Mitochondrial genome analyses uncover intriguing population structure across the invasive trajectory of an iconic invader

Running title: Intriguing population structure of an invader

Cheung K[1,2*], Amos TG[2,3], Shine R[4], Edwards RJ[2,5 #] & Rollins LA[1 #]

[1]Evolution & Ecology Research Centre, School of Biological, Earth & Environmental Sciences, University of New South Wales, Sydney, New South Wales, Australia

[2]School of Biotechnology & Biomolecular Sciences, University of New South Wales, Sydney, New South Wales, Australia

[3]Garvan Institute of Medical Research, Sydney, New South Wales, Australia

[4]Department of Biological Sciences, Macquarie University, Sydney, New South Wales, Australia

[5]Ocean Genomes Laboratory, Minderoo OceanOmics Centre, The University of Western Australia, Western Australia, Australia

Joint last authors

* Corresponding author: kelton.cheung@unsw.edu.au

Abstract

Invasive species offer insights into rapid adaptations in novel environments. The iconic cane toad (*Rhinella marina*) is an excellent model for studying rapid adaptation during invasion. Previous research using the mitochondrial NADH dehydrogenase 3 (*ND3*) gene in the Hawai'ian and Australian invasive populations found a single haplotype, indicating an extreme genetic bottleneck following introduction. Nuclear genetic diversity also exhibited reductions across the genome in these two populations. Here, we investigated the mitochondrial genome diversity of cane toads across this invasion trajectory. We created the first reference mitochondrial genome of the cane toad with long-read sequencing and constructed a phylogeny of Anura full mitochondrial genomes. We used transcriptomic data of 125 individuals from the native (French Guiana) and introduced (Hawai'i and Australia) ranges to construct nearly-complete mitochondrial genomes for population genomics analyses. As expected, the cane toad belongs to family Bufonidae, distinct from genus *Bufo*. In agreement with previous investigations of these populations, we identified genetic bottlenecks in both Hawai'ian and Australian introduced populations, alongside evidence of population expansion in the invasive ranges. Although mitochondrial genetic diversity in introduced populations was reduced, our results revealed that it had been underestimated: we identified 45 mitochondrial haplotypes in Hawai'ian and Australian samples, none of which were found in the native range. Additionally, we identified two distinct groups of haplotypes from the native range, separated by a minimum of 110 base pairs (0.6%). These findings enhance our understanding of Anura phylogenetics and how invasion has shaped the genetic landscape of this species.

Keywords

Rhinella marina, population genomics, mitochondrial genome, haplotypes, invasion biology

Introduction

Invasive species are a substantial global concern, particularly due to likely range expansions under climate change (Bonebrake et al., 2018; Seebens et al., 2017), an increasing rate of anthropologic introductions (Sm, Browne, Boudjelas, & De Poorter, 2000), and their negative impacts on native biodiversity and the economy (Bradshaw et al., 2016; Mainka & Howard, 2010). Invasive species often rapidly adapt to their new environments (Lee A. Rollins et al., 2013), creating an opportunity to study the process of rapid evolution. Biological invasions sometimes present a genetic paradox where, despite small founding population sizes that often reduce genetic diversity in the introduced range (i.e. genetic bottlenecks), they still maintain the ability to readily adapt to novel conditions (Allendorf & Lundquist, 2003; Estoup et al., 2016; Schrieber & Lachmuth, 2017).

Various genetic markers, including those from the nuclear and mitochondrial genomes, have been employed to investigate the introduction history and evolutionary processes of invasive species. Because of their distinct mode of inheritance and mutation rates, nuclear and mitochondrial markers may yield different insights into the evolutionary events underlying invasion. Genetic variation in invasive species is shaped by complex interplays among genetic drift, demography and selection. Small founding populations, common during an invasion, can promote genetic drift, leading to reduced genetic diversity. Furthermore, whilst genetic drift results in less efficient selection (Gravel, 2016), populations expanding their range into new environments may experience novel selection regimes and spatial sorting that further reduce genetic diversity.

Mitochondrial DNA (mtDNA) is more susceptible to genetic drift than nuclear markers due to its uniparental inheritance and small effective population size (Klucnika & Ma, 2019; Xiao, Nguyen, Wu, & Hao, 2017). Nonetheless, mtDNA has been extensively used to study the evolutionary history of different species or populations within the same species because of its non-recombining nature and purported near-neutrality. However, the neutrality of mtDNA has been challenged by evidence of selection, likely due to the important role of mitochondria in ATP production and cellular activities (Meiklejohn, Montooth, & Rand, 2007). Deleterious mutations are rapidly removed by purifying selection to maintain the function of the electron transport chain. Further, variation in mtDNA has been linked to adaptation to high latitudes (Chen et al., 2020; Wang et al., 2021) and temperature (Baker et al., 2019; Cheng, Liu, Yang, Sun, & Kong, 2013; Silva, Lima, Martel, & Castilho, 2014). There exists a substantial repository of mtDNA data in public databases, enabling comprehensive comparative genetic analyses, increasing its usefulness as a marker.

The cane toad (*Rhinella marina*) is one of the world's most infamous, and well-studied, invasive species (Sm et al., 2000). Cane toads are native to Central and South America (Zug, 1979) and have been introduced to multiple Caribbean islands for biological control of cane beetles, a parasite of sugarcane (Easteal, 1981). Cane toads were transported from the Caribbean to Hawai'i in 1932 for cane beetle control and, in 1935, 101 Hawai'ian cane toads were translocated from Oahu to Queensland, Australia (Turvey, 2013). However, they failed to control cane beetles in Australia and, facilitated by a high reproductive rate, subsequently spread across the continent to arid Western Australia. This species' inexorable spread throughout its introduced range in Australia has had catastrophic effects on the naïve

83 predators it has encountered (Chinchio et al., 2020). Freshwater crocodiles, snakes, lizards,
84 and quolls have been particularly badly affected by the cane toad invasion in Australia
85 (Shine, 2010). The colonization of the cane toad is still on-going; toads occupy over 1.2
86 million km^2 of Australia (Urban, Phillips, Skelly, & Shine, 2008) including tropical
87 Queensland, the Northern Territory, and the Kimberley region of northern Western
88 Australia (Radford et al., 2019).

89 Like in many other invasions, the cane toads' translocation resulted in genetic bottlenecks.
90 Australian cane toads have very low MHC class I and class II (M. Lillie, Cui, Shine, & Belov,
91 2016; Mette Lillie, Dubey, Shine, & Belov, 2017; M. Lillie, Shine, & Belov, 2014) and
92 microsatellite (Leblois, Rousset, Tikel, Moritz, & Estoup, 2000) diversity. Previous analysis of
93 the NADH dehydrogenase 3 (*ND3*) mitochondrial gene from 31 individuals sampled from
94 Hawai'i and Australia identified a single haplotype, a substantial reduction of haplotype
95 diversity compared to the native range (Slade & Moritz, 1998), further supporting the
96 presence of genetic bottlenecks in the invasive ranges. Despite this low genetic variation,
97 phenotypic traits differ substantially across the cane toad's Australian range including
98 immune function, dispersal ability, and behaviour (Alford, Brown, Schwarzkopf, Phillips, &
99 Shine, 2009; G. P. Brown, Phillips, Dubey, & Shine, 2015; Gruber, Brown, Whiting, & Shine,
100 2018; Gruber, Whiting, Brown, & Shine, 2017; Phillips, Brown, & Shine, 2010). Some of these
101 traits have been shown to be intergenerationally transmitted, consistent with genetic
102 causes. For example, toads from range-core and range-edge populations that were raised in
103 a common-garden experiment showed divergence of behavioural traits and morphology
104 (Gruber, Brown, Whiting, & Shine, 2017; Hudson, Brown, & Shine, 2016). The invasion

105 success of this species in Australia, coupled with evidence of evolutionary change since
106 introduction make it an ideal system for the study of rapid evolution during invasion.

107 Despite extensive research into toad evolution in the introduced range, we need improved
108 genetic resources to clarify the molecular mechanisms underlying these changes.
109 Transcriptome data from spleen, muscle and/or brain tissues of toads from French Guiana,
110 Hawai'i, and Australia, and liver tissue from Brazil exist (Richardson et al., 2018; L. A. Rollins,
111 Richardson, & Shine, 2015; Selechnik, Richardson, Shine, Brown, & Rollins, 2019; Yagound et
112 al., 2022) and a draft reference genome (Edwards et al., 2018) is available, but population
113 genomic analyses of cane toads across this invasion trajectory have thus far been restricted
114 to a single study of single nucleotide polymorphisms (SNPs) from the nuclear genome
115 (Selechnik, Richardson, Shine, DeVore, et al., 2019). Here, we present a complete cane toad
116 mitochondrial genome (mitogenome) and explore the species' phylogenetic relationships
117 with other Anurans. We then investigate evolution across the invasion trajectory using
118 mtDNA data derived from RNASeq and whole-genome resequencing. We compare our data
119 to those of a previously-published study of *ND3* on the same populations (Slade & Moritz,
120 1998), to determine whether our whole-mitochondrial genome analysis can extend
121 population-genetic inferences. We predicted that there would be decreasing genetic
122 diversity within the mitochondrial genome with increasing distance from the native range
123 due to genetic bottlenecks. We also predicted that we would identify more genetic diversity
124 within Australia than has been previously characterised, given the phenotypic evidence of
125 adaptation to this novel environment.

126 # Materials and Methods

127 ## Reference mitochondrial genome

128 High-molecular-weight genomic DNA was extracted from the liver of a female cane toad
129 collected in the Kimberley region of Western Australia and sequenced on Illumina HiSeq X
130 Ten and PacBio RS II platforms as described previously (Edwards et al., 2018). The
131 mitogenome was assembled from 36 SMRT cells generated from the first two SMRTBell
132 libraries (ENA experiments ERX2389178 and ERX2389179).

133 The 17,757 basepair (bp) *Bufo japonicus* complete mitochondrial genome (gi:157939553)
134 was downloaded from NCBI on 29/3/2016 and used to identify mitochondrial subreads from
135 the raw sequencing data. GABLAM (Davey, Shields, & Edwards, 2006) was used to map the
136 *B. japonicus* mtDNA onto PacBio subreads using blastn (blast+ v2.2.31) and extract 143
137 subreads that had 75%+ bidirectional coverage, *e.g.* at least 75% of the subread mapped
138 onto the mtDNA and at least 75% of the mtDNA was covered by the subread. Any subreads
139 that mapped the negative strand were reverse complemented. The subread with the best
140 global hit to the *B. japonicus* mtDNA was selected as the core of the cane toad
141 mitochondrial genome. A second GABLAM search was performed of the 143 subreads
142 against this new query, retaining 29 sequences that had >80% global sequence identity
143 within local matches in the same orientation and order, excluding any “wrapping” due to
144 circularity. These 29 sequences were aligned using MAFFT v2.273 (Katoh & Standley, 2013)
145 with a very small gap penalty to account for the high indel rate (--localpair --maxiterate 1000
146 --op 0.1 --ep 0.1 --lop -0.1 --nuc). SeqSuite (Edwards, Paulsen, Aguilar Gomez, & Perez-
147 Bercoff, 2020) was used to generate a consensus sequence from the alignment using the
148 most frequent non-gap base call at each position. Any column with fewer than 5 non-gap

149 sequences was excluded from the consensus. All 143 raw subreads were mapped onto the
150 consensus using BLASR and the sequence was polished using Quiver (SMRT Analysis v2.3.0).
151 GABLAM was used to identify the overlapping ends of the sequence, which was then
152 circularised by trimming off the first 1957 bp for a final length of 18,152 bp. A final polishing
153 step was performed using the Illumina data (ENA experiment ERX2845325), mapped with
154 Bowtie v2.3.3 (Langmead & Salzberg, 2012) and error-corrected with Pilon v1.20 (Walker et
155 al., 2014).

156 The mitochondrial genome was annotated using MITOS online (Bernt et al., 2013) with the
157 vertebrate mtDNA genetic code (Genetic Code Table 02). Annotated protein-coding genes
158 were manually curated and extended to include stop codons where required. Where no
159 canonical in-frame stop codon was found, it was assumed that a partial (T or TA) stop codon
160 would be completed upon polyadenylation. Codon usage was calculated using
161 https://www.bioinformatics.org/sms2/codon_usage.html (Accessed: 2023-03-29). The final
162 annotated mitochondrial genome was uploaded to NCBI (Accession number: NC_066225.1).
163 The mitogenome map was drawn using the CGView server
164 (http://stothard.afns.ualberta.ca/cgview_server/).

165 ## Phylogenetic analysis

166 Complete mitochondrial genomes of 203 Anura, and five non-Anura outgroup species, were
167 downloaded from NCBI BioProject database (Supplementary Table 2, Downloaded 2022-09-
168 11). Each mitochondrial genome was re-circularised and edited to end with the control
169 region for subsequent analyses. The multiple sequence alignment was conducted with the
170 cane toad reference mitogenome using MAFFT with auto mode to find the best alignment

171 algorithm (Katoh & Standley, 2013). The resulting alignment was visually inspected using
172 AliView (Larsson, 2014) without further trimming. Maximum likelihood phylogenetic
173 analysis was conducted using IQ-Tree multicore version 2.2.2.7 (default parameters,
174 ultrafast bootstrap (UFBoot) and SH-aLRT branch test with 10000 replicates) (Minh et al.,
175 2020). Clades are considered reliable if SH-aLRT value >= 80% and UFboot value >=95%. The
176 best fit model was found by ModelFinder embedded in IQ-Tree (Kalyaanamoorthy, Minh,
177 Wong, von Haeseler, & Jermiin, 2017). The phylogenetic tree was visualised and edited
178 using R package *ggtree* (Yu, Smith, Zhu, Guan, & Lam, 2017).

179 Sequence data for population- and species-level analyses

180 Previously published cane toad RNASeq data from different tissues were retrieved and
181 downloaded from NCBI BioProject PRJNA510261 and PRJNA395127 (spleen), PRJNA479937
182 (brain) and PRJNA277985 (muscle) for analysis. A total of 125 individuals with 144 tissue
183 samples were included, sampled in French Guiana (*n* = 24), Hawai'i (*n* = 8) and Australia (*n* =
184 93) (See Figure 1 and Supplementary table 1). For some individuals, RNASeq data from both
185 spleen and brain were available; in such cases, we used data from brains. Raw reads were
186 processed as described previously with STAR v2.7.2b (Dobin et al., 2013) and Genome
187 Analysis Toolkit (GATK) v4.1.9.0 (McKenna et al., 2010), using the extracted mitogenome
188 from the draft genome as a reference (Selechnik, Richardson, Shine, DeVore, et al., 2019).

189 Genomic DNA was extracted from fifteen cane toad samples, including two from toe clips,
190 two from spleen, four from liver and seven from muscle tissues using a Gentra PureGene
191 DNA extraction kit (QIAGEN) (See Supplementary Table 1).

192 For two samples (RMH006 and RM265), paired-end sequencing libraries were prepared
193 using the TruSeq DNA PCR-free kit (Illumina). These libraries were sequenced on the
194 Illumina NovaSeq 6000 platform at the Ramaciotti Centre for Genomics, University of New
195 South Wales, Sydney, Australia (UNSW Sydney). The remaining thirteen samples had paired-
196 end sequencing libraries constructed according to the manufacturer's specifications for the
197 NEBNext Ultra DNA Library Prep Kit for Illumina (NEB). Three samples (RMF031, RMF042
198 and RMF044) were sequenced on the 10X Genomics Chromium platform at the Ramaciotti
199 Centre for Genomics, UNSW Sydney. The remaining libraries were sequenced on the
200 Illumina NovaSeq 6000 platform at Deakin Genomic Centre, Victoria, Australia with a target
201 depth of 20-fold coverage per genome. Raw sequencing reads from each sample underwent
202 initial processing to remove adapter sequences. Trimmomatic v0.38 (Bolger, Lohse, &
203 Usadel, 2014) was employed using the following parameters: ILLUMINACLIP:TruSeq3-
204 PE.fa:2:30:10:4 SLIDINGWINDOW:5:30 AVGQUAL:30 MINLEN:36.

205 Mitochondrial segments, *ND3* gene with tRNA-Gly and tRNA-Arg flanking sequence (*ND3*
206 dataset hereafter) were obtained from previous studies (Acevedo, Lampo, & Cipriani, 2016;
207 Slade & Moritz, 1998) (Supplementary Table 1). We aligned our data with this *ND3* dataset
208 using the R package DECIPHER (Wright, 2016) and trimmed our data to match the existing
209 sequences.

210 Construction of mtDNA genome and variant calling

211 GATK HaplotypeCaller (Poplin et al., 2018) was used to call SNPs and indels using the
212 following criteria: soft clipped bases were not included and variants with minimum Phred-
213 scaled confidence of 20 threshold were kept. To filter low quality variants, GATK

214 VariantFiltration (Van der Auwera, 2020) was used to remove variants in the resulting VCF
215 file: clusters of 3 SNPs within a window of 35 bases, variants with Fisher Strand (FS) greater
216 than 30.0, QualbyDepth (QD) less than 2.0, depth of coverage (DP) less than 20.0 and allelic
217 frequency (AF) less than 0.05. Due to the difficulty of mapping short-read RNASeq data to
218 sequences with multiple repeats, and the risk of introducing false-positive variant calls, we
219 excluded variants from the repetitive sequences in the hypervariable region 2 (HV2) of the
220 mitochondrial control region. We determined the consensus sequence of each sample using
221 FastaAlternateReferenceMaker (McKenna et al., 2010).

222 Consensus mitogenome sequences from individuals with both RNASeq and WGS data were
223 aligned against each other using DECIPHER R packages (Wright, 2016, 2020) using default
224 settings. The alignments were further visualised and analysed with AliView (Larsson, 2014).

225 NUMT identification

226 The presence of NUMTs in the cane toad's genome has not been documented previously.
227 The cane toad draft genome was retrieved and downloaded from NCBI BioProject
228 PRJEB24695. We used RepeatModeler (http://www.repeatmasker.org/RepeatModeler/) to
229 construct species-specific repeat library and masked the genome using RepeatMasker
230 (Tarailo-Graovac & Chen, 2009). NUMTFinder v0.5.3 (Edwards et al., 2021) was used to
231 identify potential NUMTs by searching the full and repeat-masked versions of the nuclear
232 genome with our new mtDNA assembly.

233 Population genetics analyses

234 The spatial distribution of haplotypes was examined by creating a median-joining haplotype
235 network using Network and postprocessed using the maximum parsimony calculation to
236 remove unnecessary median vectors and links (Polzin & Vahdati Daneshmand, 2003). The

237 final network was drawn using Network Publisher v2.1.2.5 (Fluxus Engineering, Clare, UK).
238 This same procedure also was used to analyse the *ND3* dataset.

239 We then investigated population genetic diversity and differentiation. The following
240 diversity indices were calculated using DnaSP v6 (Rozas et al., 2017): nucleotide diversity (π),
241 number of polymorphic sites, number of parsimony informative sites, number of haplotypes
242 (*H*) and haplotype diversity (*h*). Due to the uneven sample size across populations,
243 haplotype richness (*HR*) was calculated using FSTAT v2.9.4 (Goudet, 1995) to measure the
244 number of alleles independent of the sample size. DnaSP was also used to generate a
245 haplotype list for Arlequin software v3.5.2.2 (Excoffier & Lischer, 2010) and a Roehl data file
246 for use with Network software 10.2.0.0 (Bandelt, Forster, & Rohl, 1999; Bandelt, Forster,
247 Sykes, & Richards, 1995). To study the population structure, Arlequin was used to calculate
248 genetic differentiation among and between populations by pairwise fixation indices (*Fst*)
249 and Analysis of Molecular Variance (AMOVA) to estimate the genetic variation among and
250 within populations.

251 To study the demographic history of each population, we used both neutrality statistics and
252 mismatch distribution tests to infer population expansion. Fu's *Fs* statistics was produced
253 using Arlequin and Ramos-Onsins and Rozas R_2 statistics were produced using DnaSP. Fu's *Fs*
254 test calculates the probability of observing a number of haplotype similar or smaller than
255 the observed number of samples (Fu, 1997). Both tests are based on the infinite-site model
256 without recombination, which is suitable for analysing mitochondrial genomic data. A
257 significant negative value of *Fs* indicates an excess of rare alleles and haplotypes due to a
258 sudden population growth after a population bottleneck or genetic hitchhiking while

259 positive values indicate a lack of rare mutations, resulting from balancing selection or
260 population stability. Ramos-Onsins and Rozas R_2 statistic compares the difference between
261 the number of singleton mutations and the average number of segregating sites (Ramos-
262 Onsins & Rozas, 2002). The R_2 statistic is known to have better power than Fu's *Fs* when the
263 sample size is small (< 10) or the number of segregating sites is low (Ramirez-Soriano,
264 Ramos-Onsins, Rozas, Calafell, & Navarro, 2008). A small positive value of R_2 indicates
265 population expansion. The statistical power of these tests was validated using 10,000
266 simulations. Mismatch distributions of the frequency of pairwise differences between
267 haplotype pairs in each population were also examined using the sudden expansion and
268 spatial expansion models implemented in Arlequin (Excoffier, 2004; Schneider & Excoffier,
269 1999). The sum of square deviations (SSD) was used to test the differences between
270 observed and expected values while the raggedness index was used to explore the
271 smoothness of the observed mismatch distribution. The statistical significance of these tests
272 was validated with 500 replicates in Arlequin. A non-significant SSD value and a small
273 raggedness index value with a unimodal distribution indicates no deviation of the observed
274 data from the expectation under the expansion model.

275 Results

276 Summary of the cane toad reference mitochondrial genome

277 The mitogenome of *R. marina* (Figure 2) consists of 18,152 bp, including 22 tRNA genes, 2
278 rRNA genes and 13 protein-coding genes, with 28 genes located on the forward (+) strand
279 and 9 located on reverse (-) strand (Supplementary Table 2). There is also a 2,753 bp GC-rich
280 control region containing the origin of replication. Two groups of tandem repeats were
281 observed in the latter part of the control region: 530 bps with 6-full repeats and 350 bps

282 with 3-full repeats. The gene order of the cane toad mitogenome was examined by aligning
283 with the members of its sister clades (*Bufo* and *Bufotes*) and we found no gene
284 rearrangements. The base composition of the mitogenome was A: 29%, T: 33%, G: 14% and
285 C: 24%, with a bias with 62% of A + T. The codon usages of those with nucleotide G in the
286 third codon position were lower than those with A and T. GCG (Ala) was the least used
287 codon and TGG was the only codon used for encoding Tryptophan. All protein-coding genes
288 used ATG as start codon.

289 Phylogenetics analysis

290 We produced a maximum likelihood phylogenetic tree using whole mitogenomes of Anura.
291 The best fit model was GTR+F+R9 according to lowest Bayesian information Criterion (BIC).
292 The location of *R. marina* in the phylogeny of Anura is shown in Figure 3. *Rhinella marina*
293 was grouped with *Anaxyrus americanus*, an American Bufonid and previously classified
294 under *Bufo* genus, with 100% statistical support in both UFBoot and SH-aLRT in branch tests.
295 Four species of *Bufotes* and four of *Bufo* were well resolved into two subclades. Even though
296 there was high branch support for differentiation between the genera *Rhinella* and
297 *Anaxyrus*, and high support for species within *Bufotes* and *Bufo*, the divergence among
298 genera was not significantly supported, which may indicate the signals in the whole
299 mitochondrial genomes were insufficient for resolving these clades or that wider taxonomic
300 coverage is required to resolve this phylogeny.

301 Verification of mtDNA genomes from RNASeq data

302 A near-complete mitochondrial genome with 17,261 to 17,270 bps (95.1% of reference
303 mitochondrial genome) for a total of 132 individuals was generated from the RNASeq and
304 WGS dataset. A sequence with length of 889 bases consisting of repeats in the HV2 section

305 of the control region was excluded due to difficulty in correctly mapping sequencing reads
306 on repeat regions. To determine the accuracy of our RNASeq-derived mtDNA sequences, we
307 then compared the sequences of the six individuals for which we had both RNASeq and
308 WGS data (Supplementary Table 4). We identified four sites (site 2244, site, 5162, site 7162,
309 site 7920) that showed intra-individual variation across the two datasets in multiple
310 individuals (i.e. potential RNA editing). Two sites in coding regions (site 267, site 477) had
311 intra-individual differences in only one individual. Additionally, one individual had 12 sites in
312 the control region that differed across datasets (Supplementary Table 4). We replaced the
313 sites with intra-individual variation in multiple individuals with Ns, minimizing variation for
314 subsequent population analyses that might be introduced by RNA editing.

315 NUMT analysis

316 We found 42 mtDNA fragments (38 – 140 bps in length) that mapped to 42 individual
317 contigs of the cane toad draft nuclear genome with over 80% identity (Supplementary Table
318 5). All fragments corresponded to the control region of the mitogenome. After the draft
319 genome was masked for repeat elements, no mtDNA fragments were found to map to the
320 nuclear genome.

321 Population Genetic Analyses

322 We identified 262 (1.30%) polymorphic sites in the RNASeq dataset, 198 of which were
323 parsimony-informative, including 25 from RNA genes, 145 from protein-coding genes, and
324 28 from non-coding regions and the control region (Supplementary Table 2), which resulted
325 in the identification of 58 haplotypes (Table 1, Figure 4). The thirteen native samples
326 contained twelve unique haplotypes. Two haplotypes found in all introduced populations
327 accounted for nearly 50% of the sampled individuals. Within introduced populations, a

328 “complex-star” shape topology with mostly singleton mutants was observed and 41 of the
329 46 introduced haplotypes (89%) were private to a single introduction (Hawai’i: 10; Australia:
330 31).

331 We also compared the nearly-complete mitochondrial genome dataset with the *ND3*
332 dataset to demonstrate the consistency of haplotype inference across studies, and to put
333 our data into the context of previous work. A total of 24 haplotypes were identified in this
334 analysis (Figure 6); six haplotypes were identified in French Guiana/Guiana, one haplotype
335 was identified in Indonesian samples, three in Hawai’ian and Australian samples, and the
336 remaining haplotypes were from other South America localities. Haplotypes from either side
337 of Andes clearly formed distinct lineages which is consistent to previous work (Acevedo et
338 al., 2016). The single Australian haplotype (Hap1) previously identified in (Slade & Moritz,
339 1998) was the most common haplotype in our Australian samples and the other two
340 haplotypes (Hap2 and Hap3) within Australia were a single nucleotide different to this
341 common haplotype.

342 Haplotype diversity was highest in the native range, with reduced diversity in both
343 Hawai’ian and Australian populations (*h*: 0.9782 vs 0.8800 and 0.8344, respectively) (Table
344 1, Figure 5A). Native toads had the highest nucleotide diversity (π: 0.00469) compared to
345 considerably lower diversity in Hawai’ian and Australian toads (π: 0.00010 and 0.00009,
346 respectively) (Table 1, Figure 5B). Queensland toads had similar haplotype diversity but
347 higher nucleotide diversity and haplotype richness to Hawai’ian toads. The population from
348 the Northern Territory had noticeably lower genetic diversity and haplotype richness than
349 other Australian populations (Table 1, Figure 5A and 5C). When the relationship between

350 haplotypes across the invasion trajectory is considered, the twelve haplotypes from French
351 Guiana were highly dissimilar to those from the Australian and Hawai'ian populations,
352 forming two distinct clades (Figure 4). These were separated from each other by a minimum
353 of 110 base pair changes and from the common invasive haplotype by a minimum of 38
354 base pair changes (Figure 4). Australian and Hawai'ian haplotypes were intermixed and most
355 of the haplotypes were one base pair different to other haplotypes, with invasive
356 haplotypes being separated from at least one other haplotype by no more than three base
357 pair changes.

358 AMOVA results showed that the overall genetic variation among populations (64.49%) was
359 larger than the variation within populations (36.33%) (Table 2). Pairwise *Fst* values were
360 significant between the native population and all introduced populations (*Fst*: 0.5510 –
361 0.6686; Table 3). The *Fst* values between introduced populations were all less than 0.1,
362 indicating substantially less differentiation in the invasive range than in the native range.
363 The Hawai'ian population was significantly differentiated from Australian populations (*Fst*:
364 0.02247, $p = 0.018$), but this difference was driven by the Northern Territory (*Fst*: 0.05392, p
365 $= 0.01$ and Western Australian populations (*Fst*: 0.03343, $p = 0.013$); Queensland was not
366 significantly differentiated to the Hawai'i population (*Fst*: 0.01644, $p = 0.056$). Within
367 Australia, the Queensland population showed significant differentiation to both the
368 Northern Territory and Western Australian populations (*Fst*: 0.0411, $p = 0.009$ and 0.0303, p
369 $= 0.003$, respectively), whereas the Northern Territory and Western Australian populations
370 were not genetically differentiated (*Fst*: -0.01197, $p = 0.72$). In addition, there was a high
371 degree of gene flow within Australian populations (Nm: 11.6799 – infinity).

372 Two neutrality tests were implemented to study the demographic history of cane toad
373 populations: Fu's *Fs* and Ramos-Onsins and Rozas' R_2 statistic (Table 4). The native
374 population had non-significant positive Fu's *Fs* values and a large R_2 statistic, indicating the
375 population is not expanding or shrinking (i.e. is demographically stable). Fu's *Fs* values were
376 negative in all introduced populations, indicating an excess of rare haplotypes compared to
377 the expectation from neutrality and all these were statistically significant. R_2 statistics were
378 significant in Hawai'i, Queensland and Western Australia's populations.

379 Mismatch distribution graphs of all populations are shown in Supplementary Figure 1. The
380 native population showed a multimodal distribution with marginal non-significant SSD and
381 non-significant values for *Fs* and R_2, indicating no support for spatial expansion (Table 4).
382 However, there was some support for spatial expansion in the native range (Table 4). Most
383 introduced populations showed a unimodal distribution for both expansion models with
384 non-significant SSD and raggedness indices (Table 4), supporting the presence of spatial and
385 demographic expansions. However, the Northern Territory population did not support the
386 presence of spatial expansion (SSD's *p*-value: 0.032; raggedness' *p*-value: 0.046) and the
387 raggedness value for the demographic expansion model also had a significant value for this
388 population ($p = 0.032$).

389 Discussion

390 This study uses mitochondrial genome sequences to study the evolutionary and
391 phylogenetic relationships of the iconic invasive cane toad. Here, we constructed the first
392 reference mitochondrial genome of *R. marina* and placed this species within the order
393 Anura. We utilised RNASeq-derived mitogenomes to illustrate a greater number and

394 diversity of haplotype in mtDNA present in Australia and Hawai'i, in contrast to the
395 surprisingly low genetic diversity previously described using the *ND3* gene. Our conclusions
396 are further supported by evidence of demographic and spatial expansion in invasive cane
397 toad populations. This study provides new insights into the diversity of native cane toads
398 and how they have evolved along their invasion trajectory.

399 Phylogenetic inference

400 The phylogenetic relationships of toads within Anura are understudied. More specifically,
401 the cryptic *Rhinella* species complex remains poorly understood (Pereyra et al., 2021).
402 Currently, the evolutionary relationship of *Rhinella* species in the order Anura has yet to be
403 fully documented using complete mitochondrial genome or even nuclear genes and/or
404 mitochondrial genes. To address this knowledge gap, we generated the first complete
405 mitochondrial genome from any *Rhinella* species and investigated the phylogenetic
406 relationship of *Rhinella* to other Anurans. The topology of our tree differed in some respects
407 from previous hypotheses. Feng et al. (2017) built a tree for frog species using 95 unlinked
408 nuclear genes and suggested that *Rhinella* formed a sister clade with *Anaxyrus* within the
409 Bufonidae. They also suggested the genus *Bufo* is monophyletic with respect to
410 *Duttaphrynus*. However, our whole-mitochondrial genome phylogenetic analysis indicated
411 that the genus *Bufo* is a subclade of *Duttaphrynus*. Zhang et al. (2013) attempted to study
412 the Anuran phylogeny using partial mitochondrial genomes. However, no *Rhinella* species
413 were included in their study. Our study provides one of the most comprehensive
414 phylogenetic analyses of the relationships within the order Anura to date and provides a
415 starting point for future phylogenetic analyses within the genus *Rhinella*.

416 Choosing to use a complete dataset of a few selected loci versus a large supermatrix with
417 missing data for phylogenetic inference remains a controversial issue (see for review
418 (Lemmon, Brown, Stanger-Hall, & Lemmon, 2009; Roure, Baurain, & Philippe, 2013; Wiens &
419 Morrill, 2011)). Typically, phylogenetic trees are generated using single or multiple nuclear
420 genes, single or multiple nuclear genes plus mitochondrial genes, or multiple nuclear genes
421 with morphological trait matrices, but the use of complete mitochondrial genomes for
422 phylogenetic inference is rare due to resource limitations. Missing data in phylogenetic
423 analyses can bias phylogenetic calculation and tree branch lengths (Lemmon et al., 2009;
424 Xia, 2014). Nonetheless, contrasting findings have also suggested that when a sufficient
425 number of loci and taxa are included, phylogenetic signals remain consistent regardless of
426 whether or not the dataset includes missing data (Molloy & Warnow, 2018; Shavit Grievink,
427 Penny, & Holland, 2013). Despite using complete mitochondrial genomes in this study, 11%
428 of branches remained statistically unsupported, and the genera *Bufo*, *Bufotes*, *Anaxyrus* and
429 *Rhinella* in the Bufonidae family are also not well-supported. This result suggests that the
430 resolution power of the complete mitochondrial genome may not be sufficient, and/or
431 more representative sampling of different anuran species may be needed to provide more
432 robust support for statistical calculations and improve our understanding of the phylogeny.

433 ## Comparison between RNASeq & WGS results

434 The use of transcriptomic data to build mitochondrial genomes has been suggested as a
435 cost-effective alternative to long PCRs or direct sequencing of total DNA (Forni et al., 2019),
436 but it is important to ensure that the inferred polymorphisms accurately reflect genetic
437 variation. In our study, comparing mtDNA data obtained from WGS and RNASeq, we found
438 discrepancies in multiple nucleotide positions, all of which were in rRNA, tRNA and the

439 control region, which could be the result of post-transcriptional modifications or RNA
440 editing which are well-recognised phenomena across multiple taxa (Levanon et al., 2004;
441 Porath, Knisbacher, Eisenberg, & Levanon, 2017; St Laurent et al., 2013). While common
442 substitutions such as adenosine to inosine (A-to-I) deamination and cytidine to uridine (C-to-
443 U) pyrimidine exchange were found, we also observed less common substitutions (e.g.
444 thymine to cytidine (T-to-C) that have been previously documented in another toad species,
445 *Xenopus tropicalis* (Zaranek, Levanon, Zecharia, Clegg, & Church, 2010). These processes are
446 important for creating multiple transcript isoforms and gene regulation (Tang, Fei, & Page,
447 2012) and are known to be associated with adaptation (Duan, Dou, Luo, Zhang, & Lu, 2017;
448 Duan et al., 2021). RNA editing of the tRNA anticodon recognition site would extend the
449 codon matching ability and further increase the variation of the resulting translated
450 proteins. Although we did not observe nucleotide differences in position 34 or position 37 of
451 tRNA, which are typically recognised for deamination of adenosine to inosine which can be
452 paired with A, C or U residues (Jackman & Alfonzo, 2013), we did find changes in the D arm
453 of tRNA (Aspartic acid and Lysine), the T arm of tRNA (Phenylalanine) and the acceptor stem
454 of tRNA (Tryptophan) (Supplementary Figure 1), which are critical for the recognition of
455 aminoacyl tRNA synthetases to bind appropriate amino acids to its designated tRNA
456 (Ganesh & Maerkl, 2022). Failure to form a tRNA-amino acid complex could inhibit the use
457 of a specific amino acid in subsequent protein translation. Further proteome analysis is
458 required to determine if these changes affect the protein translation machinery.

459 Post-transcriptional modifications were more diverse in the native populations than in the
460 invasive populations. In particular, one individual from French Guiana displayed

461 discrepancies in the control region between WGS and RNASeq data, all of which were
462 transitions. This observation might result from tissue-specific expression in spleen or RNA
463 editing. Studies of RNA editing in mtDNA have primarily focused on protein-coding genes
464 and tRNAs, with less research conducted on the control region of the mtDNA, which is
465 responsible for regulating RNA and DNA synthesis. While the implications of potential RNA
466 editing in cane toads remain unclear, these events may contribute to diversity at a
467 transcriptional and translational level. The low diversity of post-transcriptional modification
468 in the invasive populations might be the result of selective sweep for favourable haplotypes
469 with low activity of RNA-editing enzymes.

470 NUMTs, or nuclear insertions of mitochondrial DNA fragments (Lopez, Yuhki, Masuda, Modi,
471 & O'Brien, 1994), can potentially present another source of genetic variation and contribute
472 to false-positive signals of variation in nuclear data (Song, Buhay, Whiting, & Crandall, 2008).
473 NUMTs have been reported in many different taxa including vertebrates, plants and protists
474 (Hazkani-Covo, Zeller, & Martin, 2010). They vary in size, ranging from minimal insertions to
475 hundreds of kilobases and tend to be longer in species with larger nuclear genomes
476 (Hazkani-Covo et al., 2010; Ko & Kim, 2016). Despite NUMTs generally being perceived as
477 non-functional due to their ability to shift reading frames, the presence of in-frame stop
478 codons and the use of different genetic codes, the transcription of NUMTs into non-coding
479 RNA molecules can be captured by RNASeq and potentially falsely interpreted as
480 mitochondrial variation among or within individuals. NUMTs studies on amphibian species
481 are limited. In *Xenopus tropicalis*, over 10kb of NUMTs has been found in the nuclear
482 genome (Hazkani-Covo et al., 2010). Investigation of the presence/absence of NUMTs in

483 cane toad nuclear genome and the potential expression of NUMTs is required to distinguish
484 between genuine mtDNA variation within individuals and false-positive signals caused by
485 incorporating nuclear genetic data into mitochondrial genetic analyses.

486 A search for NUMTs in the published cane toad draft genome (Edwards et al., 2018) yielded
487 only matches with simple repeat regions within the mtDNA control region. In other species,
488 NUMTs originate across the entire mtDNA genome (Dayama, Zhou, Prado-Martinez,
489 Marques-Bonet, & Mills, 2020; Wei et al., 2022). Following the repeat-masking of the draft
490 nuclear genome, we found no evidence of NUMTs suggesting that the putative NUMTs were
491 just low-complexity regions in the nuclear genome. Similar to other gene-searching methods
492 such as BUSCO, the accuracy of the search algorithm heavily depends on the contiguity of
493 the genome and the divergence of the query from the database subjects. Given the high
494 number of contigs in the cane toad draft genome, potential NUMTs hits may have been
495 missed through breakpoints. Therefore, it is important to develop a cane toad genome
496 assembly with higher contiguity to improve our ability to investigate NUMTs. Additionally,
497 ancient NUMTs may be too divergent to be detected by the methods employed, but NUMTs
498 with such low identity would not interfere with mtDNA analysis.

499 It is worth noting that the consensus sequences generated from the bioinformatic pipelines
500 used in this study represented the most common nucleotide at positions where multiple
501 nucleotides were found. Although these discrepancies could result from sequencing errors,
502 our filtering is likely to have prevented this. As discussed above, RNA editing or NUMTs
503 could have resulted in putative polymorphisms within individuals in our RNASeq dataset. It
504 is also possible that in some cases, heteroplasmy could have caused these putative

polymorphisms. While determining the cause of this phenomenon is beyond the scope of this study, it is unlikely that this has affected population-level statistics.

Population genetic analyses

Interestingly, both the haplotype network (Figure 4) and the multimodal mismatch distribution (Supplementary Figure 1) indicate the presence of multiple genetically divergent lineages within native populations. Notably, individuals within each lineage are intermixed, rather than being grouped by geographical locations. For instance, despite an 80 km geographical separation, haplotypes H4 and H7 are found in the same lineage while haplotypes H12 and H11 are only 9 m apart geographically but belong to different lineages. Native-range cane toads can travel 60-180 m per move for sheltering and foraging purposes (Shine et al., 2021). Consequently, the possibility of hybridisation between toads from these distinct lineages exists. However, a recent study in South and Central America which utilised ddRAD sequencing has offered a different perspective on the phylogenetic relationship of cane toads. This study revealed toads from the coastal area formed two sister clades, distinct from the clade formed by toads from the rainforest regions (Mittan-Moreau et al., 2022). It becomes crucial to ascertain the precise populations of cane toads from the native range that were used as founding individuals, especially given the limited comprehensive description in prior research (Easteal, 1981). In Shine et al. (2021), cane toads from the range-edge regions (Northern Territory and Western Australia) exhibited similar dispersal patterns to native toads in one of the coastal sites. This suggests the adaptation of cane toads in hot and arid environments is likely facilitated by the standing genetic variation(s) inherited from the progenitors, rather than relying solely on the genetic adaptation and selection of novel mutation(s). A more thorough investigation into these distinct lineages

528 becomes imperative for unravelling the mechanisms underlying the adaptation of cane
529 toads in Australia.

530 Our study challenges the assumption of low genetic diversity in Hawai'ian and Australian
531 populations. Slade and Moritz (1998) reported a unique haplotype at *ND3* within Australian
532 populations, in comparison with multiple haplotypes in native populations in Central and
533 South America. This result supported the presence of founder effects following the
534 introduction of cane toads into Hawai'i and Australia. Here, we revealed two additional,
535 previously undocumented, *ND3* haplotypes (Hap2 and Hap3) in Hawai'ian and Australian
536 populations, each of which differed by only one nucleotide from the previously identified
537 haplotype (Hap1) (Figure 6). In comparison to our analysis using nearly complete
538 mitogenomes discussed above, this analysis highlights that a single gene is not sufficient to
539 accurately infer genetic diversity.

540 The significant decrease in haplotype diversity and nucleotide diversity observed in both
541 Hawai'ian and Australian populations, compared to native populations, does support that
542 genetic bottlenecks and founder effects occurred post-introduction. This finding is
543 consistent with previous studies on this species using nuclear genetic markers including
544 SNPs, microsatellite markers, and MHC class I and II loci (M. Lillie et al., 2016; M. Lillie et al.,
545 2014; Selechnik, Richardson, Shine, DeVore, et al., 2019). In this study, all three neutrality
546 indices showed strong signals of population expansion and the observed data fitted to both
547 demographic expansion and spatial expansion models in most invasive sampling sites (Table
548 4, Figure 5). This conclusion is supported by the star-like shaped topology of the haplotype
549 network created using the nearly-complete mitochondrial genome dataset. Despite reduced

diversity in the introduced range, the degree of the decrease is less than previously reported (Leblois et al., 2000). Additionally, the star-shaped topology of the network containing the introduced haplotypes, the absence of native samples here, and the significant negative Fu's *Fs* values indicate that many of the singleton haplotypes discovered in Hawai'i and Australia may be the products of post-introduction mutations.

We observed a curvilinear pattern of genetic diversity indices across Australian populations (from Queensland to Western Australia), a pattern also found in spleen gene expression (Selechnik, Richardson, Shine, Brown, et al., 2019), limb length (Hudson, McCurry, Lundgren, McHenry, & Shine, 2016), spleen and body mass (Gregory P. Brown, Kelehear, Shilton, Phillips, & Shine, 2015). The mechanism behind this reduction of genetic diversity in Northern Territory toads remains unclear, but it may be due to demographic processes that reduce the effective population size and increase genetic drift (Nei & Tajima, 1981), genetic hitchhiking that removes neutral polymorphic sites under positive selection (Fay & Wu, 2000) or directional selection due to harsh climate (Selechnik, Richardson, Shine, DeVore, et al., 2019), possibly in combination with dispersion-driven spatial sorting at the invasion front (Clarke, Shine, & Phillips, 2019). Despite our finding that the Northern Territory and Western Australia populations were not genetically differentiated, it is likely that they have experienced different local evolutionary forces and demographic processes. Northern Territory toads, unlike other populations, only exhibited demographic expansion and not spatial expansion, as supported by lower levels in all statistical measures.

Our results demonstrate significant differentiation between native and introduced populations (Table 3). We identified three genetic clusters: 1) native toads, 2) Hawai'i and

572 Queensland toads (*Fst*: 0.016, $p = 0.056$) and 3) Northern Territory and Western Australia
573 toads (*Fst*: -0.012, $p = 0.72$), consistent with previous results using SNPs from the nuclear
574 genome (Selechnik, Richardson, Shine, DeVore, et al., 2019). It is interesting that the
575 Queensland toads are not differentiated to those sampled in Hawai'i but differ from toads
576 sampled in the Northern Territory and Western Australia. Queensland toads have been
577 separated from their Hawai'ian conspecifics for almost 90 years but do share similar
578 environmental conditions [42]. Northern Territory and Western Australian toads suffer from
579 reduced genetic diversity compared to Queensland toads, so they have likely experienced
580 more genetic drift, which may explain this difference. However, it is also possible that the
581 striking reduction in precipitation and increase temperature experienced by Northern
582 Territory and Western Australian toads compared to those from Queensland [42] has
583 resulted in selective differences in the mitogenome between these populations. In support
584 of this possibility, significant negative Fu's Fs can also indicate the presence of selection.

585 ## Conclusion

586 Our study used the whole mtDNA genome as a genetic marker to investigate the evolution
587 of cane toads across their invasion trajectory from the native range to Australia, adding to
588 our knowledge of changes to genetic diversity as cane toads have been serially introduced
589 to multiple locations. Unsurprisingly, we found reduced genetic diversity in introduced
590 populations, suggesting that genetic bottlenecks have occurred in these populations.
591 Despite this reduction, our data indicate greater genetic diversity in introduced populations
592 than has been previously described using mtDNA. We identified signatures of demographic
593 and spatial expansion in most introduced localities we sampled, consistent with ongoing
594 invasion. We described 46 haplotypes in introduced populations, mostly consisting of single

base-pair changes, suggesting evolution following introduction. By the comparison of WGS- and RNASeq-derived mitogenomes, we have identified the presence of post-transcriptional modification or RNA editing, which may supply additional genetic diversity for upon which selection can act.

Acknowledgement

We thank Mark Richardson and Yini Peng Lee for assistance with genomic sequencing and processing of genomic data and the Ramaciotti Centre for Genomics, University of New South Wales (UNSW) and Deakin Genomic Centre, Deakin University who performed the sequencing in this study. This research includes computations using the computational cluster Katana supported by Research Technology Services at UNSW Sydney. This study was supported by the Australian Research Council (grants DP160102991, DP190100507, to RS and LAR), the UNSW Scientia Program (to LAR), and the UNSW Scientia PhD Scholarship (to KC).

Author Contributions

KC, RS, RJE, and LAR developed this project. LAR conducted DNA extractions. KC, TGA, RJE, and LAR analysed the data. All authors contributed to data interpretation and writing.

Data availability

The haplotype sequences will be available on the NCBI public database once the manuscript is accepted.

References

Acevedo, A. A., Lampo, M., & Cipriani, R. (2016). The cane or marine toad, Rhinella marina (Anura, Bufonidae): two genetically and morphologically distinct species. *Zootaxa, 4103*(6), 574-586. doi:10.11646/zootaxa.4103.6.7

Alford, R., Brown, G., Schwarzkopf, L., Phillips, B., & Shine, R. (2009). Comparisons through time and space suggest rapid evolution of dispersal behavior in an invasive species. *Wildlife Research, 36*, 23-28. doi:10.1071/wr08021

Allendorf, F. W., & Lundquist, L. L. (2003). Introduction: Population Biology, Evolution, and Control of Invasive Species. *Conservation Biology, 17*(1), 24-30. doi:https://doi.org/10.1046/j.1523-1739.2003.02365.x

Baker, E. P., Peris, D., Moriarty, R. V., Li, X. C., Fay, J. C., & Hittinger, C. T. (2019). Mitochondrial DNA and temperature tolerance in lager yeasts. *Sci Adv, 5*(1), eaav1869. doi:10.1126/sciadv.aav1869

Bandelt, H. J., Forster, P., & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol, 16*(1), 37-48. doi:10.1093/oxfordjournals.molbev.a026036

Bandelt, H. J., Forster, P., Sykes, B. C., & Richards, M. B. (1995). Mitochondrial portraits of human populations using median networks. *Genetics, 141*(2), 743-753. doi:10.1093/genetics/141.2.743

Bernt, M., Donath, A., Juhling, F., Externbrink, F., Florentz, C., Fritzsch, G., . . . Stadler, P. F. (2013). MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol, 69*(2), 313-319. doi:10.1016/j.ympev.2012.08.023

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114-2120. doi:10.1093/bioinformatics/btu170

Bonebrake, T. C., Brown, C. J., Bell, J. D., Blanchard, J. L., Chauvenet, A., Champion, C., . . . Pecl, G. T. (2018). Managing consequences of climate-driven species redistribution requires integration of ecology, conservation and social science. *Biological Reviews, 93*(1), 284-305. doi:10.1111/brv.12344

Bradshaw, C. J., Leroy, B., Bellard, C., Roiz, D., Albert, C., Fournier, A., . . . Courchamp, F. (2016). Massive yet grossly underestimated global costs of invasive insects. *Nat Commun, 7*, 12986. doi:10.1038/ncomms12986

Brown, G. P., Kelehear, C., Shilton, C. M., Phillips, B. L., & Shine, R. (2015). Stress and immunity at the invasion front: a comparison across cane toad (Rhinella marina) populations. *Biological Journal of the Linnean Society, 116*(4), 748-760. doi:10.1111/bij.12623

Brown, G. P., Phillips, B. L., Dubey, S., & Shine, R. (2015). Invader immunology: invasion history alters immune system function in cane toads (Rhinella marina) in tropical Australia. *Ecol Lett, 18*(1), 57-65. doi:10.1111/ele.12390

Chen, Y., Gong, L., Liu, X., Chen, X., Yang, S., & Luo, Y. (2020). Mitochondrial DNA genomes revealed different patterns of high-altitude adaptation in high-altitude Tajiks compared with Tibetans and Sherpas. *Sci Rep, 10*(1), 10592. doi:10.1038/s41598-020-67519-z

Cheng, Y. T., Liu, J., Yang, L. Q., Sun, C., & Kong, Q. P. (2013). Mitochondrial DNA content contributes to climate adaptation using Chinese populations as a model. *PLoS One, 8*(11), e79536. doi:10.1371/journal.pone.0079536

Chinchio, E., Crotta, M., Romeo, C., Drewe, J. A., Guitian, J., & Ferrari, N. (2020). Invasive alien species and disease risk: An open challenge in public and animal health. *PLoS Pathog, 16*(10), e1008922. doi:10.1371/journal.ppat.1008922

Clarke, G. S., Shine, R., & Phillips, B. L. (2019). May the (selective) force be with you: Spatial sorting and natural selection exert opposing forces on limb length in an invasive amphibian. *J Evol Biol, 32*(9), 994-1001. doi:10.1111/jeb.13504

Davey, N. E., Shields, D. C., & Edwards, R. J. (2006). SLiMDisc: short, linear motif discovery, correcting for common evolutionary descent. *Nucleic Acids Res, 34*(12), 3546-3554. doi:10.1093/nar/gkl486

Dayama, G., Zhou, W., Prado-Martinez, J., Marques-Bonet, T., & Mills, R. E. (2020). Characterization of nuclear mitochondrial insertions in the whole genomes of primates. *NAR Genom Bioinform, 2*(4), lqaa089. doi:10.1093/nargab/lqaa089

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., . . . Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics, 29*(1), 15-21. doi:10.1093/bioinformatics/bts635

Duan, Y., Dou, S., Luo, S., Zhang, H., & Lu, J. (2017). Adaptation of A-to-I RNA editing in Drosophila. *PLoS Genet, 13*(3), e1006648. doi:10.1371/journal.pgen.1006648

Duan, Y., Dou, S., Porath, H. T., Huang, J., Eisenberg, E., & Lu, J. (2021). A-to-I RNA editing in honeybees shows signals of adaptation and convergent evolution. *iScience, 24*(1), 101983. doi:10.1016/j.isci.2020.101983

Easteal, S. (1981). The history of introductions of Bufo marinus (Amphibia: Anura); a natural experiment in evolution. *Biological Journal of the Linnean Society, 16*(2), 93-113. doi:10.1111/j.1095-8312.1981.tb01645.x

Edwards, R. J., Field, M. A., Ferguson, J. M., Dudchenko, O., Keilwagen, J., Rosen, B. D., . . . Ballard, J. W. O. (2021). Chromosome-length genome assembly and structural variations of the primal Basenji dog (Canis lupus familiaris) genome. *BMC Genomics, 22*(1), 188. doi:10.1186/s12864-021-07493-6

Edwards, R. J., Paulsen, K., Aguilar Gomez, C. M., & Perez-Bercoff, A. (2020). Computational Prediction of Disordered Protein Motifs Using SLiMSuite. *Methods Mol Biol, 2141*, 37-72. doi:10.1007/978-1-0716-0524-0_3

Edwards, R. J., Tuipulotu, D. E., Amos, T. G., O'Meally, D., Richardson, M. F., Russell, T. L., . . . White, P. A. (2018). Draft genome assembly of the invasive cane toad, Rhinella marina. *Gigascience, 7*(9). doi:10.1093/gigascience/giy095

Estoup, A., Ravigné, V., Hufbauer, R., Vitalis, R., Gautier, M., & Facon, B. (2016). Is There a Genetic Paradox of Biological Invasion? *Annual Review of Ecology, Evolution, and Systematics, 47*(1), 51-72. doi:10.1146/annurev-ecolsys-121415-032116

Excoffier, L. (2004). Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Mol Ecol, 13*(4), 853-864. doi:10.1046/j.1365-294x.2003.02004.x

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour, 10*(3), 564-567. doi:10.1111/j.1755-0998.2010.02847.x

Fay, J. C., & Wu, C. I. (2000). Hitchhiking under positive Darwinian selection. *Genetics, 155*(3), 1405-1413. doi:10.1093/genetics/155.3.1405

Feng, Y. J., Blackburn, D. C., Liang, D., Hillis, D. M., Wake, D. B., Cannatella, D. C., & Zhang, P. (2017). Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous-Paleogene boundary. *Proc Natl Acad Sci U S A, 114*(29), E5864-E5870. doi:10.1073/pnas.1704632114

Forni, G., Puccio, G., Bourguignon, T., Evans, T., Mantovani, B., Rota-Stabelli, O., & Luchetti, A. (2019). Complete mitochondrial genomes from transcriptomes: assessing pros and cons of data mining for assembling new mitogenomes. *Sci Rep, 9*(1), 14806. doi:10.1038/s41598-019-51313-7

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics, 147*(2), 915-925. doi:10.1093/genetics/147.2.915

Ganesh, R. B., & Maerkl, S. J. (2022). Biochemistry of Aminoacyl tRNA Synthetase and tRNAs and Their Engineering for Cell-Free and Synthetic Cell Applications. *Front Bioeng Biotechnol, 10*, 918659. doi:10.3389/fbioe.2022.918659

Goudet, J. (1995). FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. *Journal of Heredity, 86*(6), 485-486. doi:10.1093/oxfordjournals.jhered.a111627

Gravel, S. (2016). When Is Selection Effective? *Genetics, 203*(1), 451-462. doi:10.1534/genetics.115.184630

Gruber, J., Brown, G., Whiting, M. J., & Shine, R. (2017). Is the behavioural divergence between range-core and range-edge populations of cane toads (Rhinella marina) due to evolutionary change or developmental plasticity? *R Soc Open Sci, 4*(10), 170789. doi:10.1098/rsos.170789

Gruber, J., Brown, G., Whiting, M. J., & Shine, R. (2018). Behavioural divergence during biological invasions: a study of cane toads (Rhinella marina) from contrasting environments in Hawai'i. *R Soc Open Sci, 5*(4), 180197. doi:10.1098/rsos.180197

Gruber, J., Whiting, M. J., Brown, G., & Shine, R. (2017). The loneliness of the long-distance toad: invasion history and social attraction in cane toads (Rhinella marina). *Biol Lett, 13*(11). doi:10.1098/rsbl.2017.0445

Hazkani-Covo, E., Zeller, R. M., & Martin, W. (2010). Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genet, 6*(2), e1000834. doi:10.1371/journal.pgen.1000834

Hudson, C. M., Brown, G. P., & Shine, R. (2016). It is lonely at the front: contrasting evolutionary trajectories in male and female invaders. *R Soc Open Sci, 3*(12), 160687. doi:10.1098/rsos.160687

Hudson, C. M., McCurry, M. R., Lundgren, P., McHenry, C. R., & Shine, R. (2016). Constructing an Invasion Machine: The Rapid Evolution of a Dispersal-Enhancing Phenotype During the Cane Toad Invasion of Australia. *PLoS One, 11*(9), e0156950. doi:10.1371/journal.pone.0156950

Jackman, J. E., & Alfonzo, J. D. (2013). Transfer RNA modifications: nature's combinatorial chemistry playground. *Wiley Interdiscip Rev RNA, 4*(1), 35-48. doi:10.1002/wrna.1144

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods, 14*(6), 587-589. doi:10.1038/nmeth.4285

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol, 30*(4), 772-780. doi:10.1093/molbev/mst010

Klucnika, A., & Ma, H. (2019). A battle for transmission: the cooperative and selfish animal mitochondrial genomes. *Open Biol, 9*(3), 180267. doi:10.1098/rsob.180267

Ko, Y. J., & Kim, S. (2016). Analysis of Nuclear Mitochondrial DNA Segments of Nine Plant Species: Size, Distribution, and Insertion Loci. *Genomics Inform, 14*(3), 90-95. doi:10.5808/GI.2016.14.3.90

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods, 9*(4), 357-359. doi:10.1038/nmeth.1923

Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics, 30*(22), 3276-3278. doi:10.1093/bioinformatics/btu531

Leblois, R., Rousset, F., Tikel, D., Moritz, C., & Estoup, A. (2000). Absence of evidence for isolation by distance in an expanding cane toad (Bufo marinus) population: an individual-based analysis of microsatellite genotypes. *Mol Ecol, 9*(11), 1905-1909. doi:10.1046/j.1365-294x.2000.01091.x

Lemmon, A. R., Brown, J. M., Stanger-Hall, K., & Lemmon, E. M. (2009). The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference. *Syst Biol, 58*(1), 130-145. doi:10.1093/sysbio/syp017

Levanon, E. Y., Eisenberg, E., Yelin, R., Nemzer, S., Hallegger, M., Shemesh, R., . . . Jantsch, M. F. (2004). Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat Biotechnol, 22*(8), 1001-1005. doi:10.1038/nbt996

Lillie, M., Cui, J., Shine, R., & Belov, K. (2016). Molecular characterization of MHC class II in the Australian invasive cane toad reveals multiple splice variants. *Immunogenetics, 68*(6-7), 449-460. doi:10.1007/s00251-016-0919-9

Lillie, M., Dubey, S., Shine, R., & Belov, K. (2017). Variation in Major Histocompatibility Complex diversity in invasive cane toad populations. *Wildlife Research, 44*(7), 565. doi:10.1071/wr17055

Lillie, M., Shine, R., & Belov, K. (2014). Characterisation of major histocompatibility complex class I in the Australian cane toad, Rhinella marina. *PLoS One, 9*(8), e102824. doi:10.1371/journal.pone.0102824

Lopez, J. V., Yuhki, N., Masuda, R., Modi, W., & O'Brien, S. J. (1994). Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. *J Mol Evol, 39*(2), 174-190. doi:10.1007/BF00163806

Mainka, S. A., & Howard, G. W. (2010). Climate change and invasive species: double jeopardy. *Integr Zool, 5*(2), 102-111. doi:10.1111/j.1749-4877.2010.00193.x

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., . . . DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res, 20*(9), 1297-1303. doi:10.1101/gr.107524.110

Meiklejohn, C. D., Montooth, K. L., & Rand, D. M. (2007). Positive and negative selection on the mitochondrial genome. *Trends Genet, 23*(6), 259-263. doi:10.1016/j.tig.2007.03.008

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol, 37*(5), 1530-1534. doi:10.1093/molbev/msaa015

Mittan-Moreau, C. S., Kelehear, C., Toledo, L. F., Bacon, J., Guayasamin, J. M., Snyder, A., & Zamudio, K. R. (2022). Cryptic lineages and standing genetic variation across independent cane toad introductions. *Molecular Ecology, 31*(24), 6440-6456. doi:10.1111/mec.16713

Molloy, E. K., & Warnow, T. (2018). To Include or Not to Include: The Impact of Gene Filtering on Species Tree Estimation Methods. *Syst Biol, 67*(2), 285-303. doi:10.1093/sysbio/syx077

Nei, M., & Tajima, F. (1981). Genetic drift and estimation of effective population size. *Genetics, 98*(3), 625-640. doi:10.1093/genetics/98.3.625

Pereyra, M., Blotto, B. L., Baldo, D., Chaparro, J. C., Ron, S., Elias-Costa, A., . . . Faivovich, J. (2021). Evolution in the Genus Rhinella: A Total Evidence Phylogenetic Analysis of Neotropical True Toads (Anura: Bufonidae). *Bulletin of the American Museum of Natural History, 447*, 1–156. doi:10.1206/0003-0090.447.1.1

Phillips, B. L., Brown, G. P., & Shine, R. (2010). Evolutionarily accelerated invasions: the rate of dispersal evolves upwards during the range advance of cane toads. *J Evol Biol, 23*(12), 2595-2601. doi:10.1111/j.1420-9101.2010.02118.x

Polzin, T., & Vahdati Daneshmand, S. (2003). On Steiner Trees and Minimum Spanning Trees in Hypergraphs. *Operations Research Letters, 31*, 12-20. doi:10.1016/S0167-6377(02)00185-2

Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Auwera, G. A. V. d., . . . Banks, E. (2018). Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*, 201178. doi:10.1101/201178

Porath, H. T., Knisbacher, B. A., Eisenberg, E., & Levanon, E. Y. (2017). Massive A-to-I RNA editing is common across the Metazoa and correlates with dsRNA abundance. *Genome Biol, 18*(1), 185. doi:10.1186/s13059-017-1315-y

Radford, I. J., Woolley, L.-A., Dickman, C. R., Corey, B., Trembath, D., & Fairman, R. (2019). Invasive species-driven trophic cascades: Are cane toads indirectly contributing to small mammal collapses across tropical Australia? *bioRxiv*, 616771. doi:10.1101/616771

Ramirez-Soriano, A., Ramos-Onsins, S. E., Rozas, J., Calafell, F., & Navarro, A. (2008). Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. *Genetics, 179*(1), 555-567. doi:10.1534/genetics.107.083006

Ramos-Onsins, S. E., & Rozas, J. (2002). Statistical properties of new neutrality tests against population growth. *Mol Biol Evol, 19*(12), 2092-2100. doi:10.1093/oxfordjournals.molbev.a004034

Richardson, M. F., Sequeira, F., Selechnik, D., Carneiro, M., Vallinoto, M., Reid, J. G., . . . Rollins, L. A. (2018). Improving amphibian genomic resources: a multitissue reference transcriptome of an iconic invader. *Gigascience, 7*(1), 1-7. doi:10.1093/gigascience/gix114

Rollins, L. A., Moles, A. T., Lam, S., Buitenwerf, R., Buswell, J. M., Brandenburger, C. R., . . . Sherwin, W. B. (2013). High genetic diversity is not essential for successful introduction. *Ecology and Evolution, 3*(13), 4501-4517. doi:10.1002/ece3.824

Rollins, L. A., Richardson, M. F., & Shine, R. (2015). A genetic perspective on rapid evolution in cane toads (Rhinella marina). *Mol Ecol, 24*(9), 2264-2276. doi:10.1111/mec.13184

Roure, B., Baurain, D., & Philippe, H. (2013). Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. *Mol Biol Evol, 30*(1), 197-214. doi:10.1093/molbev/mss208

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sanchez-Gracia, A. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol Biol Evol, 34*(12), 3299-3302. doi:10.1093/molbev/msx248

Schneider, S., & Excoffier, L. (1999). Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics, 152*(3), 1079-1089. doi:10.1093/genetics/152.3.1079

Schrieber, K., & Lachmuth, S. (2017). The Genetic Paradox of Invasions revisited: the potential role of inbreeding x environment interactions in invasion success. *Biol Rev Camb Philos Soc, 92*(2), 939-952. doi:10.1111/brv.12263

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., . . . Essl, F. (2017). No saturation in the accumulation of alien species worldwide. *Nat Commun, 8*, 14435. doi:10.1038/ncomms14435

Selechnik, D., Richardson, M. F., Shine, R., Brown, G. P., & Rollins, L. A. (2019). Immune and environment-driven gene expression during invasion: An eco-immunological application of RNA-Seq. *Ecol Evol, 9*(11), 6708-6721. doi:10.1002/ece3.5249

Selechnik, D., Richardson, M. F., Shine, R., DeVore, J. L., Ducatez, S., & Rollins, L. A. (2019). Increased Adaptive Variation Despite Reduced Overall Genetic Diversity in a Rapidly Adapting Invader. *Front Genet, 10*, 1221. doi:10.3389/fgene.2019.01221

Shavit Grievink, L., Penny, D., & Holland, B. R. (2013). Missing data and influential sites: choice of sites for phylogenetic analysis can be as important as taxon sampling and model choice. *Genome Biol Evol, 5*(4), 681-687. doi:10.1093/gbe/evt032

Shine, R. (2010). The ecological impact of invasive cane toads (Bufo marinus) in Australia. *Q Rev Biol, 85*(3), 253-291. doi:10.1086/655116

Shine, R., Alford, R. A., Blennerhasset, R., Brown, G. P., DeVore, J. L., Ducatez, S., . . . Phillips, B. L. (2021). Increased rates of dispersal of free-ranging cane toads (Rhinella marina) during their global invasion. *Sci Rep, 11*(1), 23574. doi:10.1038/s41598-021-02828-5

Silva, G., Lima, F. P., Martel, P., & Castilho, R. (2014). Thermal adaptation and clinal mitochondrial DNA variation of European anchovy. *Proc Biol Sci, 281*(1792). doi:10.1098/rspb.2014.1093

Slade, R. W., & Moritz, C. (1998). Phylogeography of Bufo marinus from its natural and introduced ranges. *Proc Biol Sci, 265*(1398), 769-777. doi:10.1098/rspb.1998.0359

Sm, L., Browne, M., Boudjelas, S., & De Poorter, M. (2000). 100 of the World's Worst Invasive Alien Species: A Selection From the Global Invasive Species Database. *Published by The Invasive Species Specialist Group (ISSG) a specialist group of the Species Survival Commission (SSC) of the World Conservation Union (IUCN), 12pp. First published as special lift-out in Aliens, 12*.

Song, H., Buhay, J. E., Whiting, M. F., & Crandall, K. A. (2008). Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proc Natl Acad Sci U S A, 105*(36), 13486-13491. doi:10.1073/pnas.0803076105

St Laurent, G., Tackett, M. R., Nechkin, S., Shtokalo, D., Antonets, D., Savva, Y. A., . . . Reenan, R. A. (2013). Genome-wide analysis of A-to-I RNA editing by single-molecule sequencing in Drosophila. *Nat Struct Mol Biol, 20*(11), 1333-1339. doi:10.1038/nsmb.2675

Tang, W., Fei, Y., & Page, M. (2012). Biological Significance of RNA Editing in Cells. *Molecular Biotechnology, 52*(1), 91-100. doi:10.1007/s12033-012-9498-7

Tarailo-Graovac, M., & Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinformatics, Chapter 4*, 4 10 11-14 10 14. doi:10.1002/0471250953.bi0410s25

Turvey, N. (2013). *Cane toads: a tale of sugar, politics and flawed science*: Sydney University Press.

Urban, M. C., Phillips, B. L., Skelly, D. K., & Shine, R. (2008). A toad more traveled: the heterogeneous invasion dynamics of cane toads in Australia. *Am Nat, 171*(3), E134-148. doi:10.1086/527494

Van der Auwera, G. A. (2020). *Genomics in the cloud : using Docker, GATK, and WDL in Terra* (First edition. ed.). Sebastopol, CA: Sebastopol, CA : O'Reilly Media.

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., . . . Earl, A. M. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One, 9*(11), e112963. doi:10.1371/journal.pone.0112963

Wang, X., Zhou, S., Wu, X., Wei, Q., Shang, Y., Sun, G., . . . Zhang, H. (2021). High-altitude adaptation in vertebrates as revealed by mitochondrial genome analyses. *Ecol Evol, 11*(21), 15077-15084. doi:10.1002/ece3.8189

Wei, W., Schon, K. R., Elgar, G., Orioli, A., Tanguy, M., Giess, A., . . . Chinnery, P. F. (2022). Nuclear-embedded mitochondrial DNA sequences in 66,083 human genomes. *Nature, 611*(7934), 105-114. doi:10.1038/s41586-022-05288-7

Wiens, J. J., & Morrill, M. C. (2011). Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. *Syst Biol, 60*(5), 719-731. doi:10.1093/sysbio/syr025

Wright, E. S. (2016). Using DECIPHER v2.0 to analyze big biological sequence data in R. *R Journal, 8*, 352-359.

Wright, E. S. (2020). RNAconTest: comparing tools for noncoding RNA multiple sequence alignment based on structural consistency. *RNA, 26*(5), 531-540. doi:10.1261/rna.073015.119

Xia, X. (2014). *Phylogenetic Bias in the Likelihood Method Caused by Missing Data Coupled with Among-Site Rate Variation: An Analytical Approach* (Vol. 8492).

Xiao, S., Nguyen, D. T., Wu, B., & Hao, W. (2017). Genetic Drift and Indel Mutation in the Evolution of Yeast Mitochondrial Genome Size. *Genome Biol Evol, 9*(11), 3088-3099. doi:10.1093/gbe/evx232

Yagound, B., West, A. J., Richardson, M. F., Selechnik, D., Shine, R., & Rollins, L. A. (2022). Brain transcriptome analysis reveals gene expression differences associated with dispersal behaviour between range-front and range-core populations of invasive cane toads in Australia. *Mol Ecol, 31*(6), 1700-1715. doi:10.1111/mec.16347

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T. Y. (2017). ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution, 8*(1), 28-36. doi:10.1111/2041-210x.12628

Zaranek, A. W., Levanon, E. Y., Zecharia, T., Clegg, T., & Church, G. M. (2010). A survey of genomic traces reveals a common sequencing error, RNA editing, and DNA editing. *PLoS Genet, 6*(5), e1000954. doi:10.1371/journal.pgen.1000954

Zhang, P., Liang, D., Mao, R. L., Hillis, D. M., Wake, D. B., & Cannatella, D. C. (2013). Efficient sequencing of Anuran mtDNAs and a mitogenomic exploration of the phylogeny and evolution of frogs. *Mol Biol Evol, 30*(8), 1899-1915. doi:10.1093/molbev/mst091

Zug, G. R. a. Z., Patricia B. (1979). The marine toad, Bufo marinus : a natural history resumé of native populations. doi:https://doi.org/10.5479/si.00810282.284

Tables and Figures

Figure 1: Sampling locations of the cane toad (*Rhinella marina*) in (A) Australia and (B) French Guiana (the native range) and (C) Hawai'i. Red dots in French Guiana, Hawai'i and Australia indicate sample collection sites.

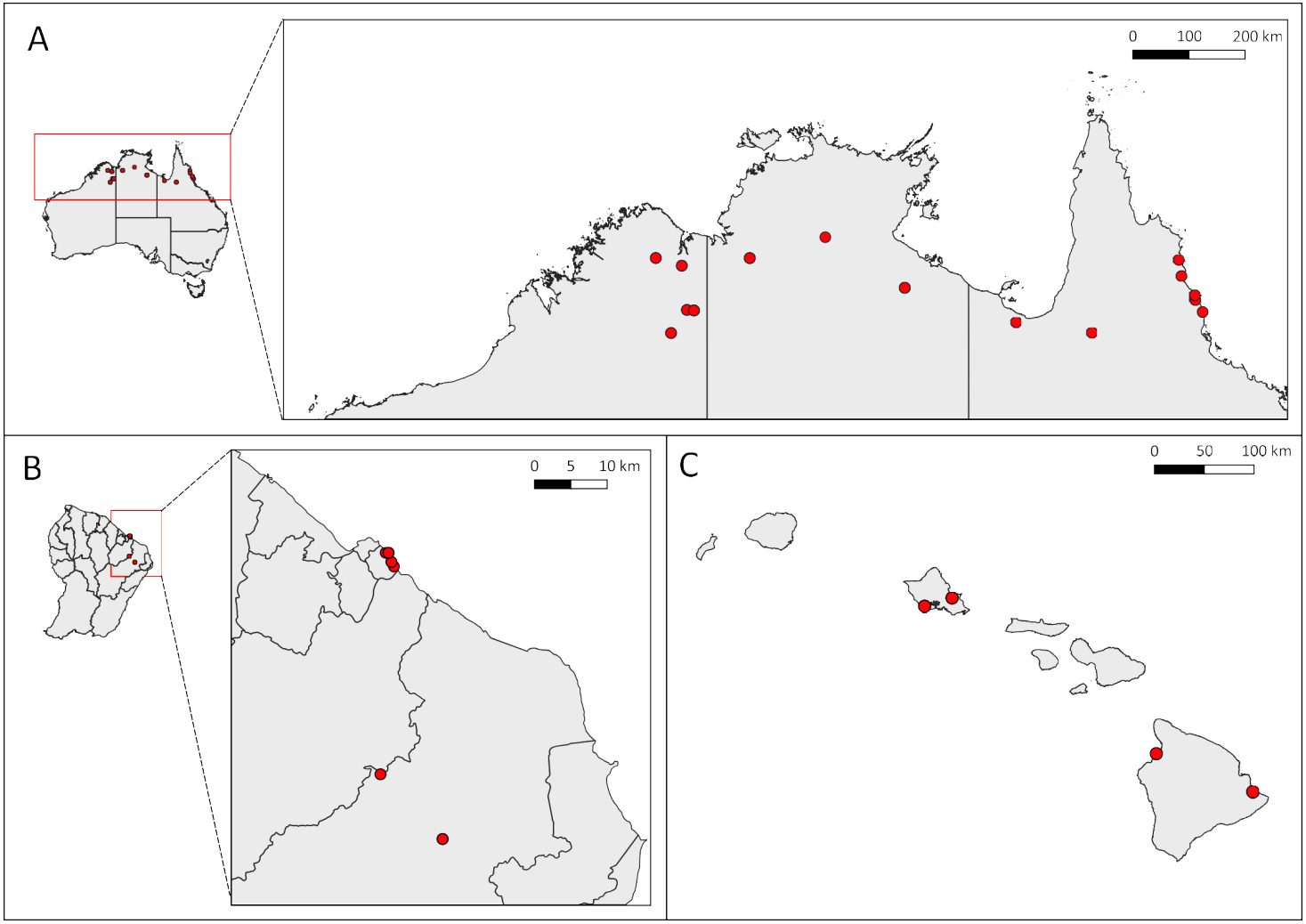


Figure 2: Mitochondrial genome structure of *R. marina*. Genes, GC skew and GC content were shown in the inner circle, middle circle and outer circle respectively. The innermost circle indicates the negative strand and the outer circle indicates the positive strand. Genes are coloured by type: red, tRNA; pink, rRNA; blue, coding sequence. Starting and stop positions of each gene are described in Supplementary Table 3.

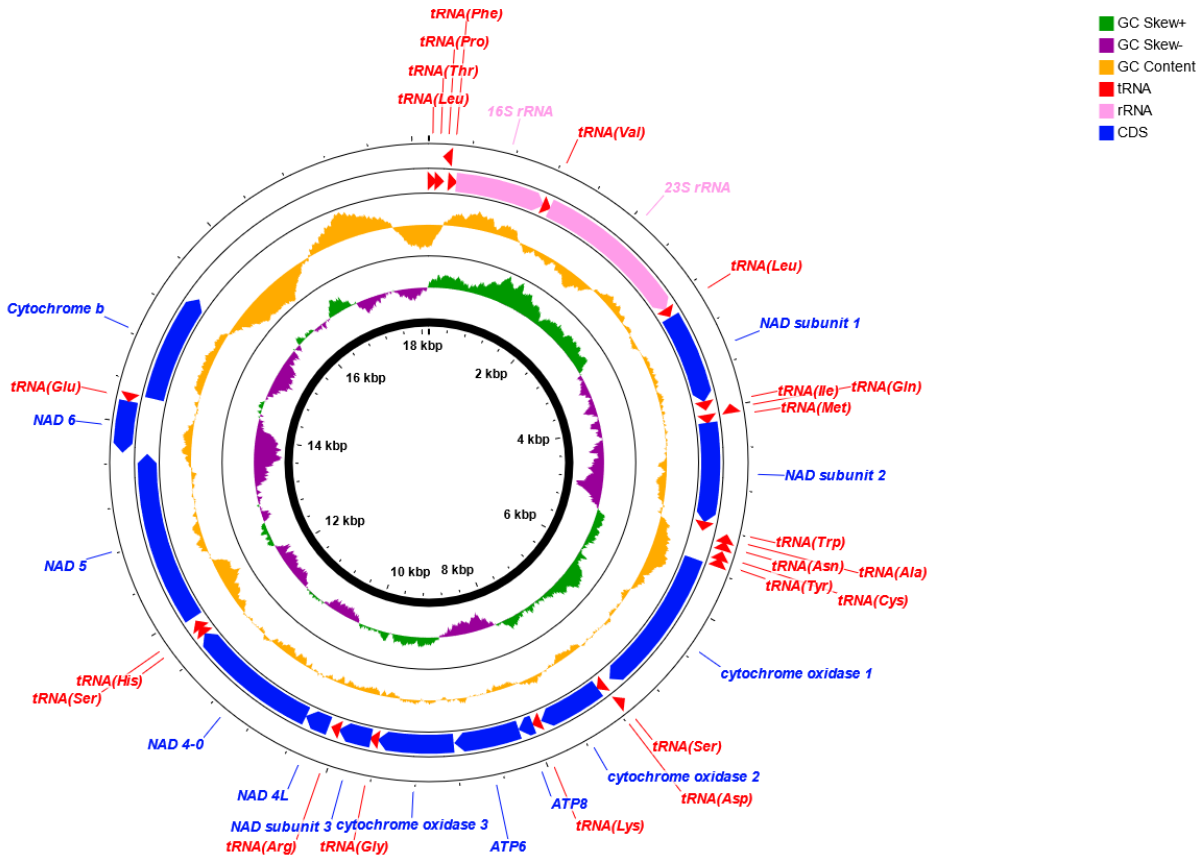


Figure 3: Maximum-likelihood phylogenetic tree of the complete mitochondrial genomes of 203 Anuran species downloaded from the NCBI BioProject public database. Five non-Anuran species were used as outgroup. The dendrogram was constructed using IQ-Tree multicore version 2.2.2.6. Branch lengths represent genetic distances. Ultrafast bootstrap and SH-aLRT branch tests were performed with 10,000 replications. Branches with SH-aLRT values ≤ 80% or UFboot values ≤ 95% were labelled with red dots. The leftmost and rightmost panels display zoomed upper and lower parts of the dendrogram, respectively. Cane toad (*Rhinella marina*) labelled with a red arrow.

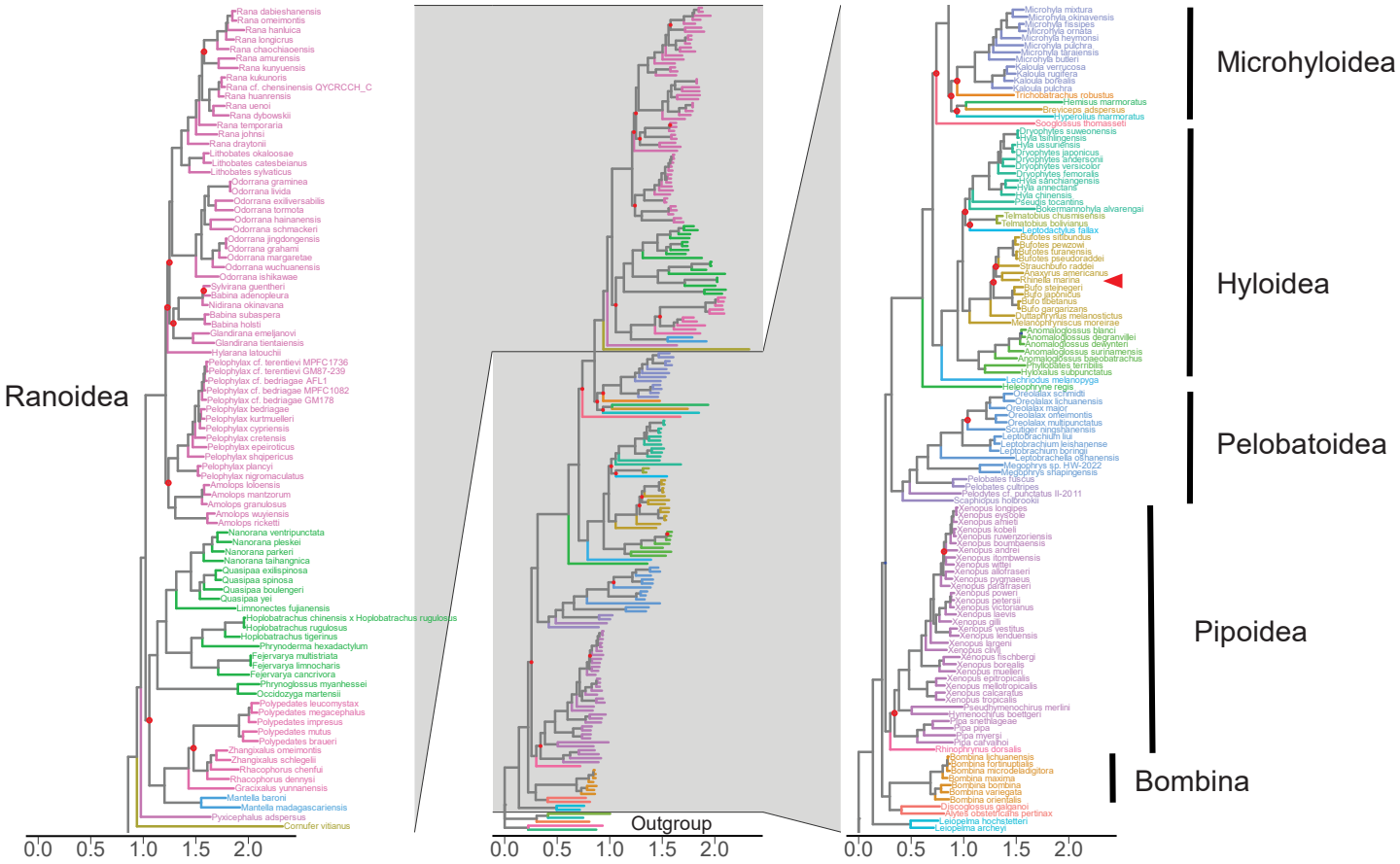
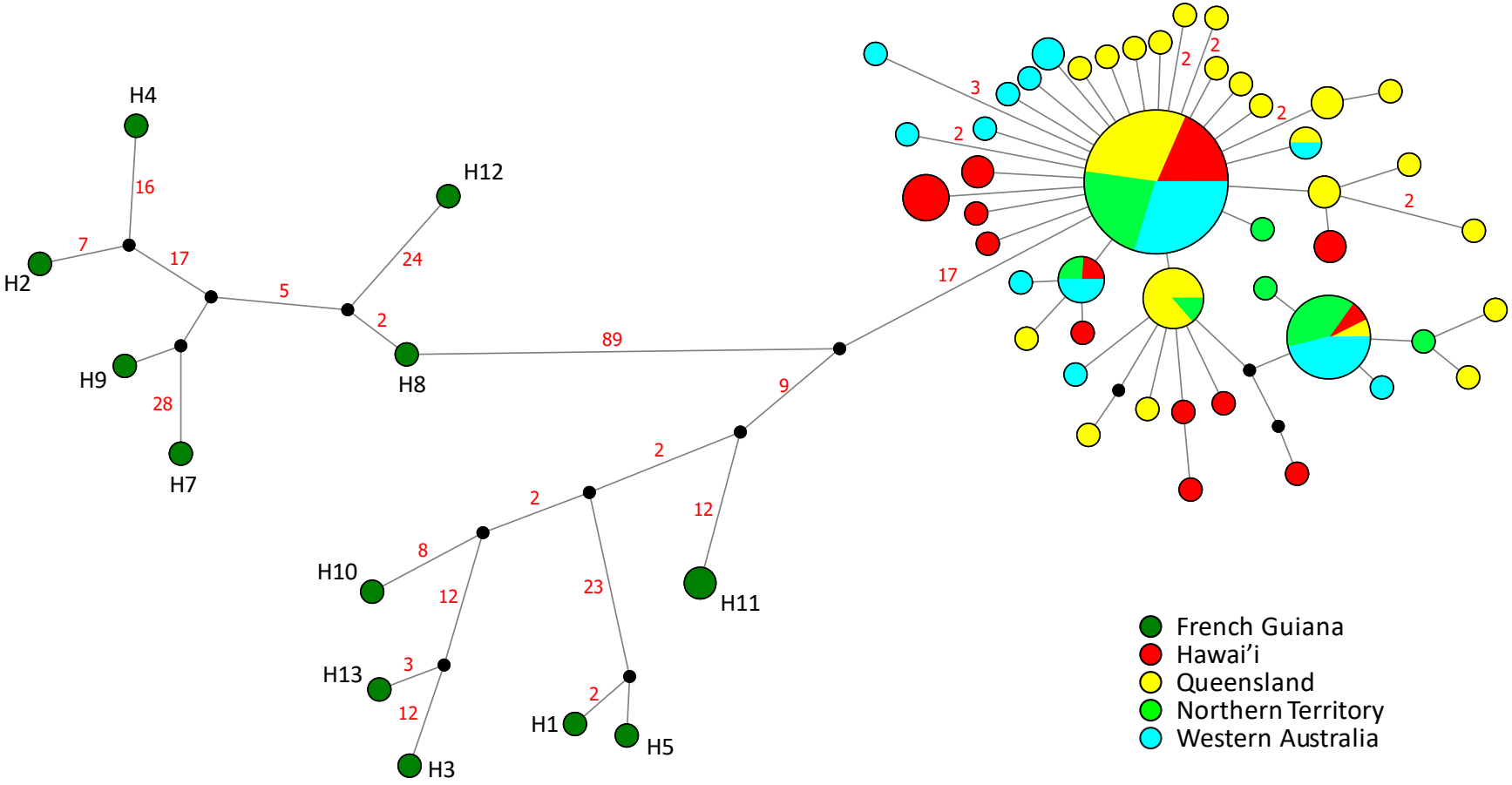
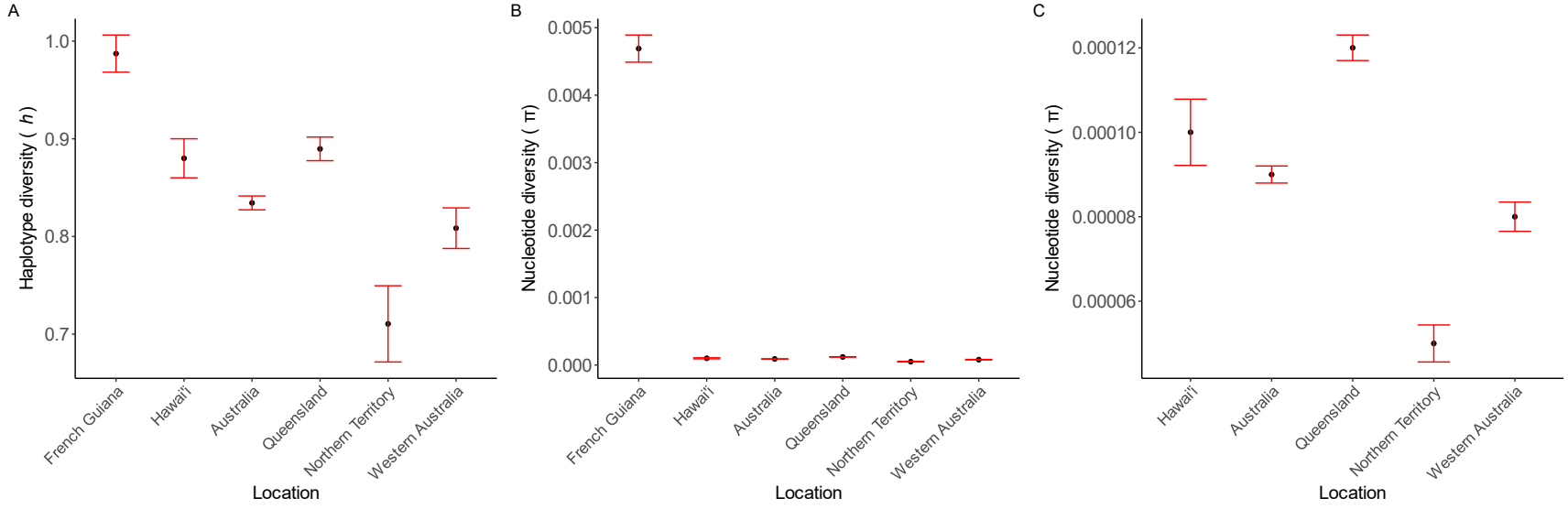


Figure 4: The haplotype network illustrating the relationships among 132 individuals from French Guinea, Hawai'i, and Australia based on near-complete mitochondrial genomes. Each sample location is represented by a different colour. A black solid circle represents a putative haplotype that was not directly sampled. The red number adjacent to a line indicates the number of nucleotide substitutions between haplotypes, while the absence of a number indicates a single nucleotide substitution.

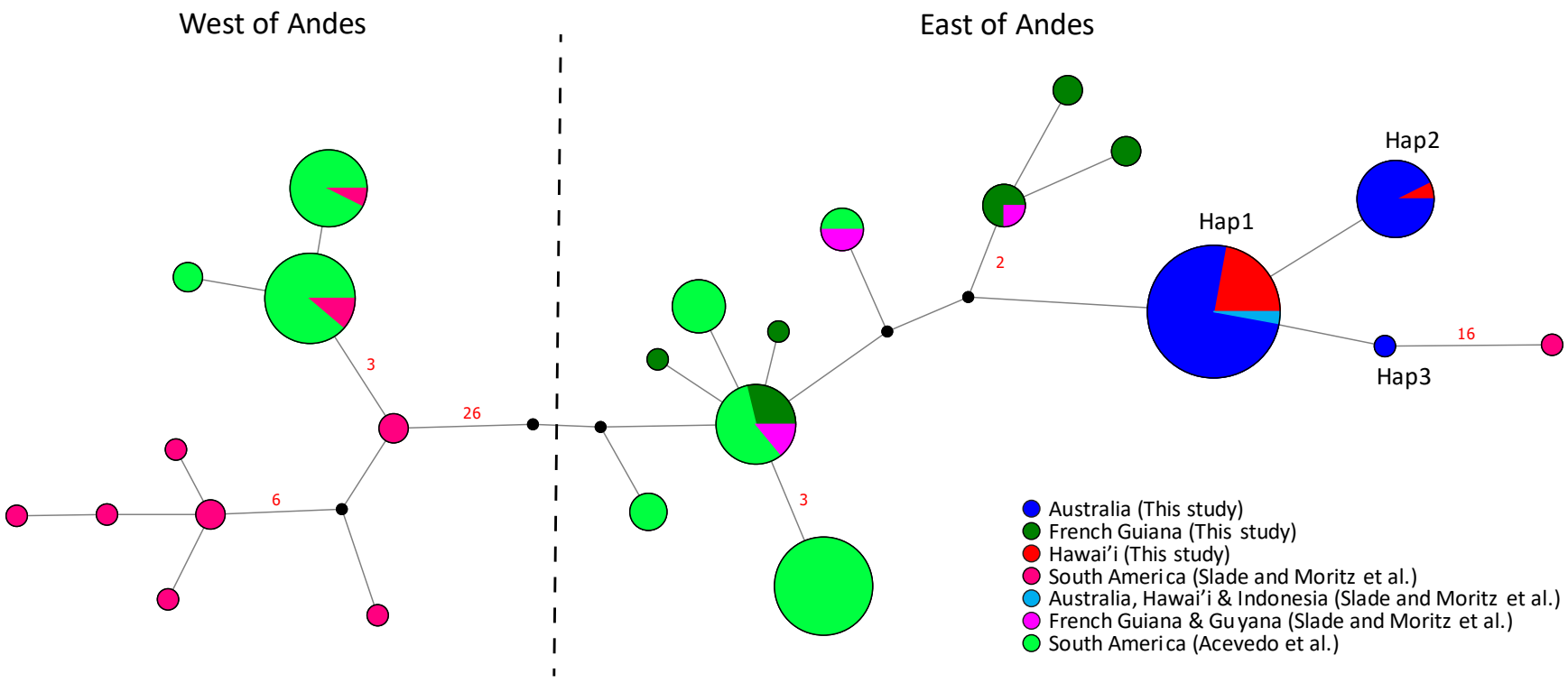


934 Figure 5: Haplotype diversity (A) and nucleotide diversities (B and C) of native population and introduced populations.



935

Figure 6: Haplotype network of a 510 bp segment of ND3 with flanking tRNAs, a total of 223 samples, including samples from this study, Slade and Moritz (1998) and Acevedo et al. (2016). Sample locations are represented by different colours. A black solid circle represents a putative haplotype that was not directly sampled. The red number adjacent to a line indicates the number of nucleotide substitutions between haplotypes, while the absence of a number indicates a single nucleotide substitution. Samples are separated by geographical locations on each side of the Andes Mountains in South America (Acevedo et al., 2016). Toads on the west side of the Andes are now known as *Rhinella horribilis*.



944 Table 1: Population genetics statistics of native range population (French Guiana) and introduced populations (Hawai'i and Australia).

945

	N	**Polymorphic sites**	***H***	**π ± SD**	***h* ± SD**	**HR**
French Guiana	13	205	12	0.00469 ± 0.00037	0.9872 ± 0.035	12
Hawai'i	25	11	13	0.00010 ± 0.00002	0.8800 ± 0.051	10.65
Australia	94	40	36	0.00009 ± 0.00001	0.8344 ± 0.035	12.50
Queensland	42	27	23	0.00012 ± 0.00001	0.8897 ± 0.040	13.54
Northern Territory	20	5	7	0.00005 ± 0.00001	0.7105 ± 0.089	6.42
Western Australia	32	16	13	0.00008 ± 0.00001	0.8085 ± 0.06	9.63

946

947 Table 2: Analysis of molecular variance (AMOVA) of all populations.
948

Source of variation	*df*	Sum of squares	% of variation	Fixation indices
All populations				
Among groups	2	482.96	64.49	
Among populations within groups	2	2.91	-0.82	
Within populations	127	579.35	36.33	F_{ST} = 0.64***

949

Table 3: Population differentiation (pairwise *Fst*) between native range and introduced ranges (Hawai'i and Australia) (below the diagonal line) The associated *p*-values were shown above the diagonal line.

	French Guiana	**Hawai'i**	**Queensland**	**Northern Territory**	**Western Australia**	**Australia**
French Guiana		< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Hawai'i	0.584		0.056	0.01	0.01	0.02
Queensland	0.669	0.016		0.009	0.003	0.34
Northern Territory	0.551	0.054	0.041		0.72	0.54
Western Australia	0.628	0.033	0.030	-0.012		0.80
Australia	0.792	0.023	0.002	-0.004	-0.006	

953 Table 4: Demographic estimators of native range population (French Guiana) and introduced populations (Hawai'i and Australia).

			Demographic expansion		Spatial expansion	
	Fu's *Fs*	**R_2**	**SSD**	**Raggedness**	**SSD**	**Raggedness**
French Guiana	1.292 (p = 0.660)	0.184 (p = 0.891)	0.026 (p = 0.188)	0.033 (p = 0.392)	0.025 (p = 0.08)	0.033 (p = 0.546)
Hawai'i	-8.783 (p < 0.001)	0.075 (p = 0.02)	0.006 (p = 0.436)	0.080 (p = 0.242)	0.006 (p = 0.31)	0.080 (p = 0.208)
Australia	-27.923 (p < 0.001)	0.019 (p < 0.001)	0.002 (p = 0.302)	0.067 (p = 0.158)	0.002 (p = 0.2)	0.067 (p = 0.144)
Queensland	-23.037 (p < 0.001)	0.030 (p < 0.001)	0.001 (p = 0.702)	0.052 (p = 0.38)	0.002 (p = 0.66)	0.053 (p = 0.418)
Northern Territory	-3.908 (p = 0.0009)	0.096 (p = 0.072)	0.032 (p = 0.068)	0.208 (p = 0.032)	0.034 (p = 0.032)	0.208 (p = 0.046)
Western Australia	-8.950 (p < 0.001)	0.051 (p = 0.001)	0.003 (p = 0.576)	0.079 (p = 0.336)	0.003 (p= 0.348)	0.079 (p = 0.264)

954