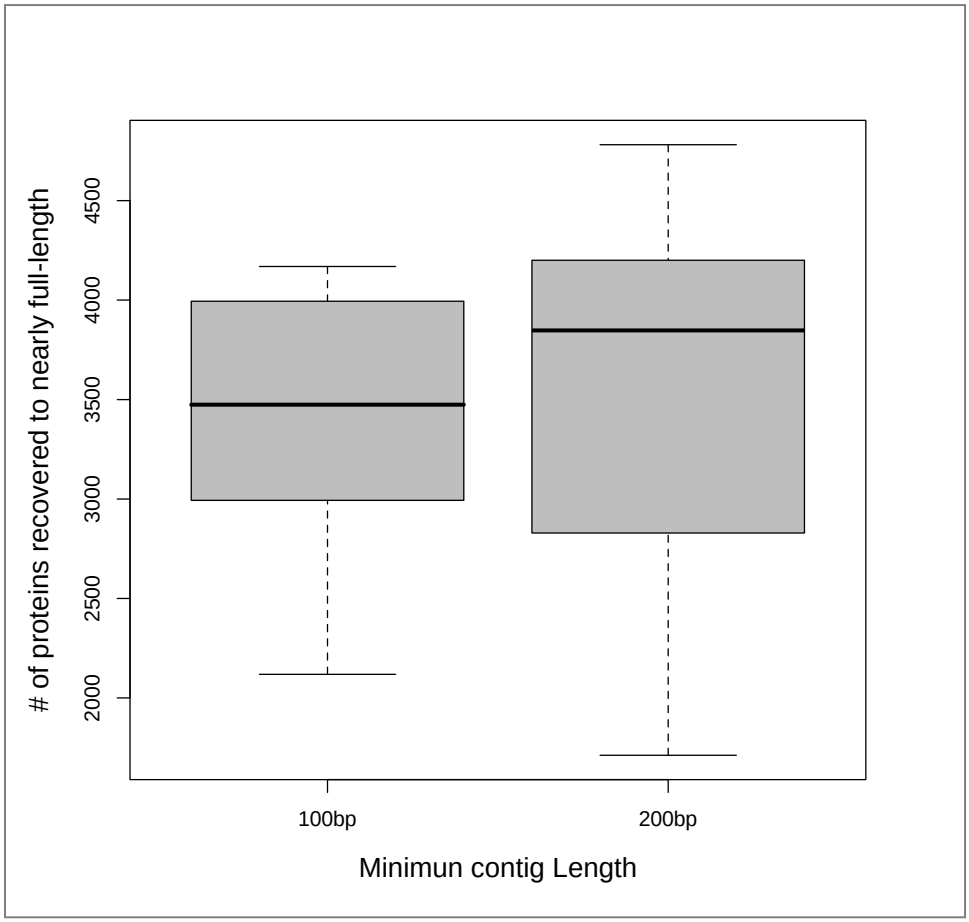

# of proteins recovered to nearly full-length
2000
2500
3000
3500
4000
4500
100bp
200bp
Minimun contig Length

Figure 3

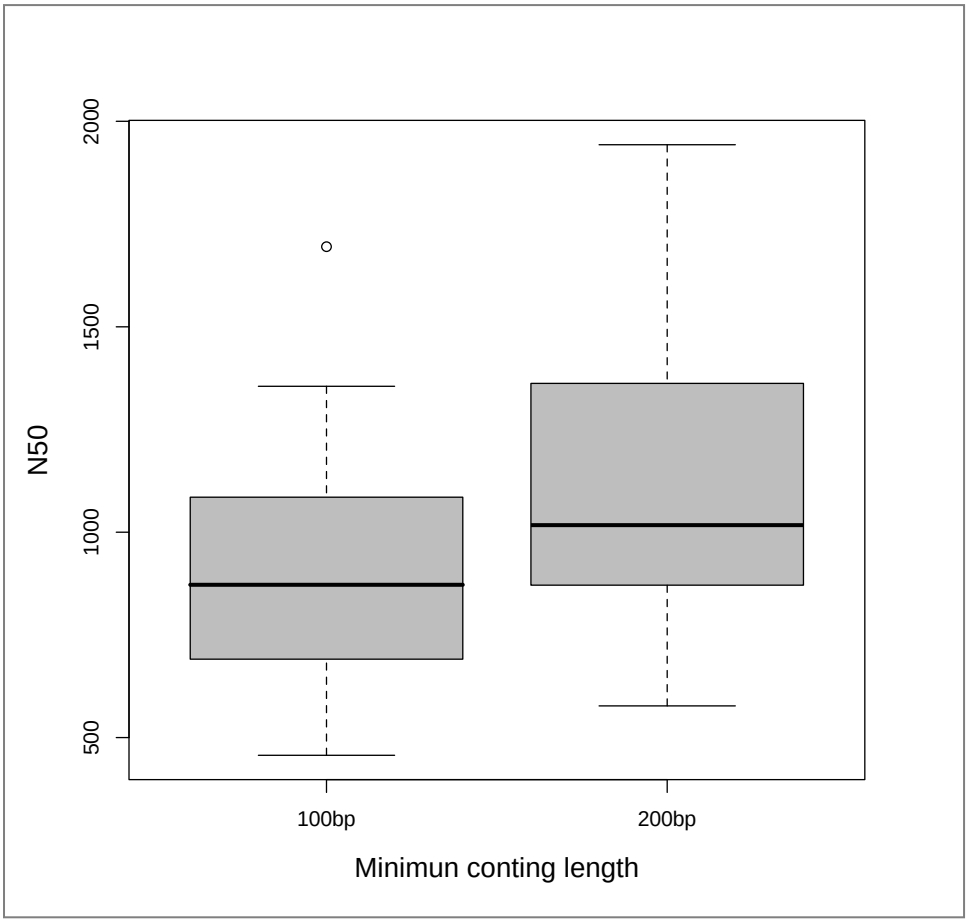


Figure 4

Na^+-K^+ ATPase α1

Amino acid positions at the ouabain-binding site

Species	**111**	**112**	**113**	**114**	**115**	**116**	**117**	**118**	**119**	**120**	**121**	**122**
Xenopustropicalis	T	A	A	T	E	E	E	P	Q	N	D	N
Engystomops pustulosus	Q	.	.	.	.	.	.	.	.	.	.	.
Leptodactylus ocellatus (sensitive)	Q	.	.	.	.	.	.	.	.	.	.	.
Leptodactylus ocellatus (resistant)	R	T	.	.	.	D	.	.	.	.	.	D
Rhinella marina	R	K	.	S	D	L	.	.	D	.	.	.
Atelopus zeteki	R	K	.	S	D	L	.	.	D	.	.	.
Atelopus sp.	R	K	.	S	D	L	.	.	D	.	.	.
Rhinella granulosa	R	K	.	S	D	L	.	.	D	.	.	.

Figure 5

Na^+-K^+ ATPase α2

Amino acid positions at the ouabain-binding site

Species	**111**	**112**	**113**	**114**	**115**	**116**	**117**	**118**	**119**	**120**	**121**	**122**
Xenopustropicalis	Q	I	A	M	D	E	P	I	N	D	N	N
Engystomops pustulosus	.	.	.	.	.	.	.	.	.	.	.	.
Leptodactylus ocellatus	.	.	.	.	.	.	.	.	.	.	.	.
Rhinella marina	T	V	.	I	E	D	D	P	T	.	.	.

Figure 6

Na^+-K^+ ATPase α3

Amino acid positions at the ouabain-binding site

Species	**111**	**112**	**113**	**114**	**115**	**116**	**117**	**118**	**119**	**120**	**121**	**122**
Xenopustropicalis	L	A	S	M	E	E	P	S	G	D	N	N
Engystomops pustulosus	Q	.	G	T	.	D	.	A	.	.	.	.
Leptodactylus ocellatus	Q	.	G	T	.	D	.	A	.	.	.	.
Rhinella marina	L	A	G	T	.	D	E	P	S	R	D	.

Figure 7.

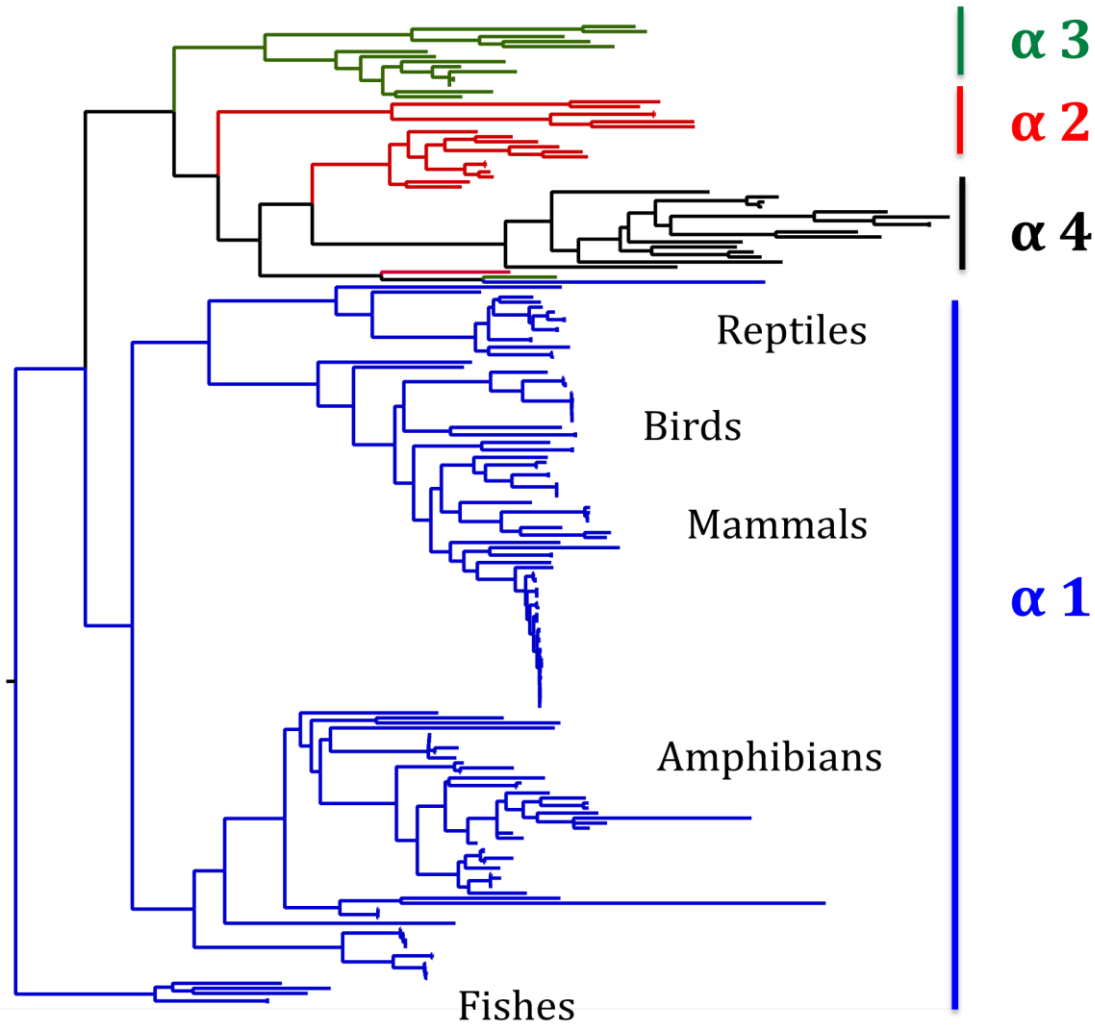
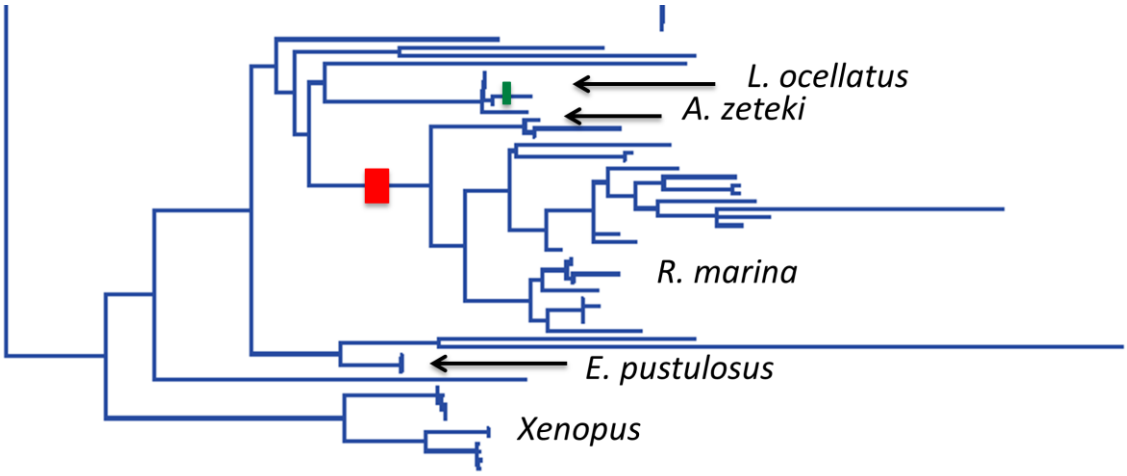


Figure 8



Supplemental Material

Figure S1. Quality scores across all bases from the raw reads obtained from stomach tissue of *Leptodactylus ocellatus* (as an example).

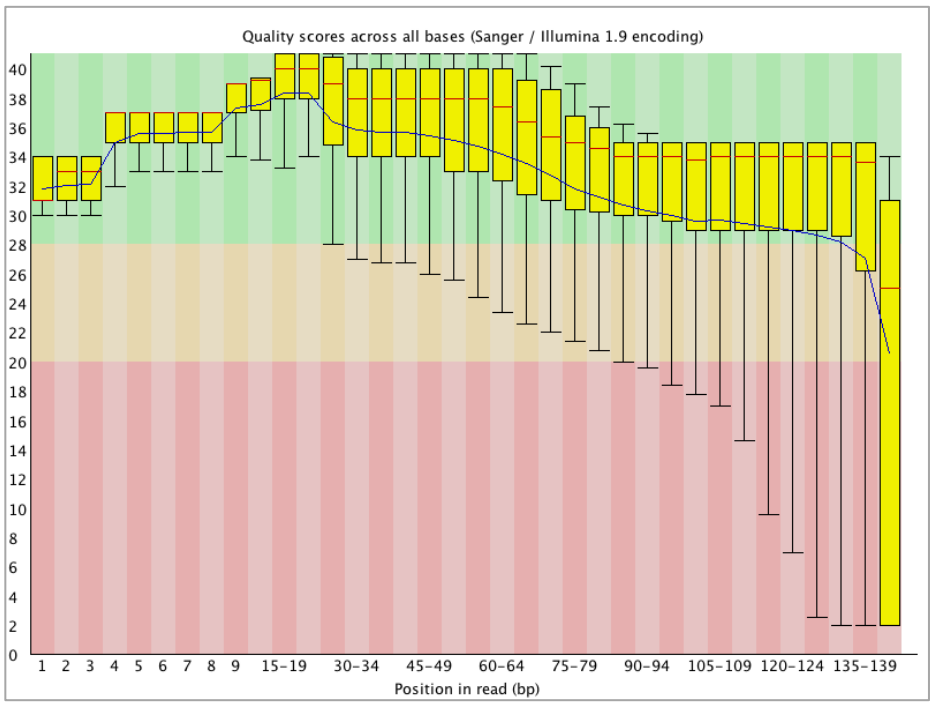


Figure S2. Alignment of the reads from R. *granulosa* stomach to the consensus transcript of the Na^{+}-K^{+} ATPase gene. Note that the outlined region corresponds to the H1-H2 domain (positions 111-122).

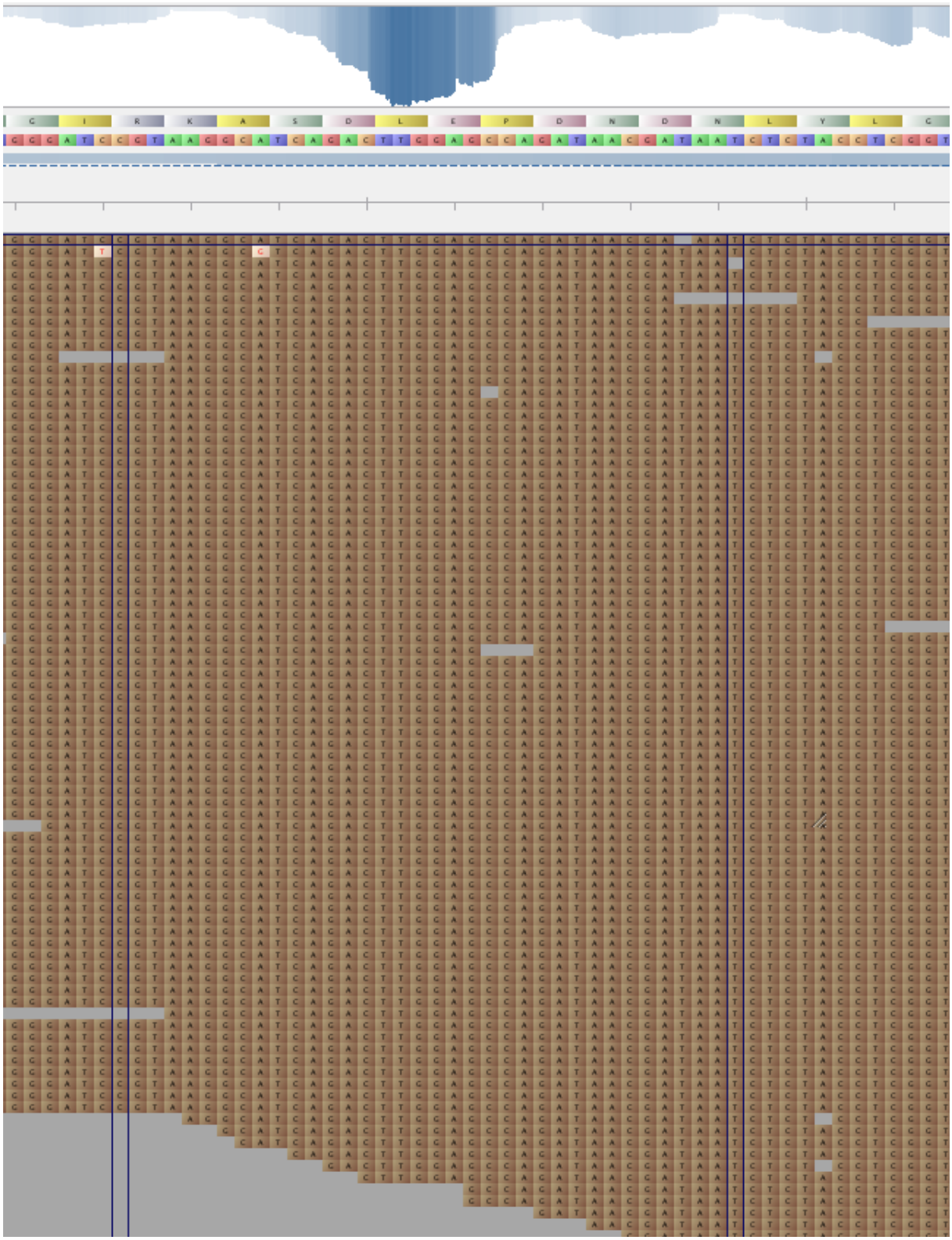


Figure S3. Alignment of the reads from R. *granulosa* stomach to the consensus transcript of the Na^{+}-K^{+} ATPaseα 1 gene. Note that the small white points are variable sites and/or sequencing errors. The outlined vertical region corresponds to the H1-H2 domain (positions 111-122). The outlined horizontal region shows the read coverage of the transcript from nucleotide position 370 to 1434.

Figure S4. Alignment of the reads from *L. ocellatus* stomach to the consensus transcript of the Na^+-K^+ ATPase α1 gene. Note that the outlined region corresponds to the H1-H2 domain (positions 111-122). The variable sites (white color) correspond to reads derived from transcripts of the ouabain-resistant Na^+-K^+ ATPase α1 gene.

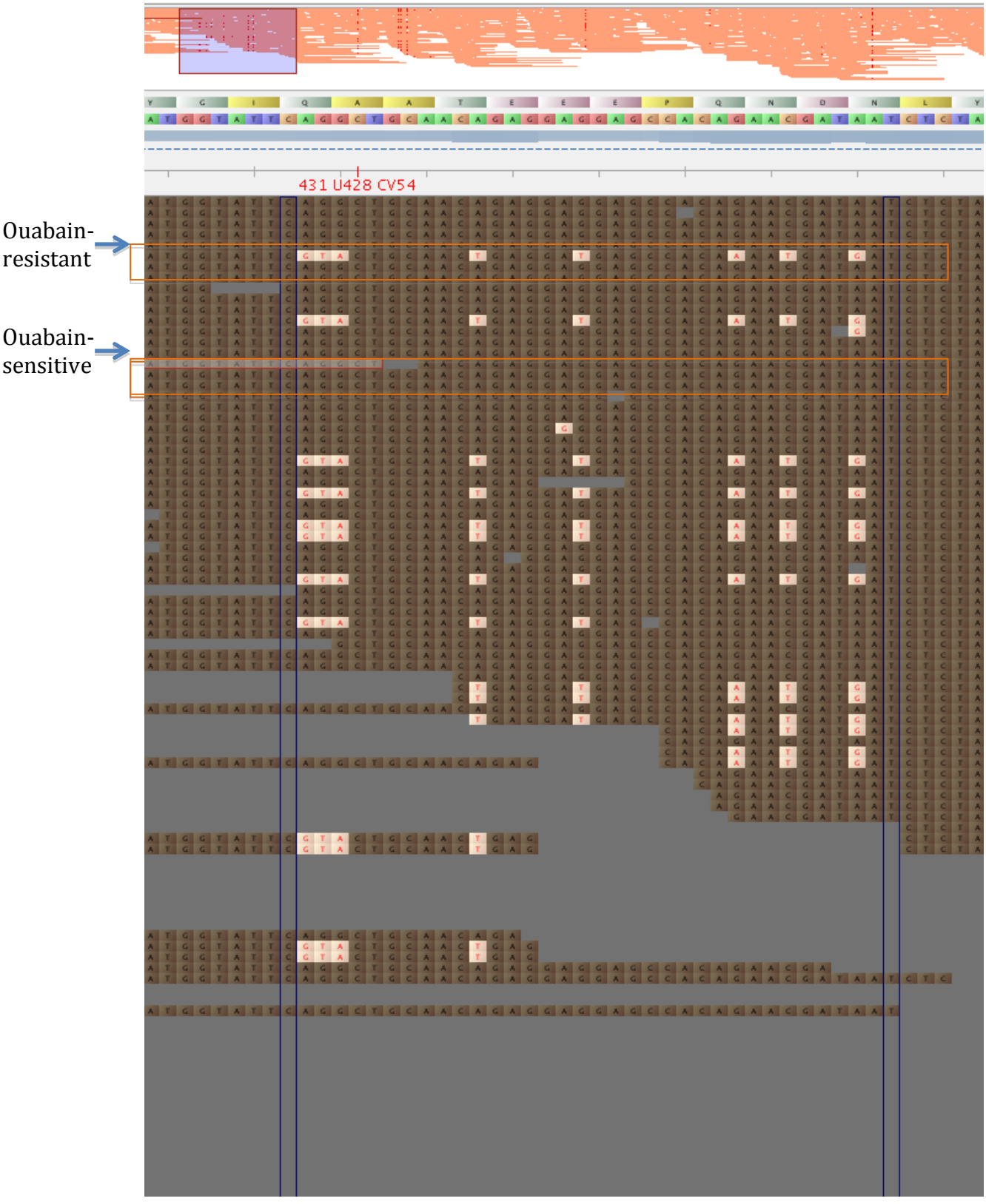
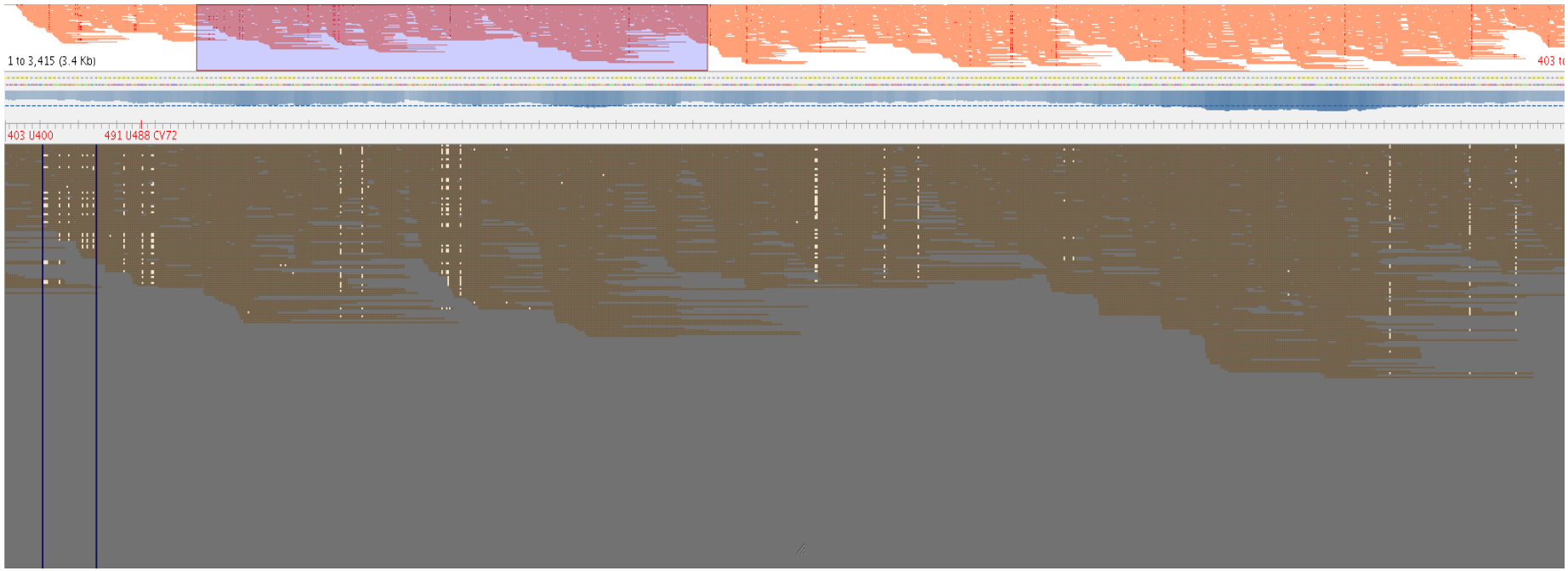


Figure S5. Alignment of the reads from *L. ocellatus* stomach to the consensus transcript of the Na^+-K^+ ATPase gene. Note that the small white points are variable sites and/or sequencing errors. The outlined vertical region corresponds to the H1-H2 domain (positions 111-122). The outlined horizontal region shows the read coverage of the transcript from nucleotide position 403 to 1466



References

Aardema M.L., Zhen Y. & Andolfatto P. (2012). The evolution of cardenolide-resistant forms of Na+,K+-ATPase in Danainae butterflies. *Molecular Ecology*, 21, 340-349.

Agrawal A.A., Petschenka G., Bingham R.A., Weber M.G. & Rasmann S. (2012). Toxic cardenolides: chemical ecology and coevolution of specialized plant–herbivore interactions. *New Phytologist*, 194, 28-45.

Ashburner M., Ball C.A., Blake J.A., Botstein D., Butler H., Cherry J.M., Davis A.P., Dolinski K., Dwight S.S., Eppig J.T., Harris M.A., Hill D.P., Issel-Tarver L., Kasarskis A., Lewis S., Matese J.C., Richardson J.E., Ringwald M., Rubin G.M. & Sherlock G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*, 25, 25-9.

Bagrov A.Y., Shapiro J.I. & Fedorova O.V. (2009). Endogenous Cardiotonic Steroids: Physiology, Pharmacology, and Novel Therapeutic Targets. *Pharmacological Reviews*, 61, 9-38.

Barrueto F., Kirrane B., Cotter B., Hoffman R. & Nelson L. (2006). Cardioactive steroid poisoning: A comparison of plant- and animal-derived compounds. *J. Med. Toxicol.*, 2, 152-155.

Blanco G. & Mercer R.W. (1998). Isozymes of the Na-K-ATPase: heterogeneity in structure, diversity in function. *American Journal of Physiology - Renal Physiology*, 275, F633-F650.

Chain F., Ilieva D. & Evans B. (2008). Duplicate gene evolution and expression in the wake of vertebrate allopolyploidization. *BMC Evolutionary Biology*, 8, 43.

Christin P.-A., Weinreich D.M. & Besnard G. (2010). Causes and evolutionary significance of genetic convergence. *Trends in Genetics*, 26, 400-405.

Conte G.L., Arnegard M.E., Peichel C.L. & Schluter D. (2012). The probability of genetic parallelism and convergence in natural populations. *Proc Biol Sci*, 279, 5039-47.

Crow K.D. & Wagner G.P. (2006). What Is the Role of Genome Duplication in the Evolution of Complexity and Diversity? *Molecular Biology and Evolution*, 23, 887-892.

Croyle M.L., Woo A.L. & Lingrel J.B. (1997). Extensive Random Mutagenesis Analysis of the Na+/K+-ATPase α Subunit Identifies Known and Previously Unidentified Amino Acid Residues that Alter Ouabain Sensitivity Implications for Ouabain Binding. *European Journal of Biochemistry*, 248, 488-495.

Dobler S., Dalla S., Wagschal V. & Agrawal A.A. (2012). Community-wide convergent evolution in insect adaptation to toxic cardenolides by substitutions in the Na,K-ATPase. *Proceedings of the National Academy of Sciences*, 109, 13040-13045.

Dostanic I., Schultz J.E.J., Lorenz J.N. & Lingrel J.B. (2004). The α1 Isoform of Na,K-ATPase Regulates Cardiac Contractility and Functionally Interacts and Co-localizes with the Na/Ca Exchanger in Heart. *Journal of Biological Chemistry*, 279, 54053-54061.

Dostanic-Larson I., Lorenz J.N., Van Huysse J.W., Neumann J.C., Moseley A.E. & Lingrel J.B. (2006). Physiological role of the α1- and α2-isoforms of the Na+-K+-ATPase and biological significance of their cardiac glycoside binding site. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 290, R524-R528.

Duan J., Xia C., Zhao G., Jia J. & Kong X. (2012). Optimizing de novo common wheat transcriptome assembly using short-read RNA-Seq data. *BMC Genomics*, 13, 392.

Emery A.M., Ready P.D., Billingsley P.F. & Djamgoz M.B.A. (1995). A single isoform of the Na+/K+-ATPase α-subunit in Diptera: evidence from characterization of the first extracellular domain. *Insect Molecular Biology*, 4, 179-192.

Feldman C.R., Brodie E.D. & Pfrender M.E. (2012). Constraint shapes convergence in tetrodotoxin-resistant sodium channels of snakes. *Proceedings of the National Academy of Sciences*, 109, 4556-4561.

Flier J.S., Maratos-Flier E., Pallotta J.A. & McIsaac D. (1979). Endogenous digitalis-like activity in the plasma of the toad Bufo marinus. *Nature*, 279, 341-343.

Fricke U. (1984). Two ouabain binding sites in guinea pig heart Na+-K+-ATPase. Differentiation by sodium and erythrosin B. *Basic Res Cardiol*, 79 Suppl, 119-27.

Garraffo H., Andriamaharavo N., Vaira M., Quiroga M., Heit C. & Spande T. (2012). Alkaloids from single skins of the Argentinian toad Melanophryniscus rubriventris (ANURA, BUFONIDAE): An unexpected variability in alkaloid profiles and a profusion of new structures. *SpringerPlus*, 1, 1-15.

Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauceli E., Hacohen N., Gnirke A., Rhind N., di Palma F., Birren B.W., Nusbaum C., Lindblad-Toh K., Friedman N. & Regev A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotech*, 29, 644-652.

Grant T., Colombo P., Verrastro L. & Saporito R. (2012). The occurrence of defensive alkaloids in non-integumentary tissues of the Brazilian red-belly toad Melanophryniscus simplex (Bufonidae). *Chemoecology*, 22, 169-178.

Hayes R.A., Crossland M., Hagman M., Capon R. & Shine R. (2009). Ontogenetic Variation in the Chemical Defenses of Cane Toads (Bufo marinus): Toxin Profiles and Effects on Predators. *J Chem Ecol*, 35, 391-399.

Hittinger C.T. & Carroll S.B. (2007). Gene duplication and the adaptive evolution of a classic genetic switch. *Nature*, 449, 677-81.

Hoekstra H.E. & Coyne J.A. (2007). The locus of evolution: evo devo and the genetics of adaptation. *Evolution*, 61, 995-1016.

Hughes A.L. (1994). The Evolution of Functionally Novel Proteins after Gene Duplication. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 256, 119-124.

Jaisser F., Canessa C.M., Horisberger J.D. & Rossier B.C. (1992). Primary sequence and functional expression of a novel ouabain-resistant Na,K-ATPase. The beta subunit modulates potassium activation of the Na,K-pump. *Journal of Biological Chemistry*, 267, 16895-903.

Jost M.C., Hillis D.M., Lu Y., Kyle J.W., Fozzard H.A. & Zakon H.H. (2008). Toxin-Resistant Sodium Channels: Parallel Adaptive Evolution across a Complete Gene Family. *Molecular Biology and Evolution*, 25, 1016-1024.

Kondrashov F., Rogozin I., Wolf Y. & Koonin E. (2002). Selection in the evolution of gene duplications. *Genome Biology*, 3, research0008.1 - research0008.9.

Langmead B. & Salzberg S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Meth*, 9, 357-359.

Li Z. & Xie Z. (2009). The Na/K-ATPase/Src complex and cardiotonic steroid-activated protein kinase cascades. *Pflugers Arch - Eur J Physiol*, 457, 635-644.

Lichtstein D., Gati I., Haver E. & Katz U. (1992). Digitalis-like compounds in the toad Bufo viridis: tissue and plasma levels and significance in osmotic stress. *Life Sci*, 51, 119-28.

Lingrel J.B. (2010). The Physiological Significance of the Cardiotonic Steroid/Ouabain-Binding Site of the Na,K-ATPase. *Annual Review of Physiology*, 72, 395-412.
Liu K., Warnow T.J., Holder M.T., Nelesen S.M., Yu J., Stamatakis A.P. & Linder C.R. (2012). SATé-II: Very Fast and Accurate Simultaneous Estimation of Multiple Sequence Alignments and Phylogenetic Trees. *Systematic Biology*, 61, 90-106.
Loreaux E.L., Kaul B., Lorenz J.N. & Lingrel J.B. (2008). Ouabain-Sensitive α1 Na,K-ATPase Enhances Natriuretic Response to Saline Load. *Journal of the American Society of Nephrology*, 19, 1947-1954.
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.
Martin A. & Orgogozo V. (2013). THE LOCI OF REPEATED EVOLUTION: A CATALOG OF GENETIC HOTSPOTS OF PHENOTYPIC VARIATION. *Evolution*, 67, 1235-1250.
Milne I., Bayer M., Cardle L., Shaw P., Stephen G., Wright F. & Marshall D. (2010). Tablet—next generation sequence assembly visualization. *Bioinformatics*, 26, 401-402.
Moore D.J., Halliday D.C.T., Rowell D.M., Robinson A.J. & Keogh J.S. (2009). Positive Darwinian selection results in resistance to cardioactive toxins in true toads (Anura: Bufonidae). *Biology Letters*, 5, 513-516.
Morris J.F., Ismail-Beigi F., Butler Jr V.P., Gati I. & Lichtstein D. (1997). Ouabain-sensitive Na+,K+-ATPase activity in toad brain. *Comparative Biochemistry and Physiology Part A: Physiology*, 118, 599-606.
Nguyen A.-N.T., Jansson K., Sánchez G., Sharma M., Reif G.A. & Blanco G. (2011). Ouabain activates the Na-K-ATPase signalosome to induce autosomal dominant polycystic kidney disease cell proliferation. *American Journal of Physiology - Renal Physiology*, 301, F897-F906.
Odom D.T., Dowell R.D., Jacobsen E.S., Gordon W., Danford T.W., MacIsaac K.D., Rolfe P.A., Conboy C.M., Gifford D.K. & Fraenkel E. (2007). Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nature Genetics*, 39, 730-732.
Pál C., Papp B. & Lercher M.J. (2006). An integrated view of protein evolution. *Nature Reviews Genetics*, 7, 337-348.
Parra G., Bradnam K. & Korf I. (2007). CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics*, 23, 1061-1067.
Petschenka G., Fandrich S., Sander N., Wagschal V., Boppré M. & Dobler S. (2013a). STEPWISE EVOLUTION OF RESISTANCE TO TOXIC CARDENOLIDES VIA GENETIC SUBSTITUTIONS IN THE NA+/K+-ATPASE OF MILKWEED BUTTERFLIES (LEPIDOPTERA: DANAINI). *Evolution*, 67, 2753-2761.
Petschenka G., Offe J.K. & Dobler S. (2012). Physiological screening for target site insensitivity and localization of Na+/K+-ATPase in cardenolide-adapted Lepidoptera. *Journal of Insect Physiology*, 58, 607-612.
Petschenka G., Pick C., Wagschal V. & Dobler S. (2013b). Functional evidence for physiological mechanisms to circumvent neurotoxicity of cardenolides in an adapted and a non-adapted hawk-moth species. *Proceedings of the Royal Society B: Biological Sciences*, 280.
Pierre S.V. & Xie Z. (2006). The Na,K-ATPase receptor complex. *Cell Biochemistry and Biophysics*, 46, 303-315.
Prince V.E. & Pickett F.B. (2002). Splitting pairs: the diverging fates of duplicated genes. *Nat Rev Genet*, 3, 827-37.

Schaefer T.L., Lingrel J.B., Moseley A.E., Vorhees C.V. & Williams M.T. (2011). Targeted mutations in the Na,K-ATPase alpha 2 isoform confer ouabain resistance and result in abnormal behavior in mice. *Synapse*, 65, 520-531.

Schmieder R. & Edwards R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics*, 27, 863-4.

Schoner W. & Scheiner-Bobis G. (2007). Endogenous and exogenous cardiac glycosides: their roles in hypertension, salt metabolism, and cell growth. *Am J Physiol Cell Physiol*, 293, C509-36.

Schulz M.H., Zerbino D.R., Vingron M. & Birney E. (2012). Oases: Robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics*.

Simon A. (2011). FASTQC. In.

Singhal S. (2013). De novo transcriptomic analyses for non-model organisms: an evaluation of methods across a multi-species data set. *Molecular Ecology Resources*, 13, 403-416.

Takeyasu K., Lemas V. & Fambrough D.M. (1990). Stability of Na(+)-K(+)-ATPase alpha-subunit isoforms in evolution. *American Journal of Physiology - Cell Physiology*, 259, C619-C630.

Vijay N., Poelstra J.W., Künstner A. & Wolf J.B.W. (2013). Challenges and strategies in transcriptome assembly and differential gene expression quantification. A comprehensive in silico assessment of RNA-seq experiments. *Molecular Ecology*, 22, 620-634.

Yotsu-Yamashita M., Kim Y.H., Dudley S.C., Jr., Choudhary G., Pfahnl A., Oshima Y. & Daly J.W. (2004). The structure of zetekitoxin AB, a saxitoxin analog from the Panamanian golden frog Atelopus zeteki: a potent sodium-channel blocker. *Proceedings of the National Academy of Sciences*, 101, 4346-51.

Zhang J. (2003). Evolution by gene duplication: an update. *Trends in ecology & evolution (Personal edition)*, 18, 292-298.

Zhang J. & Kumar S. (1997). Detection of convergent and parallel evolution at the amino acid sequence level. *Molecular Biology and Evolution*, 14, 527-536.

Zhen Y., Aardema M.L., Medina E.M., Schumer M. & Andolfatto P. (2012). Parallel Molecular Evolution in an Herbivore Community. *Science*, 337, 1634-1637.