Comparative analysis of amphibian genomes: an emerging resource for basic and applied research

Running title: Comparative analysis of amphibian genomes

Tiffany A. Kosch[1], Andrew J. Crawford[2], Rachel Lockridge Mueller[3], Katharina C. Wollenberg Valero[4], Megan L. Power[4], Ariel Rodríguez [5], Lauren A. O'Connell[6], Neil D. Young[1], and Lee F. Skerratt[1]

[1] Faculty of Science, University of Melbourne, Melbourne, Australia

[2] Departamento de Ciencias Biológicas, Universidad de los Andes, Bogotá, Colombia

[3] Department of Biology, Colorado State University, Colorado, USA

[4] School of Biology and Environmental Science, University College Dublin, Dublin, Ireland

[5] Institute of Zoology, University of Veterinary Medicine of Hannover, Hannover, Germany

[6] Department of Biology, Stanford University, California, USA

Corresponding author: tiffany.kosch@unimelb.edu.au

ABSTRACT

Amphibians are the most threatened group of vertebrates and are in dire need of conservation intervention to ensure their continued survival. They have many unique features including a high diversity of reproductive strategies, permeable and specialized skin capable of producing toxins and antimicrobial compounds, multiple genetic mechanisms of sex determination, and in some lineages even the ability to regenerate limbs and internal organs. Although genomics approaches would shed light on these unique phenotypic traits and aid in conservation management, sequencing and assembly of amphibian genomes has lagged far behind other taxa due to their comparatively large genome sizes. Fortunately, the development of long-read sequencing technologies and initiatives has led to a recent burst of new amphibian genome assemblies. Although growing, the field of amphibian genomics suffers from the lack of annotation resources, tools for working with challenging genomes, and lack of high-quality assemblies in multiple clades of amphibians. Here we analyze 32 publicly available amphibian genomes to evaluate their usefulness for functional genomics analysis. We report considerable variation in assembly quality and completeness, and report some of the highest transposable element and repeat contents of any vertebrate, which is associated with climate. We also provide evidence of conserved genome synteny despite the long divergence times of this group but show that chromosome naming and orientation have been inconsistent across genome assemblies. Additionally, we discuss sequencing gaps in the phylogeny and suggest key targets for future sequencing endeavors. Lastly, we propose increased investment in amphibian genomics research to promote their conservation.

KEYWORDS

Amphibian genomes; comparative genomics; transposable elements; repeat expansion; genome synteny

INTRODUCTION

Amphibians are an ancient lineage of vertebrates that predate amniotes by more than 100 million years. Despite the considerable age of this lineage, amphibians are now the most threatened group of vertebrates with more that 40% of species threatened by factors such as habitat change, disease, and over-exploitation (IUCN, 2022; Scheele et al., 2019). Notably, many of these threats are hard to reverse, suggesting that novel approaches that utilize genomic resources may lead to improved management decisions for some of the most endangered taxa (Kosch et al., 2022; Scheele et al., 2014).

We are only just beginning to understand the genetic basis of many of the unique features of amphibians. Amphibians exhibit a high diversity of reproductive strategies including biphasic and direct development, uniparental and biparental care, mouth and gastric brooding, and foam nesting (Brown et al., 2010; Nunes-de-Almeida et al., 2021; Schulte et al., 2020). They also have specialized skin capable of producing complex compounds of interest for drug discovery for the development of antimicrobial drugs and analgesics (Daly et al., 2000; De Angelis et al., 2021; Liu et al., 2020). Amphibians occur across habitat types from rainforests to deserts, freshwater streams to salt marshes, and tropical to arctic climates (Duellman, 1999), but it is unclear how this ecological diversity is reflected in genome composition. Salamanders are an important resource for transplant and regeneration research due to their ability to regenerate limbs and internal organs (Elewa et al., 2017; Nowoshilow et al., 2018). Amphibians also have many of the same immune components of mammals making them an important model resource for immunology (Paiola et al., 2023; Robert, 2020).

Despite the obvious value of amphibian genomes for research on ecology, evolution, medicine, and improving their conservation, until recently, the generation of amphibian reference genomes has

been markedly slower than other vertebrates (Hotaling et al., 2021a). This lag can be attributed to high costs and the computational challenges of assembling their often large and complex genomes (Sun et al., 2020). Recent advances in sequencing technologies such as long read sequencing and assembly algorithms that incorporate hybrid approaches have circumvented many of these challenges leading to a surge of high quality, chromosome-level reference genomes. The next challenge will be developing the tools for annotation and comparative analyses of these large genomes.

In this study, we provide a state-of-the-field synthesis of amphibian reference genome assemblies by analyzing 32 publicly available amphibian genomes. We evaluate assembly quality, sequencing technology, gene completeness, transposable element and repeat content and its ecological correlates, taxonomic representation, and synteny.

MATERIALS AND METHODS

Genomes

A search conducted on the NCBI genome website using search term “amphibians” conducted on November 1, 2021 revealed there were 39 amphibian genomes from 30 species. All genome files in fasta format were downloaded for assessment. One genome was selected for each species. If there was more than one draft of a genome, the most recent draft and/or the primary haplotype was selected. In cases where there were multiple versions sequenced by different groups, the best genome was selected by lowest scaffold number. We excluded the *Ranitomeya imitator* genome as the genome was later retracted from NCBI. Entire genomes (including uncharacterized contigs but excluding mitochondrial genomes) were used for assessment unless indicated otherwise.

103 Genome databases NCBI Genomes, NCBI RefSeq (O'Leary et al., 2016), Ensembl (Cunningham et al.,
104 2022), UCSC Genome Browser (Lee et al., 2022), and Genomes on a Tree (GoaT) (Sotero-Caio et al.,
105 2021) were searched for information on the 34 amphibian species with reference genomes including
106 chromosome number, annotation data, proteome availability, C-value, and sequencing technology.
107 Read length was classified as "short" for Illumina sequencing and "long" for PacBio and Nanopore
108 technologies.
109
110 A search for amphibian proteome datasets on NCBI RefSeq (O'Leary et al., 2016), Ensembl
111 (Cunningham et al., 2022), and UCSC Genome Browser (Lee et al., 2022) databases on June 24, 2022
112 revealed 11 proteomics datasets.
113
114 A search of the NCBI Organelle database on 15, February 2023 using search term "amphibian"
115 resulted in 353 mitochondrial genomes belonging to 345 species (Table S11). Seventeen
116 mitochondrial genomes overlapped with the amphibian nuclear genomes analyzed in this study.
117
118 ***Reference genome availability summary***
119 The GoaT online database (Sotero-Caio et al., 2021) was searched on January 6, 2023 to summarize
120 genomes in progress or publicly available using the search terms "tax_tree(Amphibia) AND
121 tax_rank(species) AND sequencing_status=in_progress" or "tax_tree(Amphibia) AND
122 tax_rank(species) AND sequencing_status=insdc_open". The same search terms were used to
123 summarize publicly available genomes for mammals, birds, and non-avian reptiles with the
124 "tax_tree" search term replaced by appropriate Class.
125

126 ***Genome quality analyses***

127 Genome quality assessment was performed with BBMap (v. 38.76). Benchmarking Universal Single-
128 Copy Orthologs (BUSCO) were summarized with the BUSCO tool (v. 5.1.2) (Manni et al., 2021) using
129 the OrthoDB Tetrapoda ortholog library (v. odb10) (Kriventseva et al., 2018) (N=5310 orthologs).
130 Percentage of the genome assembled to chromosomes was calculated with a custom bash script.

131

132 ***Phylogenetic tree***

133 A species to family correspondence table was obtained from Jetz and Pyron (2018)
134 (https://vertlife.org/files_20170703/) and was filtered to include only the species with the longest
135 nucleotide sequence per family. This taxa subset was used to obtain a subset of 100 phylogenetic
136 trees from the posterior distribution of the Jetz and Pyron (2018) dataset, as available
137 from http://vertlife.org/phylosubsets. A consensus tree from these 100 trees was then obtained
138 using treeannotator (Drummond & Rambaut, 2007). The species names of the tree tips were then
139 substituted with the corresponding family names using the "sub.taxa.label" function in the
140 phylotools package (https://github.com/helixcn/phylotools) in R with the aid of the species to family
141 correspondence table, which was updated with the most recent classification available in
142 AmphibiaWeb https://amphibiaweb.org and the Amphibian Species of the World
143 database https://amphibiansoftheworld.amnh.org/.

144

145 ***Repeat modelling and annotation***

146 Repeats were *de novo* modelled with RepeatModeler (v. 2.0.3) (Flynn et al., 2020). Genomes were
147 then annotated using RepeatMasker (v. 4.1.2-p1) (Smit et al., 2013) with a concatenated library of
148 genome-specific repeats generated from RepeatModeler and the Dfam amphibian repeat library (v.
149 Dfam.h5) (Storer et al., 2021). Before annotation, any previous soft masking of the genomes was
150 reversed. The results were summarized using a custom bash script.

151

152 ***Ecological correlates of transposable element content***

153 On the basis of Global Biodiversity Information Facility (GBIF) data (https://www.gbif/org/; accessed
154 May 2023), MaxEnt habitat suitability-weighted means of a range of bioclimatic variables were
155 obtained for the 32 amphibian species. As previous studies have explored the relationship between
156 amphibian genome size and environmental variables (Liedtke et al., 2022), here we focused on the
157 relationship between bioclimatic variables and amphibian transposable elements. Influence of these
158 bioclimatic variables on transposable element content was modelled using Bayesian mixed effect
159 models (Hadfield, 2010), corrected for body size and phylogenetic signal (see Supplementary
160 Methods for further information).

161

162 ***Synteny analysis***

163 Synteny of BUSCO genes for chromosome level assemblies was analyzed with R Package GENESPACE
164 (v. 0.9.4) (Lovell et al., 2022), which uses OrthoFinder (v. 2.5.4) (Emms & Kelly, 2019) to infer
165 orthology. Synteny was analyzed using BUSCO "full_table.tsv" results files that were reformatted for
166 GENESPACE input using a custom bash script. Synteny plots were generated for all chromosome level
167 assemblies, all anuran chromosome level assemblies, and for the three caecilian genomes using the
168 GENESPACE plotting tool "plot_riparianHits". Chromosomes with reversed orientation compared to
169 the reference genome were inverted to improve visualisation.

170

171 ***Quantification and statistical analysis***

172 Regression analyses and Student's t-tests for comparing genome quality measurements were
173 conducted with the R statistics package (v. 4.1.2) (Team, 2013) in R Studio (v. 2022.02.3) (Team,
174 2022). Genome quality measures, contigN50, and scaffold count, were log transformed prior to

175 analysis. R-scripts for statistical analysis and plotting are available on GitHub at

176 https://doi.org/10.5281/zenodo.7679280.

177

178 **RESULTS**

179 ***Genome quality***

180 A total of 32 nuclear amphibian genome assemblies were available for our study and were generated

181 with a variety of sequencing technologies, including Illumina (NextSeq, HiSeq), PacBio (RS11, Sequel),

182 and Oxford Nanopore. Sequenced genomes represented 20 of 73 amphibian families with reference

183 genomes distributed unevenly across the phylogeny (Fig. 1). For example, there is only one

184 salamander genome representing the 798 extant species, no genomes representing the microhylids

185 or hyperoliids, yet there are five pipid and four pelobatid genomes (Fig. 1).

186

187 Genome assembly length ranged from 0.48 Gb in *Scaphiopus couchii* to 28.21 Gb in *Ambystoma*

188 *mexicanum* and was strongly positively associated with c-value estimates of genome size (F_{30} = 270.2,

189 $p < 1 \times 10^{-11}$) (Table 1, Fig. S2). Fifteen of these genomes were assembled to the chromosome level of

190 which the percentage of the genome assigned to chromosomes ranged from 63.88 to 99.96% (Table

191 1). Percentage of the genome assigned to chromosomes was positively associated with contig N50

192 (F_{13} = 10.6, p = 0.006), number of scaffolds (F_{13} = 21.35, p < 0.001) and read length ($t_{28.4}$ = 3.65, p =

193 0.001). There are additionally mitochondrial genome assemblies for 345 species of which 17 had

194 nuclear reference genomes. Eleven of the species with genomes had proteomics data (Table S1).

195

196 The quality of the amphibian genomes varied considerably (Table 1). Genomes generated with short-

197 read technologies were of lower quality than long-read or hybrid genome assemblies as indicated by

198 significantly lower contig N50s (t_{29} = 6.05, p < 0.00001), percentage of complete Benchmarking

199 Universal Single-Copy Ortholog (BUSCO) genes ($t_{13.4}$ = 2.85, p = 0.013), and higher scaffold numbers
200 (t_{29} = -4.86, p < 0.0001).
201
202 Contig N50 ranged from 362 bp in *S. couchii* to 22.45 Mb in *Xenopus laevis* with a median of 32.80
203 Kb. Scaffold count varied considerably from 54 in *X. laevis* to more than four million in *Bombina*
204 *variegata* with a median of 27.16 Kb (Table 1). Benchmarking Universal Single-Copy Orthologs
205 (BUSCO) scores ranged from 0.7 to 96% completeness (Tables 1, S1; Fig. 2) and were positively
206 associated with contig N50 (F_{29} = 55.09, p < 0.0000001; Fig. S3) and scaffold count (F_{29} = 42.98, p <
207 0.000001). All genomes had low percentages of duplicate BUSCO genes (< 6%), suggesting they may
208 be diploid except for the known tetraploid species, *X. laevis* (Fig. 2).
209
210 The variation we report here in genome quality, contiguity, and completeness may impact the value
211 of the genomes for functional genomics research. However, the improvements in all these measures
212 seen with the utilization of long read technologies or hybrid assemblies suggests that genome quality
213 will continue to improve as these approaches are used more frequently.
214
215 ***Repeat content***
216 Overall identified repeat percentage of the genomes ranged from 23% in *Platyplectrum ornatum* to
217 82% in *Oophaga sylvatica* and was positively associated with genome size (F_{30} = 5.87, p = 0.022)
218 (Tables 1; S3). Repeat content varied across genomes with the anuran *O. sylvatica* and the
219 salamander *Ambystoma mexicanum* dominated by Long Terminal Repeats (LTRs), the three caecilians
220 dominated by Long Interspersed Nuclear Elements (LINEs), and the two bufonid anurans dominated
221 by DNA transposons (Fig. 3; Tables S2-S4). The *Ambystoma mexicanum* genome had fewer repeats
222 than might be predicted by its large size (Fig. 3).
223

224 The proportion of repeats that could be classified by RepeatMasker ranged from 6.6% in *P. ornatum*
225 to 48.2% in *O. sylvatica* (Table S4) and was positively associated with genome quality measures
226 contigN50 (F_{29} = 12.8, p = 0.001), scaffold count (F_{29} = 7.80, p = 0.009), and percent BUSCO complete
227 (F_{29} = 4.60, p = 0.041). The ability to classify repeats was also positively associated with read length,
228 with longer reads resulting in better classification ($t_{26.241}$ = 3.57, p = 0.001).
229
230 These disparities in repeat percentage and content likely reflect differing evolutionary histories
231 among species, as indicated by three of the four congeneric species pairs in our dataset having
232 similar values (i.e., *Bufo*, *Leptobrachium,* and *Xenopus*; but not *Oophaga*). The differences we
233 observed in *Oophaga pumilio* and *O. sylvatica* are likely due to assembly quality rather than genome
234 content given that these two genomes were sequenced with different technologies and have
235 dramatically different genome qualities (e.g., contig N50s of 5.8 vs. 97.8 Kbp respectively).
236
237 ***Ecological correlates of transposable element content***
238 A Bayesian mixed effect modelling approach was employed to examine the relationships
239 between proportion of transposable elements and environmental variables. Controlling for
240 phylogenetic relationships (by estimating Pagel's lambda, λ) and body size (Spearman
241 correlation with TRANSPOSABLE ELEMENTcontent ρ= -0.772, p<0.001) and excluding the
242 three globally invasive species (*Rhinella marina, X. laevis*, and *Lithobates catesbeianus*) our
243 analysis revealed a significant influence (pMCMC = 0.004) of Bio18 (precipitation in the
244 warmest quarter) on proportion of total transposable elements (Figs 4, S6; Table S7).
245 Inclusion of these three invasives for a total of 32 species showed additional significant
246 relationships with Bio15 (precipitation seasonality) and elevation (Table S6). Further analysis
247 indicated that the relationship with Bio18 was not specific to a particular class of

248 transposable elements, such as retroelements or DNA transposons (Tables S7 and S8).
249 Phylogenetic signal (Pagel's lambda, λ) was moderate when considering total transposable
250 elements and retroelements (0.440; Tables S6 and S7) and increased when we considered
251 DNA transposons alone (0.603; Table S8).
252

253 ***Genome synteny***

254 Genome synteny of BUSCO genes was highly conserved within anurans (Fig. 5) and within caecilians
255 (Fig. S8) but was less conserved across the amphibian orders (Fig. S7). However, chromosome
256 naming was inconsistent across all taxa (Fig. 5, S10-11). For example, *X. tropicalis* chr1 is chr12 in
257 *Leptobrachium ailaonicum* (but not *L. leishanense*) and chr2 in *Bufo bufo* (but not *Bufo gargarizans*)
258 (Fig. 5). Orientation of chromosomes was also inconsistent, including between species of the same
259 genus (e.g., *Bufo*, *Leptobrachium*) (Fig. 5) and among the three caecilians (Fig. S8). Multiple
260 inversions were evident including between chr3 of pipids *(Xenopus tropicalis* and *Hymenochirus*
261 *boettgeri*) and other anurans (chromosomes 1, 2, 3, 4, or 10), caecilians (chr3 and chr4/5/6), and
262 even within species of the same genus (chr7 *Bufo gargarizans*, chr 9 *B. bufo*; Figs 5, S10-11). There
263 was also evidence of several chromosomal fissions including the separation of chr1 of *Leptobrachium*
264 *leishanense* into chr3 and chr6 in *Pyxicephalus adspersus* and into chr3 and chr7 in *Engystomops*
265 *pustulosus;* however, this chromosome remained mostly intact in the other anuran genomes (Fig. 5).
266 Given that these genomes were generated using different technologies, it is unclear whether or not
267 the differences we observed in synteny are due to true differences or assembly errors.

268 **DISCUSSION**

269 In this study we analysed 32 amphibian reference genomes from the public domain to evaluate their
270 content and usefulness for functional genetics research (Fig. 1, Table 1). There are considerably
271 fewer reference genomes for amphibians than exist for birds (N=754), mammals (N=406), and non-

272 avian reptiles (N=108). This scarcity of reference genomes results in many gaps in genome
273 representation across the amphibian tree of life including many entirely unrepresented groups and
274 with only one genome representing the entire order Caudata. The unrepresented families include
275 many of interest from a conservation perspective due to their high number of IUCN RedList Critically
276 Endangered species (e.g., Cryptobranchidae, Plethodontidae, Strabomantidae, and Craugastoridae)
277 (IUCN, 2022). However, our search of the Genomes on a Tree (GoaT) database (Sotero-Caio et al.,
278 2021) indicated that there are a further 24 amphibian genome assemblies in progress (20 anurans, 4
279 caudatans) indicating that this resource will be increasing by nearly 70% in the next few years.
280
281 The quality and completeness of the genomes in our dataset varied considerably (e.g., Fig. 2). Much
282 of this variation can be attributed to the sequencing technology used to generate them, with short-
283 read sequencing approaches resulting in lower completeness and continuity (Fig. S3). These impacts
284 are a recognized limitation of short-read sequencing and have been reported to impact genome
285 quality in taxa from insects (Hotaling et al., 2021b) to other vertebrates (Rhie et al., 2021), but have
286 likely had a disproportionate impact in amphibian genomes due to the difficulty of assembling
287 genomes with high repeat content (Sun et al., 2020). Fortunately, most ongoing sequencing efforts
288 now use long-read or hybrid sequencing approaches, which along with improved sequencing
289 algorithms, should result in higher quality amphibian genomes (Hotaling et al., 2021a; Lawniczak et
290 al., 2022; Rhie et al., 2021).
291
292 Genome quality (i.e., high continuity, contiguity, accuracy, completeness (Rhie et al., 2021)) is critical
293 for applications such as quantitative genetics where assembly errors can lead to incorrect inferences
294 in genetic association or genetic prediction. Quality also enhances the usefulness of genomes. For
295 example, highly contiguous chromosome-level assemblies decrease computational requirements for
296 downstream analyses such as mapping, variant calling, and alignment (Aganezov et al., 2022).

297

298 One of the most intriguing features of amphibian genomes is the huge range they exhibit in genome
299 size (Biscotti et al., 2019). This was exemplified in our dataset where assembly length ranged from
300 0.48 Gb in *Scaphiopus couchii* to 28 Gb in *Ambystoma mexicanum*. Why gigantic genomes exist in
301 some species, but not others, remains a key evolutionary question (Kapusta et al., 2017; Wang et al.,
302 2021). Explanations include differences in genome-level processes (e.g., insertion and deletion rates)
303 (Frahry et al., 2015; Sun et al., 2012), development (e.g., developmental rate and complexity)
304 (Gregory, 2002; Liedtke et al., 2018), physiology (e.g., water loss) (Johnson et al., 2021), body size
305 (e.g. miniaturization) (Decena-Segarra et al., 2020), and demography (e.g., effective population size)
306 (Liedtke et al., 2018; Lynch & Walsh, 2007) (but see Mohlhenrich & Mueller, 2016). As more
307 amphibian genomes become available, these hypotheses can be more rigorously evaluated.
308

309 We report some of the largest estimates of repeat content of any vertebrate (82% in *Oophaga*
310 *sylvatica* and 77% in *Bufo bufo*), exceeded only by the Australian lungfish at 90% (Meyer et al., 2021).
311 As expected, genome size was correlated with repeat content affirming that much of the variation in
312 amphibian genome size is due to an excess of repeats and transposable elements (Biscotti et al.,
313 2019; Lamichhaney et al., 2021).
314

315 In contrast to mammals, whose repeat landscape is mainly dominated by LTR retrotransposons (Platt
316 et al., 2018), amphibian repeat content varied considerably with some species dominated by DNA
317 transposons (as previously reported (Suda et al., 2022), and others by non-LTR retrotransposons
318 including the three caecilian genomes which were dominated by LINEs. This somewhat agrees with
319 genomic data and transcriptomic data from the caecilian *Ichthyophis bannanicus*, where LINEs were
320 the second most abundant type of repeat (26% of the genome) behind DIRS (30%) (Wang et al.,

321 2021); this is a higher percentage of LINES than we report in the three caecilian genomes in this study
322 (7 to 14%) (Table S4).
323
324 A considerable proportion of the repeats could not be classified. This was likely due to incorrect
325 classification (e.g., genes categorized as repeats) and the lack of good amphibian-specific repeat
326 resources (Ou et al., 2019) for classification via nucleotide sequence homology. The majority of
327 amphibian curated repeat libraries are generated in reference to *Xenopus* species (e.g., Dfam); the
328 large divergence times of this genus from the other amphibian species suggests that it may be a
329 contributing factor to the lack of classification. However, we also report many unclassified repeats in
330 the two *Xenopus* genomes.
331
332 The largest genome in our dataset, *A. mexicanum*, had fewer repeats than predicted given its size
333 (Fig S3) (Nowoshilow et al., 2018). This may be due, in part, to the Dfam library used for repeat
334 annotation being anuran-based; however, we did not observe this trend in the three caecilian
335 genomes in our dataset. Also, we performed *de novo* annotation of this genome, which should have
336 captured repetitive elements missing from Dfam. More likely, this low number of repeats reflects low
337 deletion rates and, thus, persistence of repeats in the genome for extremely long periods of time,
338 leading to their mutational decay into unique sequences whose repetitive origin is obscured (Keinath
339 et al., 2015; Novák et al., 2020).
340
341 We also show that amphibian species that inhabit wet and warm climates particularly during the
342 summer months have a greater proportion of transposable elements. This observed trend does not
343 appear to be driven by a specific group of transposons suggesting it may be caused by climatic
344 factors. Recent studies indicate that transposable elements exhibit greater activity in hotter climates
345 (Baduel et al., 2021) with an increasing number of studies suggesting increased transposable element

activity contributes to genetic diversification and facilitates species adaptation (Li et al., 2018; Schrader & Schmitz, 2019; Stapley et al., 2015). The pattern observed here likewise suggests the potential for heightened transposable element activity, and may help explain transposable element accumulation and potentially the higher evolutionary rates observed in the genomes of tropical amphibians.

Our study is the first to examine chromosomal synteny across all amphibian orders. We show that overall synteny of amphibian genomes is relatively conserved, particularly within orders (Figs 5 and S8). This aligns with previous results from anurans that reported conserved genome organization in this group (Bredeson et al., 2021; Wu et al., 2022). However, chromosome content and number varied across species, which seems to have been driven by multiple occurrences of chromosomal fusions and fissions (e.g., Fig. 5). Chromosomal rearrangements have occurred throughout vertebrate evolution, including the hypothesized fusion of microchromosomes in the ancestor of tetrapods to create the larger macrochromosomes seen in amphibians and mammals and their subsequent fission to create the microchromosomes of modern birds and non-avian reptiles (Waters et al., 2021).

Some of the structural rearrangements we detected may be due to assembly errors and should be evaluated in future assemblies using long-read scaffolding approaches (e.g., Oxford nanopore sequencing), chromosome conformation capture technologies (e.g., Hi-C), or chromosome mapping approaches (e.g., FISH). We also identified incongruities with chromosome naming and orientation caused by differences in assembly methods. These were apparent even within species of the same genus (e.g., *Bufo*). We suggest potential revisions of existing genome annotations to improve congruity and that future assemblies are curated consistently against high-quality reference genomes (e.g., *Xenopus laevis*).

371 **Conclusions**

372 New sequencing technologies and assembly algorithms have resulted in enough genomes for
373 comparative analyses spanning the amphibian phylogeny. This has already begun to yield important
374 insights on the evolution (Lamichhaney et al., 2021; Wu et al., 2022), development (Schloissnig et al.,
375 2021; Stuckert et al., 2021), sex determination (Hime et al., 2019; Ma & Veltsos, 2021), and unique
376 features (Fischer et al., 2019; Nowoshilow et al., 2018; Seidl et al., 2019) of this interesting group of
377 animals.

378

379 The increased availability of amphibian genomes can also aid conservation efforts in this highly
380 threatened group by facilitating research on genome-wide functional diversity, which can be used to
381 inform management decisions such as genetic rescue or targeted genetic intervention for species
382 threatened by habitat loss or chytridiomycosis (Chestnut et al., 2014; Kosch et al., 2022).
383 Additionally, well-annotated genomes can be used to create eDNA assays for population monitoring
384 (Breton et al., 2022; Saeed et al., 2022).

385

386 Future research efforts should focus on generating more reference genomes to fill the gaps in the
387 amphibian phylogeny and the identification of advantageous genetic traits against threats. Efforts
388 should also be made to increase the quality of genomes and expand transcriptome and annotation
389 databases. We suggest that these efforts strive to follow the recommendations of initiatives such as
390 the Earth BioGenome Project (Lawniczak et al., 2022), the Darwin Tree of Life Project (Blaxter et al.,
391 2022), and the Threatened Species Initiative (Hogg et al., 2022) to sequence at least one
392 representative from each family to ensure taxonomic coverage. Species selection should prioritize
393 species of interest for understanding valuable functional genetics traits; for example, for the purpose
394 of immunological research to understand disease resistance, or for conservation purposes to
395 enhance fitness.

ACKNOWLEDGEMENTS

This research was supported by The University of Melbourne's Research Computing Services and the Petascale Campus Initiative, and Australian Research Council Grants (FT190100462, LP200301370). We are grateful to A. Stuckert, N. Brajuka and J. Sproul for their advice on repeat modelling, and J.T. Li for providing suggestions on synteny analyses. We thank M. Moore for exploratory repeat analyses.

REFERENCES

Aganezov, S., Yan, S. M., Soto, D. C., Kirsche, M., Zarate, S., Avdeyev, P., Taylor, D. J., Shafin, K., Shumate, A., Xiao, C., Wagner, J., McDaniel, J., Olson, N. D., Sauria, M. E. G., Vollger, M. R., Rhie, A., Meredith, M., Martin, S., Lee, J., Koren, S., Rosenfeld, J. A., Paten, B., Layer, R., Chin, C. S., Sedlazeck, F. J., Hansen, N. F., Miller, D. E., Phillippy, A. M., Miga, K. H., McCoy, R. C., Dennis, M. Y., Zook, J. M., & Schatz, M. C. (2022). A complete reference genome improves analysis of human genetic variation. *Science, 376*(6588), eabl3533. doi:10.1126/science.abl3533

Baduel, P., Leduque, B., Ignace, A., Gy, I., Gil Jr, J., Loudet, O., Colot, V., & Quadrana, L. (2021). Genetic and environmental modulation of transposition shapes the evolutionary potential of Arabidopsis thaliana. *Genome Biology, 22*(1), 138.

Biscotti, M. A., Carducci, F., Olmo, E., & Canapa, A. (2019). Vertebrate Genome Size and the Impact of Transposable Elements in Genome Evolution. In P. Pontarotti (Ed.), *Evolution, Origin of Life, Concepts and Methods* (pp. 233-251). Cham: Springer International Publishing.

Blaxter, M., Archibald, J. M., Childers, A. K., Coddington, J. A., Crandall, K. A., Di Palma, F., Durbin, R., Edwards, S. V., Graves, J. A. M., Hackett, K. J., Hall, N., Jarvis, E. D., Johnson, R. N., Karlsson, E. K., Kress, W. J., Kuraku, S., Lawniczak, M. K. N., Lindblad-Toh, K., Lopez, J. V., Moran, N. A., Robinson, G. E., Ryder, O. A., Shapiro, B., Soltis, P. S., Warnow, T., Zhang, G., & Lewin, H. A. (2022). Why sequence all eukaryotes? *Proceedings of the National Academy of Sciences, 119*(4), e2115636118. doi:10.1073/pnas.2115636118

Bredeson, J. V., Mudd, A. B., Medina-Ruiz, S., Mitros, T., Smith, O. K., Miller, K. E., Lyons, J. B., Batra, S. S., Park, J., Berkoff, K. C., Plott, C., Grimwood, J., Schmutz, J., Aguirre-Figueroa, G., Khokha, M. K., Lane, M., Philipp, I., Laslo, M., Hanken, J., Kerdivel, G., Buisine, N., Sachs, L. M., Buchholz, D. R., Kwon, T., Smith-Parker, H., Gridi-Papp, M., Ryan, M. J., Denton, R. D., Malone, J. H., Wallingford, J. B., Straight, A. F., Heald, R., Hockemeyer, D., Harland, R. M., & Rokhsar, D. S. (2021). Conserved chromatin and repetitive patterns reveal slow genome evolution in frogs. *bioRxiv*, 2021.2010.2018.464293. doi:10.1101/2021.10.18.464293

Breton, B.-A. A., Beaty, L., Bennett, A. M., Kyle, C. J., Lesbarrères, D., Vilaça, S. T., Wikston, M. J. H., Wilson, C. C., & Murray, D. L. (2022). Testing the effectiveness of environmental DNA (eDNA) to quantify larval amphibian abundance. *Environmental DNA, 4*(6), 1229-1240. doi:10.1002/edn3.332

Brown, J. L., Morales, V., & Summers, K. (2010). A key ecological trait drove the evolution of biparental care and monogamy in an amphibian. *The American Naturalist, 175*(4), 436-446.

Chestnut, T., Anderson, C., Popa, R., Blaustein, A. R., Voytek, M., Olson, D. H., & Kirshtein, J. (2014). Heterogeneous Occupancy and Density Estimates of the Pathogenic Fungus *Batrachochytrium dendrobatidis* in Waters of North America. *PLOS ONE, 9*(9), e106790. doi:10.1371/journal.pone.0106790

Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., Davidson, C., Dodiya, K., Donaldson, S., El Houdaigui, B., El Naboulsi, T., Fatima, R., Giron, C. G., Genez, T., Martinez, J. G., Guijarro-Clarke, C., Gymer, A., Hardy, M., Hollis, Z., Hourlier, T., Hunt, T., Juettemann, T., Kaikala, V., Kay, M., Lavidas, I., Le, T., Lemos, D., Marugán, J. C., Mohanan, S., Mushtaq, A., Naven, M., Ogeh, D. N., Parker, A., Parton, A., Perry, M., Piližota, I., Prosovetskaia, I., Sakthivel, M. P., Salam, A. I. A., Schmitt, B. M., Schuilenburg, H., Sheppard, D., Pérez-Silva, J. G., Stark, W., Steed, E., Sutinen, K., Sukumaran, R., Sumathipala, D., Suner, M. M., Szpak, M., Thormann, A., Tricomi, F. F., Urbina-Gómez, D., Veidenberg, A., Walsh, T. A., Walts, B., Willhoft, N., Winterbottom, A., Wass, E., Chakiachvili, M., Flint, B., Frankish, A., Giorgetti, S., Haggerty, L., Hunt, S. E., GR, I. I., Loveland, J. E., Martin, F. J., Moore, B., Mudge, J. M., Muffato, M., Perry, E., Ruffier, M., Tate, J., Thybert, D., Trevanion, S. J., Dyer, S., Harrison, P. W., Howe, K. L., Yates, A. D., Zerbino, D. R., & Flicek, P. (2022). Ensembl 2022. *Nucleic Acids Research, 50*(D1), D988-d995. doi:10.1093/nar/gkab1049

Daly, J. W., Garraffo, H. M., Spande, T. F., Decker, M. W., Sullivan, J. P., & Williams, M. (2000). Alkaloids from frog skin: the discovery of epibatidine and the potential for developing novel non-opioid analgesics. *Natural product reports, 17*(2), 131-135.

De Angelis, M., Casciaro, B., Genovese, A., Brancaccio, D., Marcocci, M., Novellino, E., Carotenuto, A., Palamara, A., Mangoni, M., & Nencioni, L. (2021). Temporin G, an amphibian antimicrobial peptide against influenza and parainfluenza respiratory viruses: insights into biological activity and mechanism of action. *The FASEB Journal, 35*(2), e21358.

Decena-Segarra, L. P., Bizjak-Mali, L., Kladnik, A., Sessions, S. K., & Rovito, S. M. (2020). Miniaturization, genome size, and biological size in a diverse clade of salamanders. *The American Naturalist, 196*(5), 634-648. doi:10.5061/dryad.ht76hdrcg

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology, 7*, 214. doi:10.1186/1471-2148-7-214

Duellman, W. E. (1999). *Patterns of distribution of amphibians: a global perspective*: JHU Press.

Elewa, A., Wang, H., Talavera-López, C., Joven, A., Brito, G., Kumar, A., Hameed, L. S., Penrad-Mobayed, M., Yao, Z., & Zamani, N. (2017). Reading and editing the *Pleurodeles waltl* genome reveals novel features of tetrapod regeneration. *Nature Communications, 8*(1), 1-9. doi:10.1038/s41467-017-01964-9

Emms, D. M., & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology, 20*(1), 1-14.

Fischer, E. K., Roland, A. B., Moskowitz, N. A., Tapia, E. E., Summers, K., Coloma, L. A., & O'Connell, L. A. (2019). The neural basis of tadpole transport in poison frogs. *Proceedings of the Royal Society B: Biological Sciences, 286*(1907), 20191084. doi:10.1098/rspb.2019.1084

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., & Smit, A. F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences, 117*(17), 9451-9457. doi:10.1073/pnas.1921046117

Frahry, M. B., Sun, C., Chong, R. A., & Mueller, R. L. (2015). Low levels of LTR retrotransposon deletion by ectopic recombination in the gigantic genomes of salamanders. *Journal of Molecular Evolution, 80*, 120-129. doi:10.1007/s00239-014-9663-7

Gregory, T. R. (2002). Genome size and developmental complexity. *Genetica, 115*(1), 131-146. doi:10.1023/A:1016032400147

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of statistical software, 33*, 1-22.

Hime, P. M., Briggler, J. T., Reece, J. S., & Weisrock, D. W. (2019). Genomic Data Reveal Conserved Female Heterogamety in Giant Salamanders with Gigantic Nuclear Genomes. *G3: Genes|Genomes|Genetics*, g3.400556.402019. doi:10.1534/g3.119.400556

Hogg, C. J., Ottewell, K., Latch, P., Rossetto, M., Biggs, J., Gilbert, A., Richmond, S., & Belov, K. (2022). Threatened Species Initiative: Empowering conservation action using genomic resources. *Proceedings of the National Academy of Sciences, 119*(4), e2115643118. doi:10.1073/pnas.2115643118

Hotaling, S., Kelley, J. L., & Frandsen, P. B. (2021a). Toward a genome sequence for every animal: Where are we now? *Proceedings of the National Academy of Sciences, 118*(52), e2109019118. doi:10.1073/pnas.2109019118

Hotaling, S., Sproul, J. S., Heckenhauer, J., Powell, A., Larracuente, A. M., Pauls, S. U., Kelley, J. L., & Frandsen, P. B. (2021b). Long-reads are revolutionizing 20 years of insect genome sequencing. *Genome Biology and Evolution*. doi:10.1093/gbe/evab138

IUCN. (2022). The IUCN Red List of Threatened Species. Version 2022-2. Accessed on [6 January 2023]. https://www.iucnredlist.org.

Jetz, W., & Pyron, R. A. (2018). The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. *Nature Ecology & Evolution, 2*(5), 850-858. doi:10.1038/s41559-018-0515-5

Johnson, B. B., Searle, J. B., & Sparks, J. P. (2021). Genome size influences adaptive plasticity of water loss, but not metabolic rates, in lungless salamanders. *Journal of Experimental Biology, 224*(8), jeb242196. doi:10.1242/jeb.242196

Kapusta, A., Suh, A., & Feschotte, C. (2017). Dynamics of genome size evolution in birds and mammals. *Proceedings of the National Academy of Sciences, 114*(8), E1460-E1469. doi:10.1073/pnas.1616702114

Keinath, M. C., Timoshevskiy, V. A., Timoshevskaya, N. Y., Tsonis, P. A., Voss, S. R., & Smith, J. J. (2015). Initial characterization of the large genome of the salamander *Ambystoma mexicanum* using shotgun and laser capture chromosome sequencing. *Scientific Reports, 5*(1), 16413. doi:10.1038/srep16413

Kosch, T. A., Waddle, A. W., Cooper, C. A., Zenger, K. R., Garrick, D. J., Berger, L., & Skerratt, L. F. (2022). Genetic approaches for increasing fitness in endangered species. *Trends in Ecology & Evolution, 37*(4), 332-345. doi:10.1016/j.tree.2021.12.003

Kriventseva, E. V., Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simão, F. A., & Zdobnov, E. M. (2018). OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Research, 47*(D1), D807-D811. doi:10.1093/nar/gky1053

Lamichhaney, S., Catullo, R., Keogh, J. S., Clulow, S., Edwards, S. V., & Ezaz, T. (2021). A bird-like genome from a frog: Mechanisms of genome size reduction in the ornate burrowing frog, *Platyplectrum ornatum*. *Proceedings of the National Academy of Sciences, 118*(11), e2011649118. doi:10.1073/pnas.2011649118

Lawniczak, M. K. N., Durbin, R., Flicek, P., Lindblad-Toh, K., Wei, X., Archibald, J. M., Baker, W. J., Belov, K., Blaxter, M. L., Marques Bonet, T., Childers, A. K., Coddington, J. A., Crandall, K. A., Crawford, A. J., Davey, R. P., Di Palma, F., Fang, Q., Haerty, W., Hall, N., Hoff, K. J., Howe, K., Jarvis, E. D., Johnson, W. E., Johnson, R. N., Kersey, P. J., Liu, X., Lopez, J. V., Myers, E. W., Pettersson, O. V., Phillippy, A. M., Poelchau, M. F., Pruitt, K. D., Rhie, A., Castilla-Rubio, J. C., Sahu, S. K., Salmon, N. A., Soltis, P. S., Swarbreck, D., Thibaud-Nissen, F., Wang, S., Wegrzyn, J. L., Zhang, G., Zhang, H., Lewin, H. A., & Richards, S. (2022). Standards recommendations for the Earth BioGenome Project. *Proceedings of the National Academy of Sciences, 119*(4), e2115639118. doi:10.1073/pnas.2115639118

Lee, B. T., Barber, G. P., Benet-Pagès, A., Casper, J., Clawson, H., Diekhans, M., Fischer, C., Gonzalez, J. N., Hinrichs, A. S., Lee, C. M., Muthuraman, P., Nassar, L. R., Nguy, B., Pereira, T., Perez, G., Raney, B. J., Rosenbloom, K. R., Schmelter, D., Speir, M. L., Wick, B. D., Zweig, A. S., Haussler, D., Kuhn, R. M., Haeussler, M., & Kent, W. J. (2022). The UCSC Genome Browser database: 2022 update. *Nucleic Acids Research, 50*(D1), D1115-d1122. doi:10.1093/nar/gkab959

Li, Z.-W., Hou, X.-H., Chen, J.-F., Xu, Y.-C., Wu, Q., González, J., & Guo, Y.-L. (2018). Transposable elements contribute to the adaptation of *Arabidopsis thaliana*. *Genome Biology and Evolution, 10*(8), 2140-2150. doi:10.1093/gbe/evy171

Liedtke, H. C., Cruz, F., Gómez-Garrido, J., Fuentes Palacios, D., Marcet-Houben, M., Gut, M., Alioto, T., Gabaldón, T., & Gomez-Mestre, I. (2022). Chromosome-level assembly, annotation and phylome of Pelobates cultripes, the western spadefoot toad. *DNA Research, 29*(3), dsac013.

Liedtke, H. C., Gower, D. J., Wilkinson, M., & Gomez-Mestre, I. (2018). Macroevolutionary shift in the size of amphibian genomes and the role of life history and climate. *Nature Ecology & Evolution, 2*(11), 1792-1799. doi:10.1038/s41559-018-0674-4

Liu, Y., Shi, D., Wang, J., Chen, X., Zhou, M., Xi, X., Cheng, J., Ma, C., Chen, T., & Shaw, C. (2020). A novel amphibian antimicrobial peptide, phylloseptin-PV1, exhibits effective anti-staphylococcal activity without inducing either hepatic or renal toxicity in mice. *Frontiers in Microbiology, 11*, 565158.

Lovell, J. T., Sreedasyam, A., Schranz, M. E., Wilson, M., Carlson, J. W., Harkess, A., Emms, D., Goodstein, D. M., & Schmutz, J. (2022). GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. *eLife, 11*, e78526. doi:10.7554/eLife.78526

Lynch, M., & Walsh, B. (2007). *The origins of genome architecture*. Sunderland: Sinauer Associates.

Ma, W.-J., & Veltsos, P. (2021). The Diversity and Evolution of Sex Chromosomes in Frogs. *Genes, 12*(4), 483. doi:10.3390/genes12040483

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Molecular Biology and Evolution, 38*(10), 4647-4654. doi:10.1093/molbev/msab199

Meyer, A., Schloissnig, S., Franchini, P., Du, K., Woltering, J. M., Irisarri, I., Wong, W. Y., Nowoshilow, S., Kneitz, S., Kawaguchi, A., Fabrizius, A., Xiong, P., Dechaud, C., Spaink, H. P., Volff, J.-N., Simakov, O., Burmester, T., Tanaka, E. M., & Schartl, M. (2021). Giant lungfish genome elucidates the conquest of land by vertebrates. *Nature, 590*(7845), 284-289. doi:10.1038/s41586-021-03198-8

Mohlhenrich, E. R., & Mueller, R. L. (2016). Genetic drift and mutational hazard in the evolution of salamander genomic gigantism. *Evolution, 70*(12), 2865-2878.

Novák, P., Guignard, M. S., Neumann, P., Kelly, L. J., Mlinarec, J., Koblížková, A., Dodsworth, S., Kovařík, A., Pellicer, J., Wang, W., Macas, J., Leitch, I. J., & Leitch, A. R. (2020). Repeat-sequence turnover shifts fundamentally in species with large genomes. *Nature Plants, 6*(11), 1325-1329. doi:10.1038/s41477-020-00785-x

Nowoshilow, S., Schloissnig, S., Fei, J.-F., Dahl, A., Pang, A. W. C., Pippel, M., Winkler, S., Hastie, A. R., Young, G., Roscito, J. G., Falcon, F., Knapp, D., Powell, S., Cruz, A., Cao, H., Habermann, B., Hiller, M., Tanaka, E. M., & Myers, E. W. (2018). The axolotl genome and the evolution of key tissue formation regulators. *Nature*. doi:10.1038/nature25458
https://www.nature.com/articles/nature25458#supplementary-information

Nunes-de-Almeida, C. H. L., Haddad, C. F. B., & Toledo, L. F. (2021). A revised classification of the amphibian reproductive modes. *Salamandra, 57*(3), 413-427.

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., Farrell, C. M., Goldfarb, T., Gupta, T., Haft, D., Hatcher, E., Hlavina, W., Joardar, V. S., Kodali, V. K., Li, W., Maglott, D., Masterson, P., McGarvey, K. M., Murphy, M. R., O'Neill, K., Pujar, S., Rangwala, S. H., Rausch, D., Riddick, L. D., Schoch, C., Shkeda, A., Storz, S. S., Sun, H., Thibaud-Nissen, F., Tolstoy, I., Tully, R. E., Vatsan, A. R., Wallin, C., Webb, D., Wu, W., Landrum, M. J., Kimchi, A., Tatusova, T., DiCuccio, M., Kitts, P., Murphy, T. D., & Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research, 44*(D1), D733-745. doi:10.1093/nar/gkv1189

Ou, S., Su, W., Liao, Y., Chougule, K., Agda, J. R. A., Hellinga, A. J., Lugo, C. S. B., Elliott, T. A., Ware, D., Peterson, T., Jiang, N., Hirsch, C. N., & Hufford, M. B. (2019). Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biology, 20*(1), 275. doi:10.1186/s13059-019-1905-y

Paiola, M., Dimitrakopoulou, D., Pavelka, M. S., & Robert, J. (2023). Amphibians as a model to study the role of immune cell heterogeneity in host and mycobacterial

interactions. *Developmental & Comparative Immunology, 139*, 104594. doi:https://doi.org/10.1016/j.dci.2022.104594

Platt, R. N., 2nd, Vandewege, M. W., & Ray, D. A. (2018). Mammalian transposable elements and their impacts on genome evolution. *Chromosome Res, 26*(1-2), 25-43. doi:10.1007/s10577-017-9570-z

Rhie, A., McCarthy, S. A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., Lee, C., Ko, B. J., Chaisson, M., Gedman, G. L., Cantin, L. J., Thibaud-Nissen, F., Haggerty, L., Bista, I., Smith, M., Haase, B., Mountcastle, J., Winkler, S., Paez, S., Howard, J., Vernes, S. C., Lama, T. M., Grutzner, F., Warren, W. C., Balakrishnan, C. N., Burt, D., George, J. M., Biegler, M. T., Iorns, D., Digby, A., Eason, D., Robertson, B., Edwards, T., Wilkinson, M., Turner, G., Meyer, A., Kautt, A. F., Franchini, P., Detrich, H. W., Svardal, H., Wagner, M., Naylor, G. J. P., Pippel, M., Malinsky, M., Mooney, M., Simbirsky, M., Hannigan, B. T., Pesout, T., Houck, M., Misuraca, A., Kingan, S. B., Hall, R., Kronenberg, Z., Sović, I., Dunn, C., Ning, Z., Hastie, A., Lee, J., Selvaraj, S., Green, R. E., Putnam, N. H., Gut, I., Ghurye, J., Garrison, E., Sims, Y., Collins, J., Pelan, S., Torrance, J., Tracey, A., Wood, J., Dagnew, R. E., Guan, D., London, S. E., Clayton, D. F., Mello, C. V., Friedrich, S. R., Lovell, P. V., Osipova, E., Al-Ajli, F. O., Secomandi, S., Kim, H., Theofanopoulou, C., Hiller, M., Zhou, Y., Harris, R. S., Makova, K. D., Medvedev, P., Hoffman, J., Masterson, P., Clark, K., Martin, F., Howe, K., Flicek, P., Walenz, B. P., Kwak, W., Clawson, H., Diekhans, M., Nassar, L., Paten, B., Kraus, R. H. S., Crawford, A. J., Gilbert, M. T. P., Zhang, G., Venkatesh, B., Murphy, R. W., Koepfli, K.-P., Shapiro, B., Johnson, W. E., Di Palma, F., Marques-Bonet, T., Teeling, E. C., Warnow, T., Graves, J. M., Ryder, O. A., Haussler, D., O'Brien, S. J., Korlach, J., Lewin, H. A., Howe, K., Myers, E. W., Durbin, R., Phillippy, A. M., & Jarvis, E. D. (2021). Towards complete and error-free genome assemblies of all vertebrate species. *Nature, 592*(7856), 737-746. doi:10.1038/s41586-021-03451-0

Robert, J. (2020). Experimental platform using the amphibian *Xenopus laevis* for research in fundamental and medical immunology. *Cold Spring Harbor Protocols, 2020*(7), pdb. top106625.

Saeed, M., Rais, M., Akram, A., Williams, M. R., Kellner, K. F., Hashsham, S. A., & Davis, D. R. (2022). Development and validation of an eDNA protocol for monitoring endemic Asian spiny frogs in the Himalayan region of Pakistan. *Scientific Reports, 12*(1), 5624. doi:10.1038/s41598-022-09084-1

Scheele, B. C., Hunter, D. A., Grogan, L. F., Berger, L., Kolby, J. E., McFadden, M. S., Marantelli, G., Skerratt, L. F., & Driscoll, D. A. (2014). Interventions for Reducing Extinction Risk in Chytridiomycosis-Threatened Amphibians. *Conservation Biology, 28*(5), 1195-1205. doi:10.1111/cobi.12322

Scheele, B. C., Pasmans, F., Skerratt, L. F., Berger, L., Martel, A., Beukema, W., Acevedo, A. A., Burrowes, P. A., Carvalho, T., Catenazzi, A., De la Riva, I., Fisher, M. C., Flechas, S. V., Foster, C. N., Frías-Álvarez, P., Garner, T. W. J., Gratwicke, B., Guayasamin, J. M., Hirschfeld, M., Kolby, J. E., Kosch, T. A., La Marca, E., Lindenmayer, D. B., Lips, K. R., Longo, A. V., Maneyro, R., McDonald, C. A., Mendelson, J., Palacios-Rodriguez, P., Parra-Olea, G., Richards-Zawacki, C. L., Rödel, M.-O., Rovito, S. M., Soto-Azat, C., Toledo, L. F., Voyles, J., Weldon, C., Whitfield, S. M., Wilkinson, M., Zamudio, K. R., &

Canessa, S. (2019). Amphibian fungal panzootic causes catastrophic and ongoing loss of biodiversity. *Science, 363*(6434), 1459-1463. doi:10.1126/science.aav0379

Schloissnig, S., Kawaguchi, A., Nowoshilow, S., Falcon, F., Otsuki, L., Tardivo, P., Timoshevskaya, N., Keinath, M. C., Smith, J. J., & Voss, S. R. (2021). The giant axolotl genome uncovers the evolution, scaling, and transcriptional control of complex gene loci. *Proceedings of the National Academy of Sciences, 118*(15).

Schrader, L., & Schmitz, J. (2019). The impact of transposable elements in adaptive evolution. *Molecular Ecology, 28*(6), 1537-1549.

Schulte, L. M., Ringler, E., Rojas, B., & Stynoski, J. L. (2020). Developments in Amphibian Parental Care Research: History, Present Advances, and Future Perspectives. *Herpetological Monographs, 34*(1), 71-97. doi:10.1655/HERPMONOGRAPHS-D-19-00002.1

Seidl, F., Levis, N. A., Schell, R., Pfennig, D. W., Pfennig, K. S., & Ehrenreich, I. M. (2019). Genome of *Spea multiplicata*, a Rapidly Developing, Phenotypically Plastic, and Desert-Adapted Spadefoot Toad. *G3: Genes|Genomes|Genetics, 9*(12), 3909-3919. doi:10.1534/g3.119.400705

Smit, A., Hubley, R., & Green, P. (2013). RepeatMasker Open-4.0. 2013-2015. Retrieved from <http://www.repeatmasker.org>

Sotero-Caio, C., Challis, R., Kumar, S., & Blaxter, M. (2021). Genomes on a Tree (GoaT): A centralized resource for eukaryotic genome sequencing initiatives. *Biodiversity Information Science and Standards*.

Stapley, J., Santure, A. W., & Dennis, S. R. (2015). Transposable elements as agents of rapid adaptation may explain the genetic paradox of invasive species. *Molecular Ecology, 24*(9), 2241-2252.

Storer, J., Hubley, R., Rosen, J., Wheeler, T. J., & Smit, A. F. (2021). The Dfam community resource of transposable element families, sequence models, and genome annotations. *Mobile DNA, 12*(1), 1-14.

Stuckert, A. M., Chouteau, M., McClure, M., LaPolice, T. M., Linderoth, T., Nielsen, R., Summers, K., & MacManes, M. D. (2021). The genomics of mimicry: gene expression throughout development provides insights into convergent and divergent phenotypes in a Müllerian mimicry system. *Molecular Ecology*.

Suda, K., Hayashi, S. R., Tamura, K., Takamatsu, N., & Ito, M. (2022). Activation of DNA Transposons and Evolution of piRNA Genes Through Interspecific Hybridization in *Xenopus* Frogs. *Frontiers in Genetics, 13*. doi:10.3389/fgene.2022.766424

Sun, C., Shepard, D. B., Chong, R. A., López Arriaza, J., Hall, K., Castoe, T. A., Feschotte, C., Pollock, D. D., & Mueller, R. L. (2012). LTR retrotransposons contribute to genomic gigantism in plethodontid salamanders. *Genome Biology and Evolution, 4*(2), 168-183.

Sun, Y.-B., Zhang, Y., & Wang, K. (2020). Perspectives on studying molecular adaptations of amphibians in the genomic era. *Zoological Research, 41*(4), 351.

Team, R. (2022). RStudio: integrated development for R. Boston, MA: RStudio. In: Inc.

Team, R. C. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *http://www. R-project. org/*.

Wang, J., Itgen, M. W., Wang, H., Gong, Y., Jiang, J., Li, J., Sun, C., Sessions, S. K., & Mueller, R. L. (2021). Gigantic Genomes Provide Empirical Tests of Transposable Element

Dynamics Models. *Genomics, Proteomics & Bioinformatics*. doi:https://doi.org/10.1016/j.gpb.2020.11.005

Waters, P. D., Patel, H. R., Ruiz-Herrera, A., Álvarez-González, L., Lister, N. C., Simakov, O., Ezaz, T., Kaur, P., Frere, C., Grützner, F., Georges, A., & Graves, J. A. M. (2021). Microchromosomes are building blocks of bird, reptile, and mammal chromosomes. *Proceedings of the National Academy of Sciences, 118*(45), e2112494118. doi:doi:10.1073/pnas.2112494118

Wu, W., Gao, Y. D., Jiang, D. C., Lei, J., Ren, J. L., Liao, W. B., Deng, C., Wang, Z., Hillis, D. M., Zhang, Y. P., & Li, J. T. (2022). Genomic adaptations for arboreal locomotion in Asian flying treefrogs. *Proc Natl Acad Sci U S A, 119*(13), e2116342119. doi:10.1073/pnas.2116342119

DATA ACCESSIBILITY AND BENEFIT-SHARING

Data Accessibility Statement

Genetic data:

All the genomes used in this study are available on the NCBI Genomes database (https://www.ncbi.nlm.nih.gov/genome/).

Code:

All original code has been deposited on GitHub and is publicly available at (https://doi.org/10.5281/zenodo.7679280).

AUTHOR CONTRIBUTIONS

Conceptualization, T.A.K., A.J.C., L.A.O., A.R., and K.C.W.V; methodology, T.A.K, N.D.Y, R.L.M, and A.R.; formal analysis, T.A.K., M.L.P.; investigation, T.A.K., N.D.Y, R.L.M., K.C.W.V, M.L.P.; resources, L.A.0. and A.R.; writing – original draft, T.A.K.; writing – review & editing, all authors; project administration, T.A.K; funding acquisition, T.A.K. and L.F.S.

732

733

734

735

736

737

738

739 **TABLES AND FIGURES**

740 **Table 1**. Genome quality measures.

Species	Family	N chrs	Assembly ID	N scaffolds	Genome size (Gb)	Gap %	Scaffold N50 (kb)	Contig N50 (kb)	% Chrs	% Repeats	Busco % complete
Ambystoma mexicanum	Ambystomatidae	14	GCA002915635.3	27157	28.21	0.013	1205706.7	217.996	94.77	66.26	66.06
Allobates femoralis	Aromobatidae	0	unpublished	6658	5.32	0.041	625512.0	879.1	0.00	74.02	81.37
Bombina variegata	Bombinatoridae	0	GCA905336975.1	4065339	4.68	0.564	2.0	1.9	0.00	56.99	13.03
Rhinella marina	Bufonidae	0	GCA900303285.1	31391	2.55	0	167.5	167.5	0.00	67.71	85.89
Bufo gargarizans	Bufonidae	11	GCA014858855.1	747	4.55	1.792	539822.1	1738.3	96.21	74.54	86.74
Bufo bufo	Bufonidae	11	GCA905171765.1	1307	5.04	0.196	635713.4	3959.8	98.92	77.47	90.11
Hyalinobatrachium fleischmanni	Centrolenidae	0	unpublished	227850	4.04	11.337	295403.6	10.5	0.00	61.2	66.27
Oophaga pumilio	Dendrobatidae	0	GCA009801035.1	125383	3.49	36.043	78.0	5.8	0.00	40.77	67.57
Oophaga sylvatica	Dendrobatidae	0	unpublished	15287	5.19	0.109	571702.6	97.8	0.00	81.8	53.47
Geotrypetes seraphini	Dermophiidae	19	GCA902459505.2	164	3.78	0.146	272612.2	20656.6	99.85	72.51	91.39
Nanorana parkeri	Dicroglossidae	0	GCA000935625.1	25187	2.05	3.704	1069.1	32.8	0.00	51.39	90.40
Eleutherodactylus coqui	Eleutherodactylidae	13	GCA019857665.1	105233	2.79	15.115	109467.1	10.9	63.88	53.83	79.11
Engystomops pustulosus	Leptodactylidae	11	GCA019512145.1	108304	2.56	0.38	172109.2	305.4	75.27	52.32	75.48
Platyplectrum ornatum	Limnodynastidae	0	GCA016617825.1	148035	1.07	25.246	27.3	4.7	0.00	23.4	43.45
Limnodynastes dumerilii	Limnodynastidae	0	GCA011038615.1	520896	2.38	10.162	286.0	10.6	0.00	46.25	79.10
Mantella aurantiaca	Mantellidae	0	unpublished	135214	3.41	14.489	324858.8	14.1	0.00	52.56	58.57
Leptobrachium ailaonicum	Megophryidae	13	GCA018994145.1	5370	3.54	0.149	412424.8	821.1	98.22	67.92	81.79
Leptobrachium leishanense	Megophryidae	13	GCA009667805.1	5302	3.55	0.009	394693.0	1946.3	93.34	70.63	89.55
Spea bombifrons	Pelobatidae	0	GCA009364475.1	1629896	0.77	0	0.5	0.5	0.00	33.73	3.71
Spea multiplicata	Pelobatidae	0	GCA009364415.1	49736	1.08	0.235	71.5	30.7	0.00	33.55	76.59
Pipa parva	Pipidae	0	GCA019650415.1	283592	1.37	0.596	7.2	6.2	0.00	43.01	26.72
Xenopus tropicalis	Pipidae	10	GCA000004195.4	166	1.45	0.196	153961.3	14634.3	99.86	49.09	94.44
Xenopus laevis	Pipidae	18	GCA017654675.1	54	2.74	0.061	155250.6	22451.5	99.97	45.21	95.95

Hymenochirus boettgeri	Pipidae	9	GCA019447015.1	25835	3.21	0.128	293320.9	783.8	82.53	52.72	87.38
Pyxicephalus adspersus	Pyxicephalidae	14	GCA004786255.1	5411	1.56	10.517	157522.4	30.4	98.54	32.46	87.93
Rana temporaria	Ranidae	13	GCA905171775.1	555	4.11	0.637	481763.2	6264.8	98.26	69.11	90.72
Lithobates catesbeianus	Ranidae	0	GCA002284835.2	1544635	6.25	11.584	39.4	6.0	0.00	62.17	44.05
Glandirana rugosa	Ranidae	0	GCA018402905.1	561726	7.63	10.864	747.6	11.0	0.00	62.38	79.62
Rhinatrema bivittatum	Rhinatrematidae	19	GCA901001135.2	1330	5.32	0.638	486875.1	3216.3	97.36	67.86	91.58
Scaphiopus couchii	Scaphiopodidae	0	GCA009364435.1	1378905	0.48	0	0.4	0.4	0.00	26.94	0.70
Scaphiopus holbrookii	Scaphiopodidae	0	GCA009364455.1	1509266	0.71	0	0.5	0.5	0.00	31.42	2.60
Microcaecilia unicolor	Siphonopidae	14	GCA901765095.2	1081	4.69	1.012	376147.1	3661.5	97.06	69.21	90.89

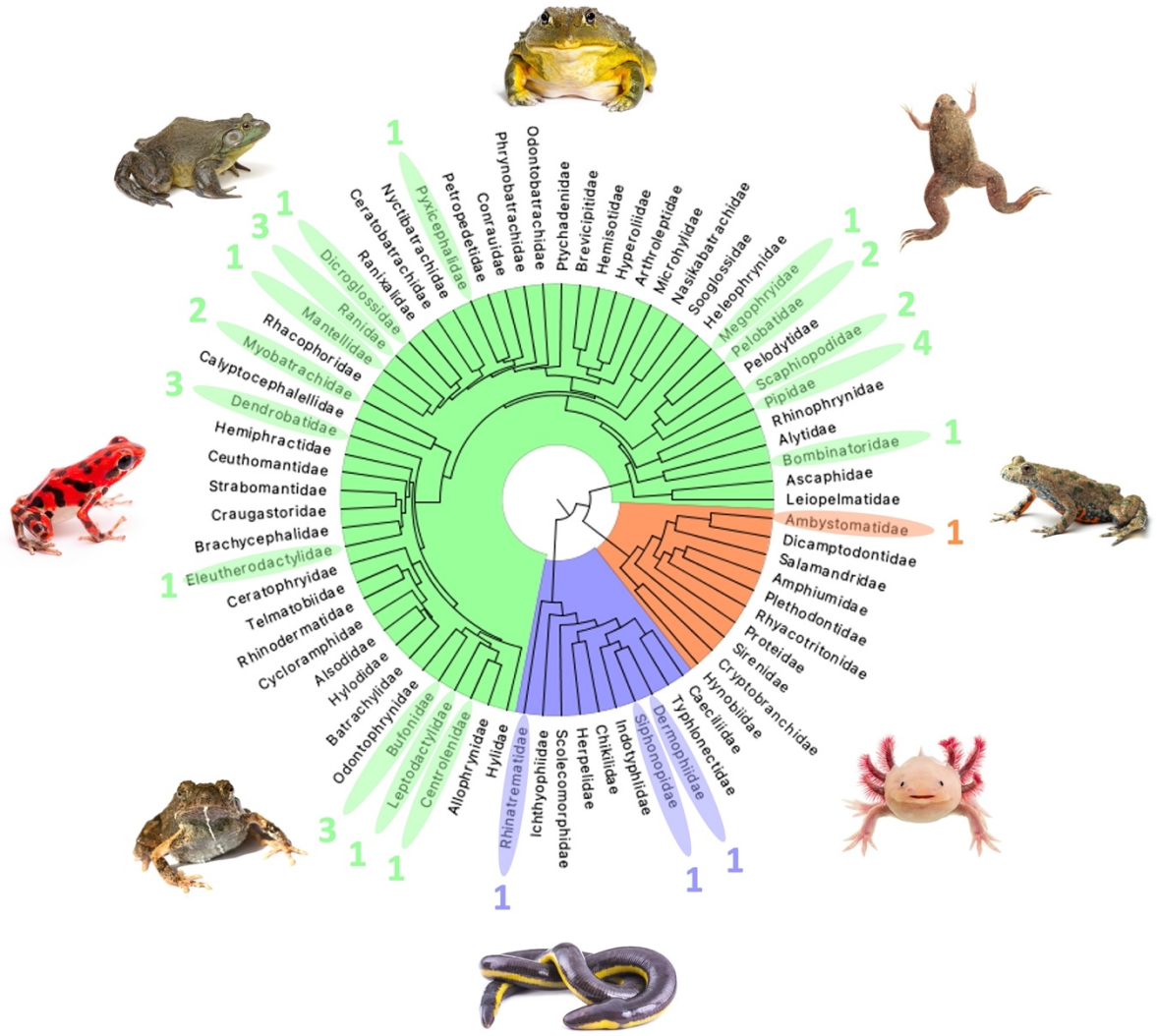

Figure 1. Phylogenetic tree of amphibian families. Amphibian families with representative genomes are highlighted and numbers indicate genome counts per family. (Green) anurans, (blue) caecilians, and (orange) salamanders. *Engystomops pustulosus* (Family) image was taken by B. Gratwicke, other amphibian images were licenced to T. Kosch by Adobe Stock and Shutterstock.

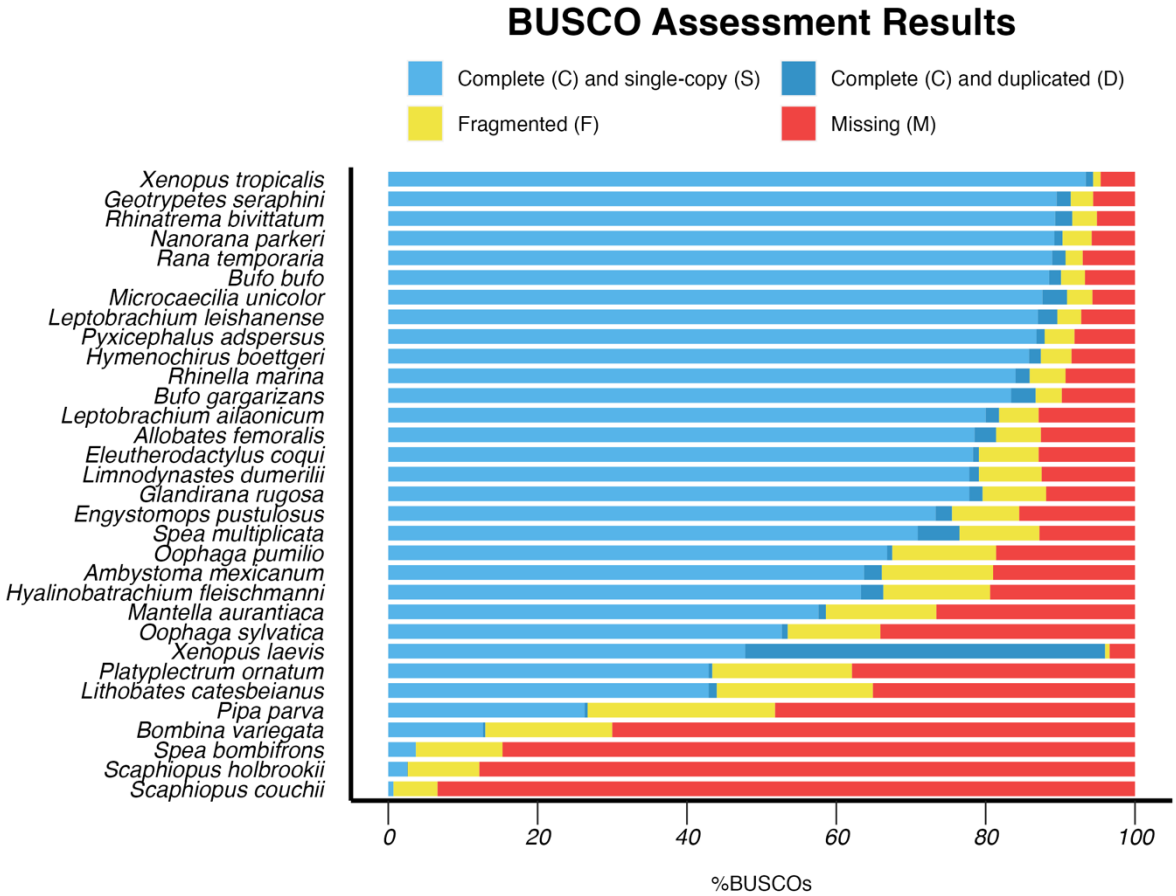

Figure 2. BUSCO (Benchmarking Universal Single-Copy Orthologs) assessment results for amphibian genomes.

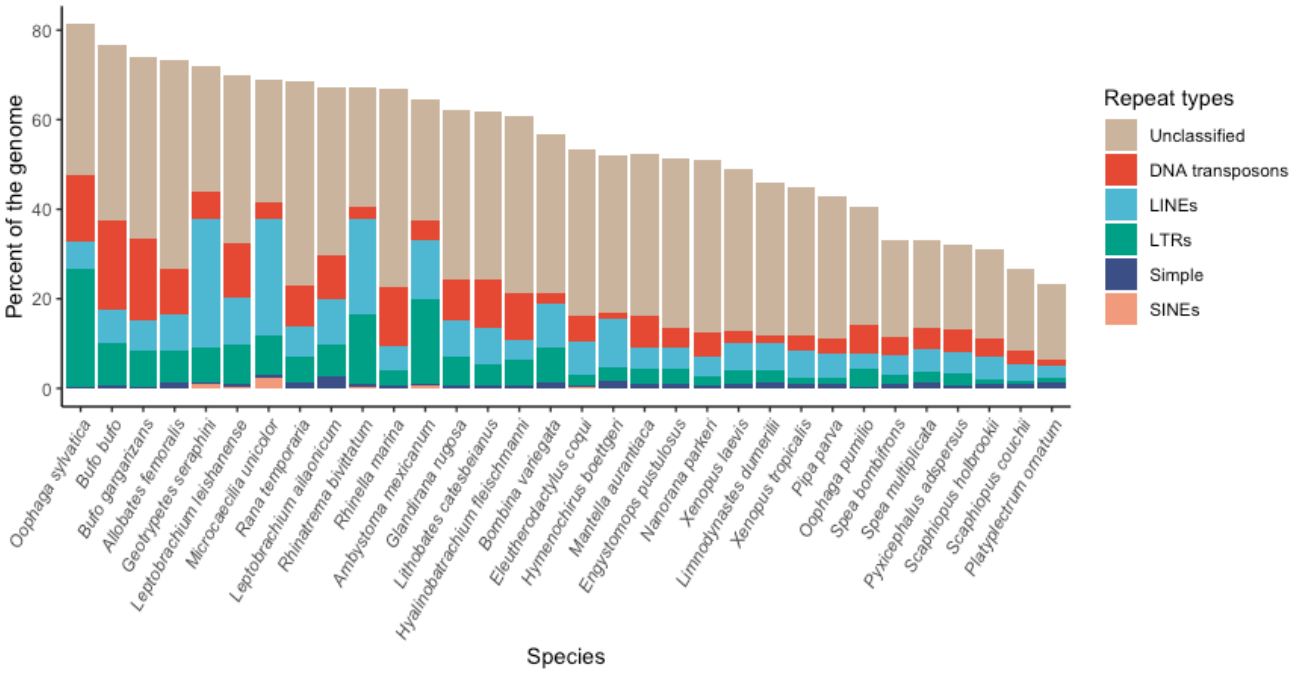

Figure 3. Repeat content across the amphibian genomes.

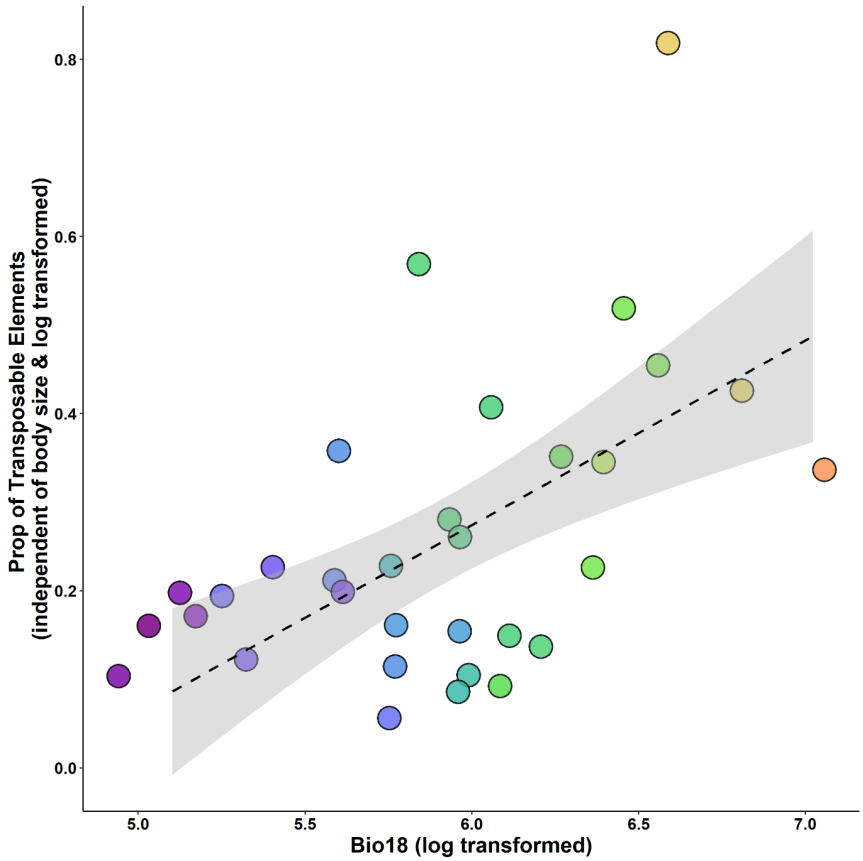

Figure 4: Relationship between proportion of total transposable elements per genome (body size corrected and log transformed) and Bio18 (precipitation of the warmest quarter, log transformed). 32 amphibian genomes have a higher proportion of transposable elements with increasing wet and warm conditions. Colours range from purple - low to medium- green to orange - high values.

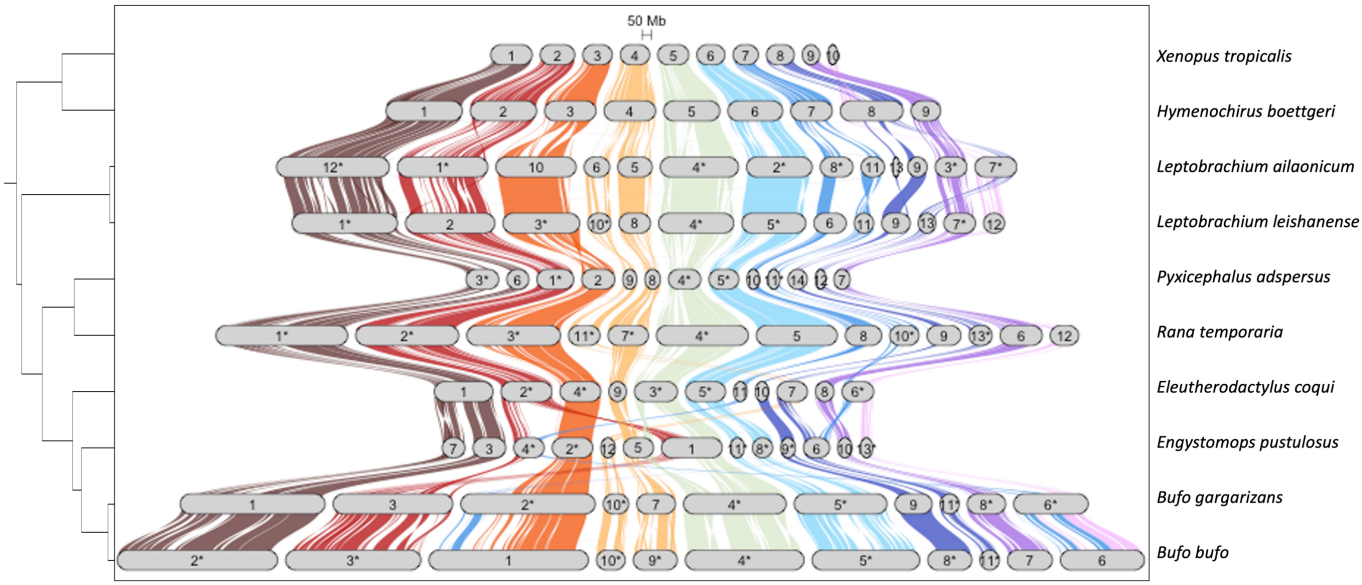

Figure 5. Synteny plot of BUSCOs (Benchmarking Universal Single-Copy Orthologs) for chromosome-level anuran genomes. The phylogenetic tree was created with Timetree.org. The reference genome is *Xenopus tropicalis*. Tetraploid anuran *Xenopus laevis* has been excluded. *Indicate inverted chromosomes. Chromosomes without BUSCOs were excluded from the plot.