広島大学学位請求論文

Molecular evolution of mitochondrial DNAs in rain frogs (Amphibia: Anura: Brevicipitidae: *Breviceps*): Discovery of the secondary largest vertebrate mitochondrial genome and the cause of genome size increase

（フクラガエル属における mtDNA の分子進化：脊椎動物で２番目に巨大なミトコンドリアゲノムの発見とその発生要因）

２０２０年
広島大学大学院理学研究科
生物科学専攻

逸見 敬太郎

Molecular evolution of mitochondrial DNAs in rain frogs (Amphibia: Anura: Brevicipitidae: *Breviceps*): Discovery of the secondary largest vertebrate mitochondrial genome and the cause of genome size increase

Keitaro Hemmi

Abstract

The mitochondrial (mt) genome of the bushveld rain frog (*Breviceps adspersus*, Brevicipitidae, Afrobatrachia) is the largest (28.8 kbp) among the vertebrates investigated to date. The major cause of genome size enlargement in this species is the duplication of multiple genomic regions. To investigate the evolutionary lineage, timing, and process of mt genome enlargement, I sequenced the complete mtDNAs of two congeneric rain frogs, *B. mossambicus* and *B. poweri*. The mt genomic organization, gene content, and gene arrangements of these two rain frogs are very similar to each other but differ from those of *B. adspersus*. The *B. mossambicus* mt genome (22.5 kbp) does not differ significantly from that of most other afrobatrachians. In contrast, the *B. poweri* mtDNA (28.1 kbp) is considerably larger: currently the second largest among vertebrates, after *B. adspersus.* The main causes of genome enlargement differ among *Breviceps* species. Unusual elongation (12.5 kbp) of the control region (CR), a single major noncoding region of the vertebrate mt genome, is responsible for the extremely large mt genome in *B. poweri*. Based on the current *Breviceps* phylogeny and estimated divergence age, it can be concluded that the genome enlargements occurred independently in each species lineage within relatively short periods. Furthermore, a high nucleotide substitution rate and relaxation of selective pressures, which are considered to be involved in changes in genome size, were also detected in afrobatrachian lineages. My results suggest that these factors were not direct causes but may have indirectly affected mt genome enlargements in *Breviceps*.

Introduction

Mitochondrial (mt) genomes of multicellular animals (metazoans) are generally closed-circular and double-stranded DNA molecules approximately 14–20 kbp in length (D. R. Wolstenholme, 1992; Pereira *et. al.*, 2010; Lavrov & Pett, 2016). However, genome size ranges from 6 to 48kbp (Pereira *et. al.*, 2010; Hou *et. al.*, 2016), and linear and fragmented mtDNAs have been found in approximately 9000 animals investigated thus far (Kayal *et. al.*, 2011; Song *et.al.*,2019; NCBI organelle genome resourcesAvailable from http://www.ncbi.nlm. http://nih.gov/genome/browse!/organelles/.). In most metazoans, this small organelle genome encodes a typical set of 37 genes: 13 protein genes, involved in the electron transport system of respiration, two rRNA genes (*rrns*), and 22 tRNA genes (*trns*). In addition, animal mt genomes contain one long noncoding region, harboring several sequence elements related to mt genomic transcription and replication, named the control region (CR) or alternatively the D-loop region (Pereira *et. al.*, 2010).

The mt genomic structure of metazoans, particularly vertebrates, tends to be conserved among closely related taxa. The same mt gene content and gene arrangement (synteny) are observed from fishes to mammals (e.g., (Boore,1999; Warzecha *et. al.*, 2018)). This genomic organization would have arisen in a common ancestor of vertebrates and has been maintained in a wide range of vertebrate taxa over 400 million years (Kurabayashi *et. al.*, 2006; Kumar *et. al.*, 2017).

However, in some vertebrate taxa, increases in gene content due to gene duplication and rearrangements of mt gene positions are often observed (e.g., (Boore,1999; Mueller & Boore, 2005)). In particular, a greater degree of structural change in mt genomes has been reported for amphibians, especially modern anurans (neobatrachians) (e.g., (Kurabayashi *et. al.*, 2010; Kakehashi *et. al.*, 2013)). Among anurans, the members of the basal group (a paraphyletic group generally called the "Archaeobatrachia") retain the typical ancestral (i.e., symplesiomorphic) mt genome organization of vertebrates (Irisarri *et. al.*, 2012; Zhang *et. al.*, 2013). In contrast, most of the

neobatrachians share the derived positions of three *trn*s translocated from their original locations (LTPF-*trn* cluster) (Irisarri *et. al.*, 2012; Sumida *et. al.*, 2001). The Neobatrachia contains two superfamilies, Hyloidea and Ranoidea. The latter includes three large groups, Microhylidae, Natatanura, and Afrobatrachia (Frost *et. al.*, 2006; Feng *et. al.*, 2017). Of these, microhylid frogs retain the mt genomic structure of typical neobatrachian, while in natatanuran and afrobatrachian taxa, mt genomic rearrangements accompanying duplications and translocations of mt genes and the CR are often observed (Kurabayashi *et. al.*, 2010; Kakehashi *et. al.*, 2013; Zhang *et. al.*, 2013; Kurabayashi *et. al.*, 2008; Xia *et. al.*, 2014). In particular, the mt genomes of afrobatrachians show largescale structural changes.

The Afrobatrachia is a group of frogs, endemic to Africa, which currently consists of 422 described species in four families: Arthroleptidae, Brevicipitidae, Hemisotidae, and Hyperoliidae (Frost *et. al.*, 2019). Kurabayashi & Sumida (2013) have sequenced the complete mt genomes of four afrobatrachians representing all four afrobatrachian families. All of these frogs tend to have large mt genomes, exceeding 20kbp, and three of them have duplicated mt genes and/or CRs (excluding *Hemisus*). In particular, the mt genome of *Breviceps adspersus* (Brevicipitidae) has a highly reconstructed mt genome with many gene rearrangements and many duplicated gene regions. Consequently, the size of the *B. adspersus* mt genome is the largest among those vertebrates for which mtDNA has been sequenced to date (also the 13th largest among metazoans) (Kurabayashi & Sumida, 2013).

To investigate the evolutionary origin and significance of the unusually large mt genome of *B. adspersus*, I analyzed the mt genomes of two additional *Breviceps* species, *B. mossambicus* and *B. poweri*. It has been suggested that the duplication of mt genes and the CRs, which results in an increase in genome size, are the result of nonadaptive evolution which, in insects, is correlated with an acceleration of nucleotide substitution rate (Shao *et. al.*, 2003) and a relaxation of purifying selective pressure, leading to a reduction in functional constraints that purge slightly deleterious mutations (Boussau *et. al.*, 2011). Thus, I investigated the changes in the substitution rate and selection pressure

among afrobatrachian mt genomes and discuss the correlation between mt genome size and the change in the substitution rate and selective pressure.

Materials and Methods

2.1. Specimens Used.

In this study, I used four frog specimens: one Mozambique rain frog, *Breviceps mossambicus* (Peters, 1854), and three Power's rain frogs, *B. poweri* (Parker, 1934) (Peters, 1854; Parker *et. al.*, 1934). These frogs were obtained via the pet trade; thus, their collection sites are unknown. The frog specimens were stored in 95.5% ethanol as part of AK's personal amphibian collection. The voucher numbers are 15-004, 15008, and 15-010 for *B. poweri* and 14-001 for *B. mossambicus*.

2.2. Molecular Experiments.

I extracted and purified the total DNA from the liver tissue of each fixed specimen with a DNeasy Blood & Tissue Kit (QIAGEN K. K., Tokyo, Japan) or using phenol/chloroform extraction with "DNA sui-sui" extraction buffer (Rizo Inc., Tsukuba, Japan) and ethanol precipitation methods (Sambrook & Russell, 2001). From the purified total DNA, the whole mtDNA was amplified by PCR and sequenced for *B. poweri* (voucher 15-004) and *B. mossambicus*. The PCR amplification and sequencing procedures from Kurabayashi *et. al.* (2006) were followed. Specifically, for each specimen, I amplified 10 overlapping fragments containing the entire mt genome using the long and accurate (LA) PCR method with LATaq (Takara Bio Inc., Shiga, Japan) according to the manufacturer's instructions. These fragments were purified using the gel extraction method with a GenElute Agarose Spin Column (Sigma-Aldrich Japan Inc., Tokyo, Japan). The purified PCR fragments were sequenced using the primer walking method with an ABI 3130xl automated DNA sequencer (Applied Biosystems, Foster City, CA, USA). In this study, I used a total of 117 PCR primers for mtDNA amplification and/or sequencing, and 108 of them were newly designed during this study. All the primers used are listed in Table 1.

2.3. NGS.

The PCR fragments containing the CRs were very long, and these fragments harbored many direct repeat sequences that could not be read by the primer walking method (see Figure 1). Thus, these fragments were sequenced using single-molecule real-time (SMRT) sequencing with the PacBio RS II next-generation sequencer (NGS), which allows exceptionally long read sequencing (max lengthperread > 40kbp) (Pacific Bioscience, "SMRT Link user guide (v5.1.0)," 2018, Available from https: //http://www.pacb.com/ documentation/smrt-link-user-guide-v5-1-0/.). I also applied the multiplex-amplicon approach (see (Pacific Bioscience, 2018)). Briefly, I amplified the CR fragments from three *B. poweri* and one *B. mossambicus* specimens (approx. 13 and 8kbp, respectively; see Figure 2) with the primers having distinct 3′ tag sequences for each specimen (consequently, the PCR fragments amplified from the same specimen have the same specific tag sequences, and thus, the PCR results were sortable from the mass NGS output-multiplex-amplicon method). Approximately 5μg of the gel-purified CR fragments of two *B. poweri* (vouchers 15-004 and 15-010) and one *B. mossambicus* was used for the library construction for NGS. I outsourced the library construction and SMRT sequencing to CoMIT (Center of Medical Innovation and Translational Research) of Osaka University. The first SMRT run allowed for the determination of the entire CR sequences of *B. mossambicus* and one *B. poweri* specimen (15-010). However, I could not obtain enough sequence reads for another *B. poweri* specimen (15-004). Thus, for this specimen, I made two internal PCR primers (named bfCSB_ Fow1 and bfCSB_Rev2, Table 1) based on the resultant CR sequence of *B. poweri* (15-010). Using these primers, two fragments of the 5′ and 3′ sides of the CR (approx. 4 and 9 kbp, respectively) were separately amplified from *B. poweri* (15-004) and the fragments were sequenced by another SMRT run. The assembled sequences of each CR fragment were reconstructed from the RS II outputs using the Long Amplicon Analysis program implemented in the SMRT Link analysis system (Pacific Bioscience, 2018).

The assembled whole mtDNA sequences of *B. mossambicus* and *B. poweri* (15-004) and the CR sequence of *B. poweri* (15-010) were deposited in the

International Nucleotide Sequence Database Collaboration (INSDC) under the accession numbers LC498571, LC498572, and LC498573, respectively.

2.4. Phylogenetic Analyses and Divergence Time Estimation.

I performed phylogenetic tree reconstructions and a divergence time estimation by adding the sequence data obtained in this study to the dataset of Kurabayashi and Sumida (Kurabayashi & Sumida, 2013), and the analytical methods used by Irisarri *et. al.*, (2012) were followed. The previous dataset includes not only mt gene sequences but also nuclear genes. The nuclear data were also used in this study. Mitochondrial sequence data of afrobatrachians reported by Zhang *et. al.* (2013) were not used here because of a lack of genome size information and sequences of some mt genes. The genes used and their accession numbers are listed in Table 2.

First, I aligned the sequences of each of the 13 protein genes, two *rrn*s, and 22 *trn*s, separately. The protein and RNA genes were aligned using the TranslatorX program with the default setting (Abascal *et. al.*, 2010) and MAFFT with the L-INS-i option (Katoh & Toh, 2008), respectively. The sequences of *trn*s were manually aligned using their secondary structures as a guide. Ambiguously aligned sites were deleted using the Gblocks program ver. 0.91b with the default parameters (Castresana, 2000). The final alignment dataset consisted of 21,063 bp (consisting of 13,938 and 7125 mt and nuclear gene sequences, respectively) from 49 operational taxonomic units (OTUs).

I used the following sequences in the phylogenetic analyses: two mt *rrn*s, 22 mt *trn*s, and the 1st and 2nd codons of the 13 mt and nine nuclear protein genes (total 15,093 bp). I did not use the 3rd codon positions of the protein genes in the phylogenetic analyses (and divergence time estimation) because it is known that their fast substitution rates could distort the reconstruction of deep anuran phylogenies (Hoegg *et. al.*, 2004; Gissi *et. al.*, 2006). I also applied a partitioning strategy in the phylogenetic analyses; i.e., the concatenated sequence data were partitioned into statistically suitable partitions and a distinct nucleotide substitution model was applied for each sequence partition. The PartitionFinder program ver. 1.1.1 (Lanfear *et. al.*, 2012) was used to estimate the best partitioning scheme using Akaike's information criterion

(AIC) (Akaike, 1973). Consequently, a setting with 17 distinct partitions was suggested as the best partitioning scheme and this partitioning scheme was used in both the maximum likelihood (ML) and Bayesian inference (BI) analyses. An independent general time reversible+gamma distribution (GTR+G) substitution model for each of the 17 partitions was applied in the ML analysis. For the BI analysis, the best substitution model was estimated for each partition using the Kakusan4 program (Tanabe, 2011). The detailed partitioning scheme and the suggested substitution models in BI are summarized in Data S1.

I performed phylogenetic reconstructions using ML and BI methods. RAxML ver. 8.2.12 and MrBayes ver. 3.2.6 software packages were used for the ML and BI analyses, respectively (Stamatakis *et. al.*, 2014; Ronquist *et. al.*, 2012). The rapid hill climbing algorithm (implemented in RAxML) with the starting tree option of 100 randomized parsimonious trees was applied for the ML analysis. For the BI analysis, 10 million generations of four Markov chains (MCs) were run with one sampling per every 1000 generations and the 1st 10% samples were discarded as burn-in. The convergence of the posterior distribution of model parameters (all parameters reached >200) was checked using Tracer ver. 1.5 (Rambaut & Drummond, 2009). The supports for the internal branches of reconstructed trees were evaluated using bootstrap percentages (BPs) calculated by 1000 pseudoreplications and Bayesian postprobabilities (BPPs) in ML and BI analyses, respectively.

The divergence times of anurans were estimated using a Bayesian dating method with the BEAST ver. 2.5.2 program package (Bouckaert *et. al.*, 2019). In the estimation, the amphibian phylogenies recovered from both ML and BI analyses were used as the topology constraint (Figure 3). The sauropsid monophyly (lizard+bird), not recovered by my ML and BI analyses, was *a priori* assumed in this dating analysis as in previous studies (Irisarri *et. al.,* 2012; Kurabayashi & Sumida, 2013). The same data partitioning scheme and substitution models used in the BI analysis were also applied here. I used the Yule process (Yule, 1925) to describe cladogenesis. The final MCs were run twice for 100 million generations with one sampling per every 10,000 generations, and the 1st one million generations were discarded as burn-in. The

posterior distributions of model parameters were checked in the same way as the above BI analysis. Following Irisarri *et. al.* (Irisarri *et. al.*, 2012), I applied seven (lower age boundaries) calibration points as follows: (A) Sauropsida-Synapsida split: >312 million years ago (Ma), (B) Archosauromorpha-Lepidosauromorpha split: >260 Ma, (C) Cryptobranchidae-Hynobiidae split: >146Ma, (D) Anura-Caudata split: >249Ma, (E) most recent common ancestor (MCA) of Discoglossoidea: >161 Ma, (F) MCA of Pipoidea: >146Ma, and (G) Calyptocephalella-Lechriodus split: >53 Ma. These were used as prior boundaries for divergence time estimation.

2.5. Relative Rate Test.

I compared the relative rates of nucleotide substitutions of mt genes (all mt genes, all mt protein-coding genes, all *rrn*s, and all *trn*s) among afrobatrachian lineages using relative rate tests (RRTs) (Li & Bousquet, 1992) with the RRTree program (Bouckaert *et. al.*, 2019). The Kimura two-parameter substitution model (Kimura, 1980) was used for the estimation of genetic distances. In this analysis, I used the gene data of 24 neobatrachians. Noncompared lineages were used as the outgroups in each comparison (e.g., when I compared *Hemisus* and the three *Breviceps* species, the lineages of the remaining 20 neobatrachians were regarded as the outgroups). The lineages for each comparison are shown in Table 3.

2.6. Detection of Changes in Lineage-Specific Selective Pressure.

It is known that the ratio of nonsynonymous/synonymous substitutions (dN/dS ratio $= \omega$) can be used to identify the changes in selective pressure among taxa (Irisarri *et. al.*, 2012; Yang, 2006; Buschiazzo *et. al.*, 2012). To understand the changes in the dN/dS ratio in ranoid lineages, I used the "branch model" analysis (Yang, 1997) using the Codeml program implemented in PAML 4.9 (Yang, 2007). In this analysis, I used the alignment data of mt protein genes from 24 neobatrachian taxa and the ML and BI tree topology. I compared five branch models with distinct assumptions about dN/dS ratios. In one model (null model), the single constant ω was assumed in all neobatrachian lineages while other models allowed the changes in the dN/dS ratios on specific ranoid

branches. The details of the models are shown in Table 4. The branch lengths were first calculated under the null model, and the resultant branch lengths were applied in the other models. The significance of the differences in log likelihoods among these models was tested using the likelihood ratio test (LRT).**<0.001.

Results and Discussion

3.1. Huge mt Genomes of Rain Frogs.

 I determined the entire mt DNA sequences of two rain frogs, *B. mossambicus* and *B. poweri*. These mt genomes are 22,553 and 28,059 bp in length, which are very large for a vertebrate mt genome (typical size range is 16–17 kbp (Wolstenholme, 1992)). Although the genome size of *B. mossambicus* is similar to those of other afrobatrachian frogs (e.g., marbled reed frog (*Hyperolius marmoratus*): 22,595 bp and hairy frog (*Trichobatrachus robustus*): 21,418 bp; see Figure 3), the *B. poweri* mt genome was the second largest among the vertebrates investigated thus far (according to NCBI organelle genome resources (NCBI organelle genome) as of August 2019). The third largest vertebrate mt genome is 25,972 bp for the prickly gecko (*Heteronotia binoei*) (Fujita *et. al.*, 2007). Therefore, the mt genome of *B. poweri* is closer in size to that of *B. adspersus*, which possesses the largest known mt genome of vertebrates (28,757bp).

In contrast to the large genome sizes, the mt gene content of both *B. mossambicus* and *B. poweri* is similar to that of many other vertebrates (Figure 3), containing the set of 37 single genes (13 proteins, two rRNA, and 22 tRNA genes) and single long and short noncoding regions commonly found in vertebrate mt genomes. The long noncoding region is referred to as CR and contains putative sequence elements involved in DNA replication and RNA transcription (e.g., terminate associate sequence (TAS) and conserved sequence blocks 1–3 (CSB1–3); see Figure 1 and (Jemt *et. al.*, 2015)),

LRT. However, model 2 had a higher −ln *L*.

and a short noncoding region is known as the *light*-strand replication origin (O_L). The presence of a small pseudogene, *trn*S(AGY), between NADH dehydrogenase subunits 4 and 5 (*nd4* and *nd5*) is the singular exception in gene content (Figure 3).

The gene content of *B. mossambicus* and *B. poweri* is not similar to that of the congeneric *B. adspersus* mt genome, which has many duplicated and triplicated gene regions and duplicated CRs, making it the largest known mt genome among vertebrates (Figure 3). Instead, the cause of the mt genome enlargement in *B. mossambicus* and *B. poweri* is the unusual expansions of the control regions. The lengths of their CRs are 6,618 and 12,537 bps, respectively, although the typical CR lengths in vertebrates are 1-2kbp (Wolstenholme, 1992). The long CRs of *B. poweri* and *B. mossambicus* result from the occurrence of multiple and long-tandem repeats (Figure 1). The *B. mossambicus* CR contains four distinct direct tandem repeats, and these repetitive sequences, totaling 4228 bp in length, occupy 63.9% of the CR. The CR of *B. poweri* (individual 4) has six distinct repeat sequence groups. Two of them are quite long (see Figure 1): one repeat group consists of 1,150 and 1,151bp, nearly complete, repeat sequences (repeat 1: totaling 2301 bp), while another is composed of 23 units of 233–405bp incomplete direct repeat sequences (repeat 3: totaling 7,339bp). Consequently, the total length of the six repeat sequences in the *B. poweri* CR is 10,625 bp (occupying 84.7% of the CR nucleotides). The long CR over 10 kbp in length is quite rare and has never been completely sequenced in other vertebrate taxa (approximately 12kbp CR occurred in a Malagasy frog (*Gephyromantis pseudoasper*) (Kurabayashi *et. al.*, 2008)).

To determine whether the unusually long CR is specific to the individual frogs used or is a common characteristic of the frog species, I checked the CRs of two additional *B. poweri* individuals. The PCR fragments having similar lengths (approx. 13 kbp, including the whole CR and 5′ and 3′ franking gene regions) are commonly amplified in all three individuals (Figure 2). I also sequenced the CR of one of the two additional individuals. Although the CR lengths differ by 627bp between the *B. poweri* individuals (12,537 bp vs. 11,910 bp), the CR sequences of the two *B. poweri* individuals are quite similar, excluding the

number of repeat units. Therefore, the very long CR, with numerous repetitive sequences, seems to be a common feature of this rain frog species.

It is almost impossible to precisely sequence long repetitive regions exceeding 10 kbp, such as those of the *B. poweri* CR by conventional sequencing methods (i.e., primer walking and construction of deletion mutants (Kurabayashi *et. al.*, 2008)) or by using the NGS technique with a short-read sequencing strategy (Mahmoud *et. al.*, 2019). In this study, a long-read strategy with PacBio RS II allowed me to relatively easily sequence such long repetitive sequences. These results demonstrate the superiority and necessity of long-read sequencing in analyzing the repetitive sequences occasionally found in both organelles and nuclear genomes.

3.2. Phylogeny, Timing, and Distinct Causes of mt Genome Enlargement in These Congeneric Rain Frogs.

To infer the phylogenetic lineages and evolutionary times of the mt genome enlargement of afrobatrachians, including *Breviceps*, I performed molecular phylogenetic analyses and divergence time estimation. The ML and BI methods reconstructed the same tree topology, and the maximum log likelihood (ln *L*) of the resultant ML and the average ln *L* of the BI tree were -224667.83 and -221797.17, respectively. The time tree of amphibians having the ML/BI tree topology is shown in Figure 4. The major anuran phylogenies and the split ages agreed completely or to a substantial degree with those from previous studies: e.g., the monophyly of neobatrachians and paraphyly of archaeobatrachians with respect to neobatrachians (most recent common ancestor of Neobatrachia = 162.3Ma), monophyly of ranoids (MCA = 127.0Ma), and the three major clades of ranoids (Natatanura, Microhylidae, and Afrobatrachia, which started to split from 127.0 Ma) (e.g., (Irisarri *et. al.*, 2012; Zhang *et. al.*, 2013; Feng *et. al.*, 2017; Pyron & Wiens, 2011)).

The resultant phylogenies of afrobatrachian taxa completely match those of recent studies (Zhang *et. al.*, 2013; Kurabayashi & Sumida, 2013; Nielsen *et. al.*, 2018) (Figure 5), excluding the Afrobatrachia+Microhylidae clade not supported in Feng *et. al.* (Feng *et. al.*, 2017). Specifically, (1) Afrobatrachia is monophyletic, (2) Microhylidae becomes the sister taxon of Afrobatrachia, (3)

13

Brevicipitidae forms a clade with Hemisotidae (this clade is called Xenosyneunitanura axis is in million years. (Frost *et. al.*, 2006)), (4) *Breviceps* is monophyletic and *B. poweri* is the basal taxon of the three species sampled in this genus, and (5) *B. mossambicus* and *B. adspersus* have a closer affinity within this sample. The estimated divergence ages of the corresponding nodes are as follows: (1) 118.1Ma for the divergence between Afrobatrachia and Microhylidae, (2) 107.9 Ma for the most recent common ancestor of afrobatrachians, (3) 87.4Ma for the split of Brevicipitidae and Hemisotidae, (4) 47.1 Ma for the divergence of *B. poweri* from the lineage of *B. mossambicus*+*B. adspersus*, and (5) 34.0Ma for the split of *B. mossambicus* and *B. adspersus*. The divergence ages estimated here largely agree with those from recent studies (Zhang *et. al.*, 2013; Kurabayashi & Sumida, 2013). However, for the *Breviceps* divergences, younger ages have been estimated by Nielsen *et. al.* (Nielsen *et. al.*, 2018) (20 and 18 Ma for the divergences of 4 and 5, respectively). They used a total of 24 *Breviceps* taxa and applied some young calibration points within afrobatrachian lineages in their dating analyses. These differences in the data analysis could account for the differences in estimated divergence ages between my study and that of Nielsen *et. al.* (Nielsen *et. al.*, 2018).

As described above, the mt genomes of microhylid frogs, the sister group of Afrobatrachia, exhibit average genome sizes for vertebrates (16.7–17.2 kbp) (Zhang *et. al.*, 2005; Hwang & Lee, 2012). In contrast, the afrobatrachian mt genomes exceed 20 kbp for all six species examined. Thus, the mt genome enlargement appears to be an evolutionary trend that has arisen in the lineages leading to the living afrobatrachians after the split from the highlighted by orange and red colors, respectively.

microhylid lineage (Figures 4 and 5). In particular, the mt genome sizes of the two *Breviceps* species, *B. adspersus* and *B. poweri*, are over 28 kbp, making them as the largest and second largest known mt genomes of vertebrates. With regard to the interspecific phylogeny of *Breviceps*, these two species are not monophyletic; i.e., *B. adspersus* is closer to *B. mossambicus* than to *B. poweri* (Figure 3). Overall, my results indicate that the huge mt genomes arose in two

independent rain frog lineages within relatively short periods (>47 and >30 million years for the lineages leading to *B. poweri* and *B. adspersus*, respectively). Furthermore, it is noteworthy that the physical causes of mt genome enlargement differ between these two rain frogs. The duplication of multiple gene/CRs and the accumulation of repetitive sequences are the main causes of mt genome enlargement in *B. adspersus* and *B. poweri*, respectively.

There are several examples of mt genome enlargement shared by congeneric species. For example, Malagasy poison frogs (*Mantella* spp.) commonly have >22 kbp mt genomes enlarged by the duplication of genes and CRs (Kurabayashi *et. al.*, 2008; Boussau *et. al.*, 2011), and two *Scapharca* ark shells exhibit >46kbp mt genomes, mainly caused by elongations (>30kbp) of noncoding sequences (Hou *et. al.*, 2016; Liu *et. al.*, 2013). However, there has been no similar example to that of *Breviceps* where the mt genome enlargement occurs in independent congeneric lineages, from distinct causes. Thus, the mt genomic structure is highly variable in this frog taxon.

3.3. Gene Rearrangements and Evolutionary Inference.

Although a highly rearranged mt genome is present in *B. adspersus* (Kurabayashi & Sumida, 2013), the mt gene content and their arrangements of the congeneric *B. poweri* and *B. mossambicus* do not deviate largely from the typical mt genomes of vertebrates (Figure 3). The latter *Breviceps* species have almost the same gene arrangement, with the exception of two small translocations of *trn*N and light-strand replication origin (O_L) within the WN-O_L-ACY *trn* cluster (Figure 3). Because the order of *trn*sWN-O_L found in *B. poweri* (and also in *B. adspersus*) is a primitive (plesiomorphic) gene order commonly shared by other afrobatrachians (Figure 3) (Zhang *et. al.*, 2013; Kurabayashi & Sumida, 2013), the O_L-*trn*sNW order of *B. mossambicus* appears to be newly emerged in the lineage leading to this species.

The *B. poweri* and *B. mossambicus* mt gene arrangements are similar to those of the mt gene orders of neobatrachians, especially *Hemisus*. Excluding the two minor *trn* translocations that occurred within the same *trn* clusters (specifically O_L-*trn*sNW in *B. mossambicus* and *trn*sPLTF in *Hemisus*), the translocation of *trn*sHS is the only distinctive difference between *Breviceps* and *Hemisus*.

Although the *trns*HS is located between *nd4* and *nd5* in most neobatrachians, including *Hemisus*, this *trn* block lies between the cytochrome apoenzyme *b* gene (*cytb*) and CR in *B. poweri* and *B. mossambicus*. The novel *cytb-trns*HS-CR arrangement could have arisen in a common ancestor of *Breviceps* after the split with other brevicipitid genera, for the following reasons: (1) the original *nd4-trns*HS-*nd5* arrangement remains in another brevicipitid, *Callulina kreffti*, mt genome (Zhang *et. al.*, 2013) and (2) the pseudogene of *trn*S is found at its original position between *nd4* and *nd5* in *B. poweri* and *B. mossambicus* (Figure 3).

The gene rearrangements in animal mt genomes are considered to reflect animal evolution (e.g., (Kurabayashi & Sumida, 2013; Boore *et. al.*, 1995; Boore *et. al.*, 1998; Zhang *et. al.*, 2018)), and the rearranged gene orders found in this study can be regarded as novel phylogenetic markers for brevicipitid taxa. Specifically, the O_L-*trns*NW can be considered characteristic of some members of the *B. mossambicus* group (Nielsen *et. al.*, 2018), while the *cytbtrns*HS-CR can be regarded as a synapomorphy of *Breviceps*.

In the *B. adspersus* mt genome, the *trns*HS block is further translocated and positioned within the triplicated *trns*LTPF-12S*rrn-trn*V-16S*rrn* cluster. In addition, duplications and translocations of the *trns*WN-OL cluster and CR were also found in this species (Figure 3) (Kurabayashi & Sumida, 2013). The detailed phylogenetic lineage and evolutionary period of these large genomic modifications have not been clarified. This study shows that the mt genomes of *B. poweri* and *B. mossambicus* are not markedly rearranged from the typical neobatrachian type, although *B. adspersus* branches between these species' lineages (Figure 3 and see (Nielsen *et. al.*, 2018)). Therefore, it can be concluded that the unique mt genome of *B. adspersus* evolved only in the lineage leading to this species. Large mt genomic modifications resulting in stepwise gene rearrangements along with multiple lineage splits have been reported in Malagasy frogs (Mantellidae) (Kurabayashi *et. al.*, 2008). However, in *Breviceps*, largescale genomic changes occurred in a single species lineage in a relatively short period (>47 Ma in this study and >20 Ma according to (Nielsen *et. al.*, 2018)), suggesting that the genomic structure of vertebrate mtDNA

could harbor higher structural variability than is generally believed (e.g., (Warzecha *et. al.*, 2018)).

3.4. Substitution Rates and Selective Pressure on the Afrobatrachian mt Genomes and Their Correlations with Genome Size Expansion.

As mentioned above, the duplication of genes/CRs and the occurrence of numerous repetitive sequences in the CR are responsible for the huge mt genomes of *B. adspersus* and *B. poweri*, respectively. Most of these events occurred in tandem, except for the nontandem duplications of CR and *trn*S(AGY) (Figure 3), which could have been caused by a nontandem duplication mechanism (Kurabayashi *et. al.*, 2008; Kurabayashi & Sumida, 2013). In animal mt genomes, it is believed that both duplicated gene regions and tandem repeats in the CR have emerged because of errors in mtDNA replication, such as imprecise replication termination and strand slippage of the nascent DNA strand (e.g., (Moritz & Brown, 1987; Xia *et. al.*, 2016)). Furthermore, it has been suggested that accelerated nucleotide changes lead to the frequent tandem duplications via frequent substitutions at the initiation and termination points of mtDNA replication (Shao *et. al.*, 2003). To investigate the correlation between genome size and nucleotide substitution rates, I compared the relative substitution rates of mt genes (four categories: all 37 genes, 13 protein genes, two *rrn*s, and 22 *trn*s) among afrobatrachian-related taxa using the RRT (Table 3).

The RRT showed that the substitution rates of afrobatrachian mt genes commonly having large genome sizes (>20 kbp) are significantly faster than those of the sister taxon, the microhylids, with normal genome sizes in all compared categories. Furthermore, the substitution rates of xenosyneunitanurans (Brevicipitidae+Hemisotidae), including the longest mt genomic species, are significantly higher than those of another afrobatrachian group, the laurentobatrachians (Arthroleptidae+Hyperoliidae) for all mt genes and protein genes (but are not significant for *rrn*s and *trn*s). These results suggest a correlation between the substitution rate and mt genome size. However, the substitution rates of the *Hemisus* mtDNA with a 20 kbp genome size are not significant but are faster than those of *Breviceps* species in all four

compared categories. Among *Breviceps* species, the *B. poweri* mt genes show faster substitution rates compared to those of *B. mossambicus* and *B. adspersus*, although the latter species has the largest mt genome among vertebrates. From these results, it is concluded that mt genomes with large genome sizes also tend to have fast nucleotide substitution rates but the latter factor is not a direct cause of genome enlargement in afrobatrachian frogs.

It has been shown that animal mt genomes are subject to a strong purifying pressure that suppresses mutations leading to functional changes because the functions of mt coding genes are essential for respiration (e.g., (Li *et. al.*, 2019)). Nevertheless, it has been reported that the purifying pressure of mt genomes is relaxed in neobatrachians, especially in the ranoid lineages (Irisarri *et. al.*, 2012; Kurabayashi & Sumida, 2013). A relaxation of purifying selective pressure, leading to a reduction in functional constraints that purge slightly deleterious mutations, has been suggested as one of the causes of mt genome enlargement (Boussau *et. al.*, 2011). In general, the ratio of nonsynonymous and synonymous substitutions (dN/dS ratio, ω) is useful to understanding the conditions for selection of the genes: positive selection ($\omega > 1$), neutral evolution ($\omega = 1$), and purifying selection ($\omega < 1$). Here, I estimate the dN/dS ratio using the branch model to investigate the changes in selective pressures in the ranoid lineages. In this analysis, I calculated the ω value(s) of the five branch models and compared the resultant log likelihood values among these models (Table 4). The LRT showed that model 4, with variable ω for all neobatrachian branches, is the best fitted among the models tested ($P = 6.6 \times 10^{-21}$ vs. model 3 with the 2nd largest likelihood value).

Figure 5 indicates the estimated ω values for the ranoid lineages under the best branch model (model 4). In all neobatrachian lineages, the estimated ω values are less than 1 (0.039–0.096) and confirm that the mt genomes of neobatrachian frogs are under strong purifying pressure. Compared to the background ω value (0.053) of the nonranoid neobatrachian lineage, the ω values are high in 25 of 31 ranoid branches (the branches shown in black and red in Figure 5), indicating that the purifying selection has been relaxed in these lineages. Largely relaxed branches mainly correspond to the ancestral lineages, specifically common ancestral lineages leading to *Breviceps* (i.e.,

ranoids, afrobatrachians, and microhylids, afrobatrachians, and xenosyneunitanurans: $\omega = 0.096$, 0.090, 0.89, and 0.89, respectively). In contrast, the ω values of the branches leading to the extant frogs (i.e., terminal branches) tend to be lower. On the other hand, the ω values of six out of 16 terminal branches are less than the background ω value, indicating that the purifying selection increased again in these lineages.

Among anurans, the mt genomes with >20 kbp size have only been reported from ranoid taxa (see (NCBI organelle genome)) and obvious increases in the dN/dS ratio are found in most ranoid lineages. These facts appear to support the idea that the sizes of the mt genome can change concomitantly with lower selection strengths (Boussau *et. al.*, 2011). However, it is remarkable that reincreased purifying pressures are found in species lineages leading to *Mantella* and *Hoplobatrachus*, having >20 kbp mt genomes ($\omega = 0.051$ and 0.052, respectively). Furthermore, selective pressures of the species lineages of *B. adspersus* and *B. poweri* ($\omega = 0.071$ and 0.068, respectively), i.e., the exact lineages with increased genome size, are not as relaxed as those of their ancestral lineages. Consequently, my findings suggest that low selective pressure does not directly cause the huge mt genomes as does substitution rate (see the above). Rather, the existence of relaxed purifying selection in the ancestral lineages seems to have an indirect effect on the mt genome size. This indirect correlation might be caused by the accelerated accumulation of deleterious mutations under low selective pressure. Significantly, the mutations occur at the DNA replication-related sites which could induce numerous repetitive regions, leading to large mt genomes, as pointed out in Shao *et. al.* (Shao *et. al.*, 2003).

Conclusion

In this study, I show that *B. poweri* has the second largest known mt genome among the vertebrates after its congeneric *B. adspersus*. I also found that the unusually large mt genomes did not arise in a common ancestor of these *Breviceps* species but rather that the genome enlargement occurred independently in each species lineage within relatively short periods.

Consistent with this evolutionary inference, the causes of the genome enlargement differed between these species. At present, 19 nominal species are known in *Breviceps* (Frost *et. al.*, 2019), but only three mt genomes have been analyzed. It is remarkable that two of these mt genomes are the largest known among the vertebrates. Also, the mt genomic structures differ significantly between these congeneric species, suggesting high variability in the mt genomic structures in this genus. Future investigations of *Breviceps* mt genomes may provide new insights into the hidden diversity of vertebrate mt genomes.

Data Availability

The data used to support the findings of this study are available from the INSDC under the accession numbers LC498571, LC498572, and LC498573.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

Abascal F., Zardoya R., and elfordM. J. T, "TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations," *Nucleic Acids Research*, vol. 8, pp. 7–13, 2010.

Akaike H., "Information theory and an extension of the maximum likelihood principle," *Proceedings of the 2nd International Symposium on Information Theory*, B. N. Petrov and C. F. Budapest, Eds., , pp. 267–281, Akadimiai Kiado, Budapest, 1973.

Boore J. L., Collins T. M., Stanton D., Daehler L. L., and Brown W. M., "Deducing the pattern of arthropod phytogeny from mitochondrial DNA rearrangements," *Nature*, vol. 376, no. 6536, pp. 163–165, 1995.

Boore J. L. and Brown W. M., "Big trees from little genomes: mitochondrial gene order as a phylogenetic tool," *Current Opinion in Genetics and Development*, vol. 8, no. 6, pp. 668– 674, 1998.

Boore J. L., "Animal mitochondrial genomes," *Nucleic Acids Research*, vol. 27, no. 8, pp. 1767–1780, 1999.

Bouckaert R., Vaughan T. G., Barido-Sottani J. et al., "BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis," *PLoS Computational Biology*, vol. 15, no. 4, p. e1006650, 2019.

Boussau B., Brown J. M., and Fujita M. K., "Nonadaptive evolution of mitochondrial genome size," *Evolution*, vol. 65, no. 9, pp. 2706–2711, 2011.

Buschiazzo E., Ritland C., Bohlmann J., and Ritland K., "Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms," *BMC Evolutionary Biology*, vol. 12, no. 1, p. 8, 2012.

Castresana J., "Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis," *Molecular Biology and Evolution*, vol. 17, no. 4, pp. 540–552, 2000.

Feng Y. J., Blackburn D. C., Liang D. et al., "Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous–Paleogene boundary," *Proceedings of the National Academy of Sciences*, vol. 114, no. 29, pp. E5864–E5870, 2017.

Frost D. R., Grant T., Faivovich J. et al., "The amphibian tree of life," *Bulletin of the American Museum of Natural History*, vol. 297, pp. 1–291, 2006.

Frost D. R., *Amphibian species of the world: an online reference. Version 6.0*, American Museum of Natural History, New York, USA, 2019.
(Frost *et al.*, 2019)

Fujita M. K., Boore J. L., and Moritz C., "Multiple origins and rapid evolution of duplicated mitochondrial genes in parthenogenetic geckos (Heteronotia binoei; Squamata, Gekkonidae)," *Molecular Biology and Evolution*, vol. 24, no. 12, pp. 2775–2786, 2007.

Gissi C., San M. D., Pesole G., and Zardoya R., "Mitochondrial phylogeny of Anura (Amphibia): a case study of congruent phylogenetic reconstruction using amino acid and nucleotide characters," *Gene*, vol. 366, no. 2, pp. 228–237, 2006.

Hoegg S., Vences M., Brinkmann H., and Meyer A., "Phylogeny and comparative substitution rates of frogs inferred from sequences of three nuclear genes," *Molecular Biology and Evolution*, vol. 21, no. 7, pp. 1188–1200, 2004.

Hou Y., Wu B., Liu Z.-H. et al., "Complete mitochondrial genome of ark shell *Scapharca subcrenata*," *Mitochondrial DNA Part A*, vol. 27, no. 2, pp. 939-940, 2016.

Hwang D. S. and Lee J. S., "Complete mitochondrial genome of the boreal digging frog Kaloula borealis (Anura, Microhylidae)," *Mitochondrial DNA*, vol. 23, no. 4, pp. 301-302, 2012.

Irisarri I., Mauro D. S., Abascal F., Ohler A., Vences M., and Zardoya R., "The origin of modern frogs (Neobatrachia) was accompanied by acceleration in mitochondrial and nuclear substitution rates," *BMC Genomics*, vol. 13, no. 1, p. 626, 2012.

Jemt E., Persson Ö., Shi Y. et al., "Regulation of DNA replication at the end of the mitochondrial D-loop involves the helicase TWINKLE and a conserved sequence element," *Nucleic Acids Research*, vol. 43, no. 19, pp. 9262–9275, 2015.

Kimura M., "A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences," *Journal of Molecular Evolution*, vol. 16, no. 2, pp. 111–120, 1980.

Kakehashi R., Kurabayashi A., Oumi S., Katsuren S., Hoso M., and Sumida M., "Mitochondrial genomes of Japanese *Babina* frogs (Ranidae, Anura): unique gene arrangements and the phylogenetic position of genus *Babina*," *Genes & Genetic Systems*, vol. 88, no. 1, pp. 59–67, 2013.

Katoh K. and Toh H., "Recent developments in the MAFFT multiple sequence alignment program," *Briefings in Bioinformatics*, vol. 9, no. 4, pp. 286–298, 2008.

Kayal E., Bentlage B., Collins A. G., Kayal M., Pirro S., and Lavro D. V., "Evolution of linear mitochondrial genomes in medusozoan cnidarians," *Genome Biology and Evolution*, vol. 4, no. 1, pp. 1–12, 2011.

Kumar S., Stecher G., Suleski M., and Hedges S. B., "TimeTree: a resource for timelines, timetrees, and divergence times," *Molecular Biology and Evolution*, vol. 34, no. 7, pp. 1812–1819, 2017.

Kurabayashi A., Usuki C., Mikami N. et al., "Complete nucleotide sequence of the mitochondrial genome of a Malagasy poison frog Mantella madagascariensis: evolutionary implications on mitochondrial genomes of higher anuran groups," *Molecular Phylogenetics and Evolution*, vol. 39, no. 1, pp. 223–236, 2006.

Kurabayashi A., Sumida M., Yonekawa H., Glaw F., Vences M., and Hasegawa M., "Phylogeny, recombination, and mechanisms of stepwise mitochondrial genome reorganization in mantellid frogs from Madagascar," *Molecular Biology and Evolution*, vol. 25, no. 5, pp. 874–891, 2008.

Kurabayashi A., Yoshikawa N., Sato N. et al., "Complete mitochondrial DNA sequence of the endangered frog *Odorrana ishikawae* (family Ranidae) and unexpected diversity of mt gene arrangements in ranids," *Molecular Phylogenetics and Evolution*, vol. 56, no. 2, pp. 543–553, 2010.

Kurabayashi A. and Sumida M., "Afrobatrachian mitochondrial genomes: genome reorganization, gene rearrangement mechanisms, and evolutionary trends of duplicated and rearranged genes," *BMC Genomics*, vol. 14, no. 1, p. 633, 2013.

Lanfear R., Calcott B., Ho S. Y. W., and Guindon S., "PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses," *Molecular Biology and Evolution*, vol. 29, no. 6, pp. 1695–1701, 2012.

Lavrov D. V., and Pett W., "Animal mitochondrial DNA as we do not know it: mt-genome organization and evolution in nonbilaterian lineages," *Genome Biology and Evolution*, vol. 8, no. 9, pp. 2896–2913, 2016.

Li P. and Bousquet J., "Relative rate test for nucleotide substitutions between two lineages," *Molecular Biology and Evolution*, vol. 9, pp. 1185–1189, 1992.

Li J.-Y., Zeng C., Yan G.-Y., and He L.-S., "Characterization of the mitochondrial genome of an ancient amphipod Halice sp. MT-2017 (Pardaliscidae) from 10,908 m in the Mariana Trench," *Scientific Reports*, vol. 9, no. 1, p. 2610, 2019.

Liu Y. G., Kurokawa Sekino T., M., Tanabe T., and Watanabe K., "Complete mitochondrial DNA sequence of the ark shell _Scapharca broughtonii_ : An ultra-large metazoan mitochondrial genome," *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, vol. 8, no. 1, pp. 72–81, 2013.

Mahmoud M., Zywicki M., Twardowski T., and Karlowski W. M., "Efficiency of PacBio long read correction by 2nd generation Illumina sequencing," *Genomics*, vol. 111, no. 1, pp. 43–49, 2019.

Moritz C. and Brown W., "Tandem duplications in animal mitochondrial DNAs: variation in incidence and gene content among lizards," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 84, no. 20, pp. 7183–7187, 1987.

Mueller R. L. and Boore J. L., "Molecular mechanisms of extensive mitochondrial gene rearrangement in plethodontid salamanders," *Molecular Biology and Evolution*, vol. 22, no. 10, pp. 2104–2112, 2005.

NCBI organelle genome resourcesAvailable from http://www .ncbi.nlm. http://nih.gov/genome/browse!/organelles/.

Nielsen S. V., Savel R., Daniels S. R., Conradie W., Heinicke M. P., and Noonan B. P., "Multilocus phylogenetics in a widespread African anuran lineage (Brevicipitidae: *Breviceps*) reveals patterns of diversity reflecting geoclimatic change," *Journal of Biogeography*, vol. 45, no. 9, pp. 2067–2079, 2018.

Pacific Bioscience, "SMRT Link user guide (v5.1.0)," 2018, Available from https: //http://www.pacb.com/ documentation/smrt-link-user-guide-v5-1-0/.

Parker H. W., *A monograph of the frogs of the family Microhylidae*, Trustees of the British Museum, London, 1934.

Pereira F., Carneiro J., and van Asch B., "A guide for mitochondrial DNA analysis in non-human forensic Investigations~!2010-01-07~!2010-04-02~!2010-05-17~!," *The Open Forensic Science Journal*, vol. 3, no. 2, pp. 33–44, 2010.

Peters W. C. H., "Diagnosen neuer Batrachier, welche zusammen mit der früher (24. Juli und 17. August) gegebenen Übersicht der Schlangen und Eidechsen mitgetheilt werden," *Bericht über die zur Bekanntmachung geeigneten Verhandlungen der königlich-preussischen Akademie der Wissenschaften zu Berlin*, vol. 1854, pp. 614–628, 1854.

Pyron R. A. and Wiens J. J., "A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians," *Molecular Phylogenetics and Evolution*, vol. 61, no. 2, pp. 543–583, 2011.

Rambaut A. and Drummond A. J., 2009, Tracer v. 1.5 available from http://tree.bio.ed.ac.uk/software/tracer/.

Robinson R. M. and Huchon D., "RRTree: relative–rate tests between groups of sequences on a phylogenetic tree," *Bioinformatics*, vol. 16, no. 3, pp. 296-297, 2000.

Ronquist F., Teslenko M., van der Mark P., et al., "MRBAYES 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space," *Systematic Biology*, vol. 61, no. 3, pp. 539–542, 2012.

Sambrook J. and Russell D. W., *Molecular Cloning: A Laboratory Manual*, vol. 1, Cold Spring Harbor Laboratory Press, New York, 3rd Edition edition, 2001.

Shao R., Dowton M., Murrell A., and Barker S. C., "Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects," *Molecular Biology and Evolution*, vol. 20, no. 10, pp. 1612–1619, 2003.

Song F., Li H., Liu G. H. et al., "Mitochondrial genome fragmentation unites the parasitic lice of eutherian mammals," *Systematic Biology*, vol. 68, no. 3, pp. 430–440, 2019.

Stamatakis A., "RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies," *Bioinformatics*, vol. 30, no. 9, pp. 1312-1313, 2014.

Sumida M., Kanamori Y., Kaneda H. et al., "Complete nucleotide sequence and gene rearrangement of the mitochondrial genome of the Japanese pond frog *Rana nigromaculata*," *Genes & Genetic Systems*, vol. 76, no. 5, pp. 311–325, 2001.

Tanabe A. S., "Kakusan4 and Aminosan: two programs for comparing nonpartitioned, proportional, and separate models for combined molecular phylogenetic analyses of multilocus sequence data," *Molecular Ecology Resources*, vol. 11, no. 5, pp. 914–921, 2011.

Warzecha J., Fornal A., Oczkowicz M., and Bugno-Poniewierska M., "A molecular characteristic of the Anatidae mitochondrial control region – a review," *Annals of Animal Science*, vol. 18, no. 1, pp. 3–15, 2018.

Wolstenholme D. R., "Animal mitochondrial DNA: structure and evolution," in *Mitochondrial Genomes*, D. R.

Wolstenholme and K. W. Jeon, Eds., pp. 173–216, Academic Press, New York, 1992.

Xia Y., Zheng Y., Murphy R. W., and Zeng X., "Intraspecific rearrangement of mitochondrial genome suggests the prevalence of the tandem duplication-random loss (TDLR) mechanism in *Quasipaa boulengeri*," *BMC genomics*, vol. 17, no. 1, p. 965, 2016.

Xia Y., Zheng Y., Miura I., Wong P. B. Y., Murphy R. W., and Zeng X., "The evolution of mitochondrial genomes in modern frogs (Neobatrachia): nonadaptive evolution of mitochondrial genome reorganization," *BMC Genomics*, vol. 15, no. 1, p. 691, 2014.

Yule G. U., "A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 213, no. 402-410, pp. 21–87, 1925.

Yang Z., *Computational molecular evolution*, Oxford University Press, New York, 2006.

Yang Z., "PAML: a program package for phylogenetic analysis by maximum likelihood," *Computer Applications in the Biosciences*, vol. 13, no. 5, pp. 555-556, 1997.

Yang Z., "PAML 4: phylogenetic analysis by maximum likelihood," *Molecular Biology and Evolution*, vol. 24, no. 8, pp. 1586–1591, 2007.

Zhang P., Liang D., Mao R.-L., Hillis D. M., Wake D. B., and Cannatella D. C., "Efficient sequencing of anuran mtDNAs and a mitogenomic exploration of the phylogeny and evolution of frogs," *Molecular Biology and Evolution*, vol. 30, no. 8, pp. 1899–1915, 2013.

Zhang P., Zhou H., Chen Y. Q., Liu Y. F., and Qu L. H., "Mitogenomic perspectives on the origin and phylogeny of living amphibians," *Systematic Biology*, vol. 54, no. 3, pp. 391–400, 2005.

Zhang J. Y., Zhang L.-P., Yu D.-N., Storey K. B., and Zheng R.Q., "Complete mitochondrial genomes of *Nanorana taihangnica* and *N. yunnanensis* (Anura: Dicroglossidae) with novel gene arrangements and phylogenetic relationship of Dicroglossidae," *BMC Evolutionary Biology*, vol. 18, no. 1, p. 26, 2018.

Table 1. List of primers used in this study

Primer name	Ssequence (5' > 3')
CO2NeFow	CAAGAYGCRRYHTCHCCNATYATAGAAGA
CO3 150Rev	CCTTCWCGRAYNAYRTCTCGYCAYCAYTG
ND4 650 Rev*	GADCCDGCRATDGGDGCYTCDACRTG
ND3 120 Fow*	AAACTNTCMCCNTAYGARTGYGGHTTTGAYCC
ND5 720 Rev*	GCDGADACWGGKGTDGGRCCYTCYAT
Thr Fow	AAAGCAYYGGTCTTGTAARCC
R51**	GGTCTGAACTCAGATCACGTA
F51*	CCCGCCTGTTTACCAAAAACAT
ND1 660Rev*	ACRTTRAANCCNGANACHAGTTCWGAYTC
ND1 420 Fow*	CGRGCHGTHGCHCAAACNATYTCHTAYGA
AsnFow*	AAGCTCKCTGGAWWGAGYGTTTAGCTGTTAA
Trp Fow*	YHARACYAARRGCCTTCAAAGC
Bmos_16S_700Rev	TGAGTTCCGGTGATTTGTTC
Bmos_ATP6_120Rev	GTTTTTAAGAATCATGATTG
Bmos_ATP6_630Fow	GCCGTCGCAATTATTCAAGC
Bmos_CO1Nend_Rev	GTGCCGATGTCTTTGTGGTTAGT
Bmos_CO2_200Rev	AGAATGATTGCTGGTATAAC
Bmos_Met_Fow	TAAGCTCTTGGGCCCATACC
Bmos_ND2_40Fow	CTAACATTAAGCTTAGCTGTAGG
Bmos_ND3_300Rev	5'CATAAATAAAGCCAATGGT
Bmos_ND6_220Rev	TTAGTTGCTGAGCCATATCC
Bmos_upOL_Rev	GGGATATAATGTTGGTTGAC
Bpow_CO1_CendRev	GGCGAATGAGGGTTCTTC
Bpow_CO3_770Fow	TTCTATCTACTGATGAGG
Bpow_Cytb_400Rev	GGATATTTGTCCTCATGG
Bpow_ND3_310Rev	TCATTCATAGATGAAGCC
Afrobat_16S_100Rev	CTCCATGGGGTCTTYTCGTC
Bmos_CytbFow1	TCTGAGGAGGATTCTCAGTTGAC
Bmos_CytbFow2	CCAAACATCTTAGGAGACCCAGA

Bmos_CytbRev2	ATAGCCGACGAAGGCTGTTGCCA
Bmos_upSerUCNpsgene_Rev	AAGTAATACTTGATCTAGCT
Bmos_ND5Fow1	GCTAATACCGCTGCTCTTCAAGC
Bmos_ND5Fow2	ATAGTAAGGCTAGGACCAGACCT
Bmos_12S_Rev2	ACTGGTGTGCGGAGACTTGCATG
Bmos_12S350Fow	GTGAAACACTTGTTAATTAG
Bmos_ATP6_140Rev	GTTTTTAAGAATCATGATTG
Bmos_ATP6_630Fow	GCCGTCGCAATTATTCAAGC
Bmos_CO2_200Rev	AGAATGATTGCTGGTATAAC
Bmos_Lys_Fow	GATAGCAACAGCCTTTTAAG
Bmos_ND4_460Fow	CTATATACTCTAGTAGCAGC
Bmos_ND3_300Rev	CATAAATAAAGCCAATGGT
Bmos_16S_700Rev	TGAGTTCCGGTGATTTGTTC
Bmos_Met_Fow	TAAGCTCTTGGGCCCATACC
Bmos_upOL_Rev	GGGATATAATGTTGGTTGAC
Bmos_CO1_540Fow	CAAACTCCTCTATTTGTATG
Bmos_ND2_40Fow	CTAACATTAAGCTTAGCTGTAGG
Bmos_CO1_Nend_Rev	GTGCCGATGTCTTTGTGGTTAGT
Bmos_CR_Fow1	GTTTCATTAACTCTAGGATATC
Bmos_CR_Fow2	CATTAGTTTCATTAACTCTAGG
Bmos_ND6_220Fow	TTAGTTGCTGAGCCATATCC
Bmos_ND5_1420Fow	CTCCAATGACACACACAATG
Bmos_ND2_930Fow	TCATCTAAAGTAACATGACG
Bmos_upLeuCUN_Rev	AGCGTGGGCATAGCATTTAG
Afrobat_16S_100Rev	CTCCATGGGGTCTTYTCGTC
Afrobat_16S_200Fow	GTACCTTTTGCATCATGGTC
Afrobat_16S_550Rev	AGRTGGCTGCTYTYAGGCC
Bmos_CO3_440Fow	CGRGCHGTHGCHCAAACNATYTCHTAYGA
Bpow_Lys_Fow	TAGCAACAGCCTTTTAAGC
Bpow_CO2_170Rev	TTTCGATTTCTTGGGCGTC
Bpow_CO3_NeFow	ACCCATCCCTACCACATAGT
Bpow_ATP6_230Rev	GTATTAGGGATGTTAGGAT

Bpow_ND2_NeFow	ATTAACCCCACAGCACTA
Bpow_ND2_550Fow	GTCACCTAGGCTGAACAA
Bpow_afAsnFow_Cst	AAGCTTCTTTCTACCACCC
Bpow_asAsnRev	GTGGTAGAAAGAAGCTTG
Bpow_ND4_320Rev	TCAGAGGTAGTAAAGGCTA
Bpow_12S_NeRev	AGGCTAGGACCAGACCTTTG
Bpow_16S_700Rev	GCGGTGAGTTCCGGTGAT
Afrobat_Cytb_150Fow	GGYCTNTTCCTNGCNATACACTAYAC
Afrobat_Cytb_1030Rev	ATRAAWGGRTNTTCTACWGGYTG
Bpow_ND4_430Fow	ATTAGCTGGCAACTATATTGT
Bpow_Cytb_400Rev	GGATATTTGTCCTCATGG
Bpow_ND4_1150Fow	TATAACAACTATCACTGC
Bpow_ND3_310Rev	TCATTCATAGATGAAGCC
Bpow_CO3_770Fow	TTCTATCTACTGATGAGG
Bpow_CO1_CendRev	GGCGAATGAGGGTTCTTC
Bp_CO1_1060Rev	GTTAGCTAGAATAATACC
Bp_ND6_30Fow	GGGGGTTTCTTCAAATCC
Bp_ND5_180Fow	TCATGAATATTCCTTAC
Bpow_ND5_610Fow	AATTACTTCCTCCTGGC
Bpow_ND6_200Fow	GCTTTAGTGGCCGAGCC
Bpow_CO1_680Rev	GTAGAGGATGGGGTCTCC
Bre_Val_Fow	CTTACACCGAGCCAATG
Bre_Val_Rev	CATTGGCTCGGTGTAAGC
Bp_12S_800F	GACACGTACAAACCGCC
Bp_CO1_400R	CCATATGAGCTGTATTCC
Bp_ND5_1340F	CCCCGCTCTAACGTTTC
Brevi_OL_Fow1	CATAWAACACAATGYYAA
Brevi_OL_Fow2	AWMTYCCMYCCAGWGAAC
Brevi_OL_Fow3	CAGWGAACTAAAMTACTTCT
Afro_OL_LoopStem_Rev1	GRCTTCTCCCGTTTTTAA
Afro_OL_LoopStem_Rev2	GRCTTCTCCCGTTTTWAA
Afro_OL_LoopStem_Rev3	CYCSGGGRCTTCTCCCGTT

32

ME1_Fmod	ACAGTARTYCAAGYCTAYRCYCCAAC
ME2_Rmod	CWGCAWATCTGAGGTTGTTRAKATTTCT
Bpow_bfOL_C2_Fow	CAACAAGCTTCTTTCTACCACC
Bpow_bfOL_C3_Fow	CAACAAGCTTCTTTCTACCACCC
Bpow_bfOL_C4_Fow	CAACAAGCTTCTTTCTACCACCCC
Bpow_afbfOL_C2_Rev	TACACCCCCGCGGACCTTATTCC
Bpow_afbfOL_C3_Rev	TACACCCCCGCGGACCTTATTCCC
Bpow_afbfOL_C4_Rev	TACACCCCCGCGGACCTTATTCCCC
Bp_Cytb_630Fow_AD21	AGTCGCATGACTGTGTTCCAATCCAACAGGGCTAAACTC
Bp_Cytb_630Fow_AD22	CAGTACTGCACGATCGTCCAATCCAACAGGGCTAAACTC
Bp_Cytb_630Fow_AD23	GTGCTGAGCATCAGACTCCAATCCAACAGGGCTAAACTC
Bm_Cytb_F1_AD24	CACTGATCGATATGCATCTGAGGAGGATTCTCAGTTGAC
Bp_phe_12S_Rev_AD21	AGTAGTGCTACTCGACCCAGACCTTTGTGCTTATAAGAC
Bp_phe_12S_Rev_AD22	ATGCGAGATCTGCTCACCAGACCTTTGTGCTTATAAGAC
Bp_phe_12S_Rev_AD23	TGAGACATACTGAGTGCCAGACCTTTGTGCTTATAAGAC
Bm_12S_Rev_AD24	ATGTGCACTAGTGTACTGTAAATTTAAGATTAGCTG
Bpow_bfCSB_Fow1	GTGTGTTTTGAAATGGACGGGC
Bpow_bfCSB_Fow2	GCCCCGCCCCTGAAGATTCTGGAGC
Bpow_bfCSB_Rev1	AAACCCGTAAAATCAGACATG
Bpow_bfCSB_Rev2	GCCCGTCCATTTCAAAACACAC
bfCSB_Fow0	ATTGTAGAATCCTCTAATCA
intCSB_Rev0	GTAAGGGGGGGGTAGGCTCC
afCSBRev1	GTTCCTAATATACTGAGGGA
Brevi_CR_intCSB_Rev0N	GGTAAGGGGGGGGTAGGCTCCAG

*Kurabayashi A, Sumida M: PCR Primers for the Neobatrachian Mitochondrial Genome. Current Herpetol 2009, 28:1–11.

**Kurabayashi A, Usuki C, Mikami N, Fujii T, Yonekawa H, Sumida M, Hasegawa M (2006) Complete nucleotide sequence of the mitochondrial genome of a Malagasy poison frog *Mantella madagascariensis*: Evolutionary implications on mitochondrial genomes of higher anuran groups.   223–236.

Molecular Phylogenetics and Evolution 39: 223–237.

Table. 2 Details on OTU and gene data used

Table 3

Comparisons of the relative rates of nucleotide substitutions among afrobatrachian-related lineages. The lineages with faster substitution rates are shown in bold.

Compared genes	Compared lineages		Relative substitution rates		Probability	Significance
	Lineage 1	Lineage 2	Lineage 1	Lineage 2		** < 0.001
Microhylids vs. Afrobatrachians						
All 37 mt genes	Microhylids	**Afrobatrachians**	0.340	**0.380**	1.0×10^{-7}	**
All 13 protein genes	Microhylids	**Afrobatrachians**	0.391	**0.428**	1.0×10^{-7}	**
All rRNA genes	Microhylids	**Afrobatrachians**	0.165	**0.204**	1.0×10^{-7}	**
All tRNA genes	Microhylids	**Afrobatrachians**	0.225	**0.286**	1.0×10^{-7}	**
Laurentobatrachia (Arthroleptidae + Hyperoliidae) vs. Xenosyneunitanura (Hemisotidae + Brevicipitidae)						
All 37 mt genes	Laurentobatrachians	**Xenosyneunitanurans**	0.343	**0.361**	2.1×10^{-5}	**
All 13 protein genes	Laurentobatrachians	**Xenosyneunitanurans**	0.389	**0.408**	2.6×10^{-4}	**
All rRNA genes	Laurentobatrachians	**Xenosyneunitanurans**	0.178	**0.192**	0.098	
All tRNA genes	Laurentobatrachians	**Xenosyneunitanurans**	0.248	**0.267**	0.114	
Hemisus vs. *Breviceps* (Hemisotidae vs. Brevicipitidae)						
All 37 mt genes	***Hemisus***	*Breviceps*	**0.369**	0.358	0.058	
All 13 protein genes	***Hemisus***	*Breviceps*	**0.413**	0.404	0.207	
All rRNA genes	***Hemisus***	*Breviceps*	**0.203**	0.184	0.083	
All tRNA genes	***Hemisus***	*Breviceps*	**0.279**	0.266	0.470	
Breviceps mossambicus vs. *B. adspersus*						
All 37 mt genes	*B. mossambicus*	***B. adspersus***	0.351	**0.351**	0.993	
All 13 protein genes	*B. mossambicus*	***B. adspersus***	0.393	**0.394**	0.906	
All rRNA genes	*B. mossambicus*	***B. adspersus***	0.185	**0.190**	0.514	
All tRNA genes	***B. mossambicus***	*B. adspersus*	**0.279**	0.266	0.237	
Breviceps mossambicus vs. *B. poweri*						
All 37 mt genes	*B. mossambicus*	***B. poweri***	0.351	**0.375**	4.6×10^{-6}	**
All 13 protein genes	*B. mossambicus*	***B. poweri***	0.393	**0.423**	5.9×10^{-6}	**
All rRNA genes	*B. mossambicus*	***B. poweri***	0.185	**0.193**	0.311	
All tRNA genes	*B. mossambicus*	***B. poweri***	0.279	**0.287**	0.551	
Breviceps adspersus vs. *B. poweri*						
All 37 mt genes	*B. adspersus*	***B. poweri***	0.351	**0.375**	3.7×10^{-6}	**
All 13 protein genes	*B. adspersus*	***B. poweri***	0.394	**0.423**	6.5×10^{-6}	**
All rRNA genes	*B. adspersus*	***B. poweri***	0.190	**0.193**	0.695	
All tRNA genes	*B. adspersus*	***B. poweri***	0.266	**0.287**	0.087	

Table 4

Branch models used to estimate the dN/dS ratio (ω) of ranoid lineages.

Model	Constraint of the model	$-\ln L^*$ of the Model	LRT** vs. Model 0	1	2	3
0 Constant ω	All neobatrachian branches have single ω	148549.3	-	-	-	-
1 Two ω (Afrobatrachia)	Afrobatrachian branches have one unique ω	148536.2	3.2×10^7	-	-	-
2 Two ω (Ranoides)	Ranoid branches have one unique ω	148539.0	5.5×10^6	NC***	-	-
3 Eleven ω	All afrobatrachian branches have distinct ω	148515.9	5.1×10^{10}	1.3×10^5	1.4×10^6	-
4 Thirty-one ω	All ranoid branches have distinct ω	148443.6	9.6×10^{29}	2.7×10^{24}	2.6×10^{25}	6.6×10^{21}

Figure 1. Control reagion of Breviceps poweri and B. mossambicus.

The numbers of repeat units and the total length of the repititive sequence foe each repeat region are shown in brankets.

Putative conserved sequence elements, Terminal associated sequences (TAS) and conserved sequence block 1-3 (CSB1-3) known in animal mtDNAs were also indicated.

* indicated the conserved nucleotides in TAS and CSBI.

A. Control reagion of Breviceps poweri (voucher 15-004)

The total length of this control region is 12537 bp. The six major repetitive elements are shown.

1

AATCTACCCTCACTCATCTACCCTCACTCATCTACCCACACTCATCTACCCACACTCATCTACCCACACATCTAAACCCTTATCTCTCATTATTATTCGCCTTGTCGTTCTACTTTCAACCCTTCTCACTCTTCCCTTAA
CCCTTACGGTAACCACTACCTCATTTCTTTTCTTTCATTTATACCCCCATAAATGTCCTAAATTTTTATTTATCTCACCTATT

[Repeat 1: 85 bp x 2 units + 3' partial unit,   236 bp]
                    - putative TAS -
                                                                                              **   ***

AACCATAAGTGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCACCACCCTTACACTTATTAAT
AACCATAAGTGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCACCACCCTTACACTTATTAAT
AACCATAAATGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCA

[Repeat 2: 1150 bp x 1 unit, 1150 bp]
TTAACTAATATTATGATACTTTTACTTCTTAAATTACCCCACATAACATGATTCTCCCTTATTATATCATGTCTTATCCCATACTGACAATACTAGTTTGTACCTCTCTGATTTTCCATACTGTAATACTCTCCAAAGT
TAATCCACTCCTTGATCGTATCCAACTCCATAATCACTATCAACTCTTGATTAAAAAAATAAAACCATAATCAACAACTTCACAATCAACAGAAATGAATACCGTATTTTATGATTGAACAATCCCAAACTCCTAACT
ATCAATCGTTACATTAACTAAACCATGAATATGATCACTAATTAATAATGAAAAATAACCTTCCTTTAGTAAAATACTTCAAGAATGTAAATCAAAATCGTACTAGCTTCTTAATTTTACATAAAAAAACCACATTGAAA
ACTTATCAAATAACTTTGTGTTACTCCACTAAAGACTTTGTTTTAGGGACCTTACACAAGTTTCATGGCGTTGGTAAGGTTTCGCGATAATCCCAATCAGCCAGTCTACCTATCTTCATTGCGAAACCTCACCAATTT
ATTTTAACCAATCACGTGTGTCCCACTCCCTGTTATATTAACACACTGATAAGGTATAACTACTCATTAGATTCAGAAAACATCTGGTAATGATGCAAACTGACCACATTAAATTATTAATAACCCTCAATTGACTTCC
CAGGGGAACAAGGATGCGAGTTCCATATAATACATTCAAGTCCGTGATCGCGGCATCACTCCCAATCGAGAATCGCTCGTACTTTTTTTTTTTTTTACCTGTTGCCTCGCGCTCGAAAAATTAAGCCACGGAAATCCGC
TCTATCTTCGTATTTACAACTAGGCTTCAATTTCTCTCGTCTGCTCGGCAAAGGATTTTGGGCGTCGGTGCATACATCTCTGTATTACACGTTAGTCCGCTCTTCGGGGGCCTACATTCATCTCGAGTAAAAAAGA
AAAGCTGATGCTCTATAAAAACCAAAAGAAATATCTAACATTCTTAATGGAGTCGGTTCGAAAATC

   - putative CSB1-
    **   ***
AATGAAATACAGCATATAGGGCCTGGCATGTACATTATGCTTCAACTGATAAATATCAACAGGGGATATCATTATCCTTGCATTAATACATTAAAATAATGCTTTTTAAACATTA

CAAAACCCTTTTTTCAATCATTCATGCAACCTTCCTCATATGCATTATCCCCTTTATTTTTTATCAATTATTATCATTCACCTTGTCGTTCTACTTTCAACCCTTCTCACTCTTCCCTTAACCCTTACGGTAACCACTA
CCTCATTTCTTTTCTTTCATTTATACCCCCATAAATGTCCTAAATTTTTATTTATCTCACCTATC

[Repeat 1: 85 bp x 1 unit + 3' partial unit, 150 bp]
                    - putative TAS -
                                                                                              **   ***

AACCATAAGTGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCACCACCCTTACACTTATTAAT
AACCATAAGTGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCA

[Repeat 2: 1151 bp x 1 unit, 1151 bp]
TTAACTAATATTATGATACTTTTACTTCTTAAATTACCCCACATAACATGATTCTCCCTTATTATATCATGTCTTATCCCATACTGACAATACTAGTTTGTACCTCTCTGATTTTCCATACTGTAATACTCTCCAAAGT
TAATCCACTCCTTGATCGTATCCAACTCCATAATCACTATCAACTCTTGATTAAAAAAATAAAACCATAATCAACAACTTCACAATCAACAGAAATGAATACCGTATTTTATGATTGAACAATCCCAAACTCCTAACT
ATCAATCGTTACATTAACTAAACCATGAATATGATCACTAATTAATAATGAAAAATAACCTTCCTTTAGTAAAATACTTCAAGAATGTAAATTAAAATCGTACTAGCTTCTTAATTTTACATAAAAAAAACCACATTGAAA
ACTTATCAAATAACTTTGTGTTACTCCACTAAAGACTTTGTTTTAGGGACCTTACACAAGTTTCATGGCGTTGGTAAGGTTTCGCGATAATCCCAATCAGCCAGTCTACCTATCTTCATTGCGAAACCTCACCAATTT
ATTTTAACCAATCACGTGTGTCCCACTCCCTGTTATATTAACACACTGATAAGGTATAACTACTCATTAGATTCAGAAAACATCTGGTAATGATGCAAACTGACCACATTAAATTATTAATAACCCTCAATTGACTTCC
CAGGGGAACAAGGATGCGAGTTCCATATAATACATTCAAGTCCGTGATCGCGGCATCACTCCCAATCGAGAATCGCTCGTACTTTTTTTTTTTTTTTACCTGTTGCCTCGCGCTCGAAAAATTAAGCCACGGAAATCCG
CTCTATCTTCGTATTTACAACTAGGCTTCAATTTCTCTCGTCTGCTCGGCAAAGGATTTTGGGCGTCGGTGCATACATCTCTGTATTACACGTTAGTCCGCTCTTCGGGGGCCTACATTCATCTCGAGTAAAAAAG
AAAAGCTGATGCTCTATAAAAACCAAAAGAAATATCTAACATTCTTAATGGAGTCGGTTCGAAAATC

   - putative CSB1-
    **   ***
AATGAAATACAGCATATAGGGCCTGGCATGTACATTATGCTTCAACTGATAAATATCAACAGGGGATATCATTATCCTTGCATTAATACATTAAAATAATGCTTTTTAAACATTA

                                                  -- putative CSB2--   --- putative CSB3--
TTGTAGAATCCTCTAATCATACCTCTTTTCCCATGTTTCCTCCCCCTCCCACCCCGCGTGTGTTTGAAATGGACGGGCCCCGCCCCTGAAGATTCTGGAGCCTACCCCCCCTTACCCCCCAAACATGGACTTTCCC
TCAGTATATTAGGAACCCCCCCCGGAGCTAATAACTTTTTTGAGGGCCCATGTCTGATTTTACGGGTTTTCCTGCGGCCCCAGCCTCACTTCCCCCCCCCGGTTTCTGGAGCCTACCCCCCCTTACCCCCCCAAACA
TGGACTTTTCCTCAGTAAATTAGCCACCCCCCCGGAGCTAATAACTTTTTCGAGGGTCCACGTCTGTGATTTTACGGGTTTTCCTGCAGCCCCAGCCTCACTTTTTCACCCCCTCTCCCCCG

[Repeat 3: 233 - 405 bp x 23 units + 3' partial unit, total 7339 bp]
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTAAACTATGCAGATTAACTTCTCAGGTCCCCCGTGACTCATACCCACCCCCTATAAT
CCTATCCCACTATCTTATCCCGCAGTATTAACCCGCAGTATTTCTCTGCGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTAAACTATGCAGATTAACTTCTCAGGTCCCCCGTGACTCATACCCACCCCCATAATC
CTATCCCACTATCTTATCCCGCAGTATTAACCCGCAGTATTTCTCTGCGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTAAACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTAACCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTAACCCGCAGTATTTCTCTGCGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTAAACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTAACCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTAACCCGCAGTATTTCTCTACGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTCTACCTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCAACCCCCTATAA
TCCTATCCCACTATCTTATCCCGCAGTATTAACCCGCAGTATTTCTCTACGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTAACCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCC
CGCAGTATTATCCCACAGTATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT

CTTAACCTGCAGTATTAACCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTTCACTACGTACTTACTATTTCCTTCACCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTATCCCGTAGTATTATTCCGTAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCC
GCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTTCACTACGTACTTACTATTTCCTTCACCCCA
CCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTATCCCGTAGTATTATTCCGTAGTATTATTCCGTAGTATTATTCCGTAGTATTATCCC
GCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAATCTTTCACTACGTACTTTCTATTTCCTTCACCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTATCCCGCAGTATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCA
CCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCCACCCCTGCAAT
CTTATCCCACAATCTTATCCCACAATATTAACCCTCAATATTAACCCTCAATATTAACCCGCAACATTATCCCGCAACATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCACCACTGAACTCACCC
ACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTCCTCCTGTAACTCATACCCACCCCTGCAAT
CTTATCCCACAATCTTATCCCACAATATTAACCCTCAATATTAACCCTCAATATTAACCCGCAACATTATCCCGCAACATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCACCACTGAACTCACCC
ACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTCCTCCTGTAACTCATACCCACCCCTGCAAT
CTTATCCCTCAATCTTATCCCTCAATCTTATCCCTCAATCTTATCCCTCAATCTTATCCCACAATCTTATCCCACAATCTTATCCCACAATATTAACCCACAATATTAACCCACAATATTAACCCTCAACATTATCCCGC
AATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCCACCCCTGCAAT
CTTATCCCTCAATCTTATCCCTCAATCTTATCCCTCAATCTTATCCCACAATATTAACCCTCAATATTAACCCGCAACATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCACTGAACTCACC
CACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTATCCCGTAGTATTATTCCGTAGTATTATTCCGTAGTATTATCCTCAGTATTATCCTCAGTATTATCCC
TCAGTATTATCCCGCAGTATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCTCCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTATCCGTAGTATTATTCCGTAGTATTATCCCTCAGTATTATCCTCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAATCTTTCACT
ACGTACTTACTATTTCCTTCACCCCCCTCCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCCACCCCTGCAAT
CTTATCCCTCAATCTTATCCCTCAATCTTATCCCACAATCTTATCCCACAATATTAACCCACAATATTAACCCTCAACATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCCA
CCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTTCAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTATCTTCTCAGGTCCCTCCTGTAACTCATACCCACCCCTGCAAT
CTTATCCCTCAATCTTATCCCACAATATTAACCCGCAACATTAACCCGCAACATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTAC
CCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTTCAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTATCTTCTCAGGTCCCTCCTGTAACTCATACCCACCCCTGCAATC
TTATCCCACAATCTTATCCCACAATATTAACCCACAATATTAACCCACAATATTAACCCACAATATTAACCCACAATATTAACCCACAATATTAACCCACAATATTAACCCACAATATTAACCCTCAA
CATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCCACCCCTGCAAT
CTTATCCCACAATCTTATCCCACAATATTAACCCTCAATATTAACCCTCAATATTAACCCGCAACATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCACCACTGAACTCACCCACCTTAATGTCGA
GTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCCACCCCTGCAAT
CTTATCCCTCAATCTTATCCCTCAATCTTATCCCTCAATCTTATCCCTCAATCTTATCCCACAATCTTATCCCACAATCTTATCCCACAATCTTATCCCAC
AATATTAACCCTCAACATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCCACCCCTGCAAT
CTTATCCCTCAATCTTATCCCTCAATCTTATCCCTCAATCTTATCCCACAATATTAACCCTCAATATTAACCCTCAATATTAACCCGCAACATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCAC
CACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTTCCTCCTGTAACTCATAATT

[Repeat 4: 18 - 19 bp x 14 units + 3' partial unit, total 272 bp]
ATTCTTCATACTTACTCTT
ATTCTTCATACTTACTCTT
ATTCTTCATACTTACTCTT
ATTCTTTATACTTACTCTT
ATTCTTTATACTTACTCTT
ATTCTTTATACTTACTCTT
ATTTCACACTCTTATTCT
ATTTTTACATACTTATTCT
ATTTTTACATACTTATTCT
ATTTTTACATACTTATTCT
ATTTTTACATACTTACTCT
ATTTTTACATACTTATTCT
ATTTTTACGTACTTATTCT
ATTTTTACGTACTTATTTT
ATTCTTA

TTCTGTGGACTTCTTTACCTCTACTGAACTTTCTCATCTTAAGTCACAAGACATCTGCCCCAATATATCGTGCTTTTAAACACGCTTATTCTATGAAACTTATTAGGACCACGAAGATTCTCCACAGATCTTTCTGCT
TGCTCCAAAATTTATAACACAAATTATGGTTGATAGATTATTTTTATCGCTTGTTCATTCATGAGATAGTTAGTTAAAACCTGACTCAACCAGATGCCTTACCTAGTCCGCTTGGTCCTCTTGTTCCACTTGTTCTAC

[Repeat 5: 9 bp x 19 units + 3' partial repeat unit, total 178 bp]
TTGGCCCGC
TTGGCCCGC
TTGGCCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC

TTGGTCCGC
TTGGTCCGC
CTGGTCCGC
CTGGTCCGC
CTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGTCC

ATAAACACCTCTCCCAAATCCTCAATGAAAAACTTCATTTCTCCGACCCTCTATCAATTTCTAGTCTTCGTTGGTGCCTTTTTCTCTACCACTCTACTTATTTTCTTTCTAACTATTGCATCTCCACTAATTAAATATC
TAACTTGCAAACTATTCTACAAACTGGCTTATAAACTAATACGCAAACTAATCCGTGAACTAATCCGTAAACTTGTTCATAAACTGATTCCTAAGCTACGCCGTATTCTAAACCCTATTCTCATCCTTGCTTTTATCCC
CCCTCTT

[Repeat 6: 12 bp x 12 units + 3' partial repeat unit, total 149 bp]
AACCCCGATGTT
AACCCCGATGTT
AACCCCGATGTT
AACCCCAATGAT
AACCCCAATGAT
AACCCCAATGTT
AACCCCAATGAT
AACCCCAATGTT
AACCCCAATGTT
AACCCCGATGTT
AACCCCGATGTT
AACCCCGATGTT
AACCC

TAATCCTATTCCTAACCCCGGGCCTGAGCCTGATCTAAATATTCCTCTCCCCGATGACAATCCTCTCCTAAATTTTATTCTACATGTTTTTAACGGCCTTCTCGACGATATTATCCTCGAATTCCTCCTTGATTGAGA
TGACTCTTCCAGCTCAAGCTCAGAAGACCTAGAAGTCGAACTAGAAAATATTATAAACGTCTTCGGCCCTCAAGTACCAACAGACTCCGGGAACGAATCCGGAAGTGCTTCAGAAGACAATCCTCACCCTCATGCCC
CTGCAAATAACTCTCAATCTGACTCAGATTACGAATCTGATGCTAGCCCTTACCGCCACTTCATTCCCCGCTAACACGTTTTCTTTTGCTTTTTTCCCCACTAACACGTTTTCTTTTTACTTTTTTCCCACTAACACGT
TTTCTTTTACTTTTTTCCCACTAACACGTTTTCTTTTACTTTTTCCCCACTAACACGTTTTCTTTTACTTTTCCCACCCGCATAAATTTTTTTTTATTAAAAATTTTCTTAGCGCATCA
12537

B. Control reagion of Breviceps poweri (voucher 15-010)
The total length of this control region is 11910 bp. The six major repetitive elements are shown.
1
AATCTACCCACACTCATCTACCCACACTCATCTACCCACACTCATCTACCCACACATCTAAACCCTTATCTCTCATTATCATTCACCTTGTCGTTCTACTTTCAACCCTTCTCACTCTTCCCTTAA
CCCTTACGGTAACCACTACCTCATTTCTTTTTCTTTCATTTATACCCCCATAAATGTCCTAAATTTTTATTTATCTCACCCTATC

[Repeat 1: 85 bp x 3 units + 1 3' partial unit, total 320 bp]
                         - putative TAS -
                                                                                                      ***   ***

AACCATAAGTGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCACCACCCTTATACTTATTAAT
AACCATAAGTGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCACCACCCTTATACTTATTAAT
AACCATAAGTGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCACCACCCTTATACTTATTAAT
AACCATAAGTGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCA

[Repeat 2: 1150 bp x 1 unit, 1149 bp]
TTAACTAATATTATGATACTTTTACTTCTTAAATTACCCCACATAACATGATTCTCCCTTATTATATCATGTCTTATCCCATACTGACAATACTAGTTTGTACCTCTCTGATTTTCCATACTGCAATACTCTCCAAAGT
TAATCCACTCCTTGATCGTATCCAACTCCATAATCACTATCAACTCTTGATTAAAAAAAATAAAATCATAATCAACAACTTCACAATCAACAGAAAATGAATACCATATTTTATGATTGAACAATCCCAAACTCCTAACT
ATCAATCGTTACATTAACTAAACCATGAATATGATCACTAATTAATAATGAAAAATAACCTTCCTTTAGTAAAATACTTCAAGAATGTAAATCAAAATCGTACTAGCTTCTTAATTTTACATAAAAAAACCACATTGAAA
ATTTATCAAATAACTTTGTGTTACTCCACTAAAGACTTTGTTTTAGGGACCTTACACAAGTTTCATGGCGTTGGTAAGGTTTCGCGATAATCCCAATCAGCCAGTCTACCTATCTTCATTGCGAAACCTCACCAATTT
ATTTTAACCAATCACGTGTGTCCCACTCCCTGTTATATTAACACACTGATAAGGTATAACTACTCATTAGATTCAGAAACATCTGGTAATGATGCAAACTGACCACATTAAATTATTAATAACCCTCAATTGACTTCC
CAGGGGAACAAGGATGCGAGTTCCATATAATACATTCAAGTCCGTGATCGCGGCATCACTCCCAATCGAGAATCGCTCGTACTTTTTTTTTTTTACCTGTTGCCTCGCGCTCGAAAAATTAAGCCACGGAAATCCGCT
CTATCTTCGTATTTACAACTAGGCTTCAATTTCTCTCGTCTGCTCGGCAAAGGATTTTGGGCGTCGGTGCATACATCTCTGTATTACACGTTAGTCCGCTCTTCGGGGGCCTACATTCATCTCGAGTAAAAAAGAA
AAGCTGATGCTCTATAAAAACCAAAAGAAATATCTAACATTCTTAATGGAGTCGGTTCGAAAATC
 -putative CSB1-
 ***   ***
AATGAAATACAGCATATAGGGCCTGACATGTACATTATGCTTCAACTGATAAATATCAACAGGGGGTATCATTATCCTTGCATTAATACATTAAAATAATGCTTTTTAAACATTA

CAAAACCCTTTTTTCAATCATTCATGCAACCTTCCTCATATGCATTATCCCCTTTATTTTTTATCAATTATTTATCATTCACCTTGTCGTTCTACTTTCAACCCTTCTCACTCTTCCCTTAACCCTTACGGTAACCATTA
CCTCATTTCTTTTCTTTCATTTATACCCCCATAAATGTCCTAAATTTTTATTTATCTCACCTATCAACCATAAATGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCACCACCCTTAT
ACTTATTAAT

[Repeat 1: 85 bp x 2 units + 1 3' partial unit, 236 bp]
                         - putative TAS -
                                                                                                      ***   ***
AACCATAAGTGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCACCACCCTTATACTTATTAAT
AACCATAAGTGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCACCACCCTTATACTTATTAAT
AACCATAAGTGCATTTCCTCATACTATGTTTAATCCCCATTCATCTATATACCCCTAGCATATCA

[Repeat 2: 1150 bp x 1 unit, 1150 bp]

TTAACTAATATTATGATACTTTTACTTCTTAAATTACCCCACATAACATGATTCTCCCTTATTATATCATGTCTTATCCCATACTGACAATACTAGTTTGTACCTCTCTGATTTTCCATACTGCAATACTCTCCAAAGT
TAATCCACTCCTTGATCGTATCCAACTCCATAATCACTATCAACTCTTGATTAAAAAAAATAAAATCATAATCAACAACTTCACAATCAACAGAAAATGAATACCATATTTTATGATTGAACAATCCCAAACTCCTAACT
ATCAATCGTTACATTAACTAAACCATGAATATGATCACTAATTAATAATGAAAAATAACCTTCCCTTAGTAAAATACTTCAAGAATGTAAATCAAAATCGTACTAGCTTCTTAATTTTACATAAAAAAACCACATTGAAA
ATTTATCAAATAACTTTGTGTTACTCCACTAAAGACTTTGTTTTAGGGACCTTACACAAGTTTCATGGCGTTGGTAAGGTTTCGCGATAATCCCAATCAGCCAGTCTACCTATCTTCATTGCGAAACCTCACCAATTT
ATTTTAACCAATCACGTGTGTCCCACTCCCTGTTATATTAACACACTGATAAGGTATAACTACTCATTAGATTCAGAAAACATCTGGTAATGATGCAAACTGACCACATTAAATTATTAATAACCCTCAATTGACTTCC
CAGGGGAACAAGGATGCGAGTTCCATATAATACATTCAAGTCCGTGATCGCGGCATCACTCCCAATCGAGAATCGCTCGTACTTTTTTTTTTTTTTACCTGTTGCCTCGCGCTCGAAAAATTAAGCCACGGAAATCCGC
TCTATCTTCGTATTTACAACTAGGCTTCAATTTCTCTCGTCTGCTCGGCAAAGGATTTTGGGCGTCGGTGCATACATCTCTGTATTACACGTTAGTCCGCTCTTCGGGGGCCTACATTCATCTCGAGTAAAAAAGA
AAAGCTGATGCTCTATAAAAACCAAAAGAAATATCTAACATTCTTAATGGAGTCGGTCCGAAAATC
 -putative CSB1-

  ***   ***

AATGAAATACAGCATATAGGGCCTGACATGTACATTATGCTTCAACTGATAAATATCAACAGGGGGTATCATTATCCTTGCATTAATACATTAAAATAATGCTTTTTAAACATTA
                                                            -- putative CSB2--   --- putative CSB3--
ATTGTAGAATCCTCTAATCATACCTCTTTTCCCATGTTTCCTCCCCCTCCCACCCCGCGTGTGTTTTGAAATGGACGGGCCCCGCCCCTGAAGATTCTGGAGCCTACCCCCCCCTTACCCCCCAAACATGGACTTTC
CCTCAGTATATTAGGAACCCCCCCCGGAGCTAATAACTTTTTTGAGGGCCCATGTCTGATTTTACGGGTTTTCCTGCGGCCCCAGCCTCACTTCCCCCCCCCGGTTTCTGGAGCCTACCCCCCCTTACCCCCCCAA
ACATGGACTTTTCCTCAGTAAATTAGCCACCCCCCCGGAGCTAATAACTTTTTCGAGGGTCCACGTCTGATTTTACGGGTTTTCCTGCAGCCCCAGCCTCACTTTTTTCACCCCCTCTCCCCCG

[Repeat 3: 215 - 425 bp x 22 units + 3' partial unit, total 6186 bp]

CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTAAACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCAACCCCCTATAA
TCCTATCCCACTATCTTATCCCGCAGTATTAACCCGCAGTATTTCTCTGCGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTAAACTATGCAGATTAACTTCTCAGGTCCCCCCGTGACTCATACCCACCCCCTATAAT
CCTATCCCACTATCTTATCCCGCAGTATTAACCCGCAGTATTTCTCTACGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTTCAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTAAACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCAACCCCCTATAA
TCCTATCCCACTATCTTATCCCGCAGTATTAACCCGCAGTATTTCTCTACGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTAAACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCAACCCCCTATGA
TCCTATCCCACTATCTTATCCCGCAGTATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTAGATTATGTGTTATTTTACCCGCACATCTAAACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCAACCCCCTATGA
TCCTATCCCACTATCTTATCCCGCAGTATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCAACCCCCTATGA
TCCTATCCCACTATCTTATCCCGCAGTATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCAACCCCCTATGA
TCCTATCCCACTATCTTATCCCGCAGTATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCAACCCCCTATGA
TCCTATCCCACTATCTTATCCCGCAGTATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTAACCCGCAGTATTATCCCGCAGTATTATCCCGCAGTCTTTCACTACGTACTTACTATTTCCTTCACCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTGTTATTTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAATCTTAACCGCAGTATTAACCCGCAGTATTA
ACCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTCTTTCACTACGTACTTACTATTTCCTTCACCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCTGTGTTATTTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAATCTTAACCCGCAGTATTAACCCGCAGTATTA
ACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTATCCTGCAGTATTATCCTGCAGTATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCACCACTGAACTCACCCACCTTAATGTCGAGT
GACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCAACCCCCTATAA
TCCTATCCCACTATCTTATCCCGCAGTATTAACCCGCAGTATTTCTCTACGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTAAACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCAACCCCCTATAA
TCCTATCCCACTATCTTATCCCGCAGTATTAACCCGCAGTATTTCTCTACGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTACTATGCATAACTCTCAGGTCCCTCCTGTAACTCATACCAACCCCCTATAATCCTAT
CCCACTATCTTATCCCGCAGTATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCAACCCCCTATAA
TCCTATCCCACTATCTTATCCCGCAGTATTATCCGCATTCACGTACTATTCTCACCCCTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCGCACTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAATCTTAA
CCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTATCTGCAGTATTATCTGCAGTATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCACCACT
GAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTTCACTACGTACTTACTATTTCCTTCACCCCA
CCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCC
CGCAGTATTATCCCGCAGTATTTCACTACGTACTTACTATTTCCTTCACCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTCTTAGATTATGTGTTATTTTACCCGCACATCTCTACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCTACCCCTGTAAT
CTTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTAACCCGCAGTATTATCCCGCAGTATTATCCCGCAGTATTATCCCGTAGTATTATCCCGCAGTAATATCCCGCAGTAATATCCC
TCAGTATTATCCCTCAGTATTATCCCTCAGTATTATCCCGCAGTATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCCCCCCACCACTGAACTCACCCACCTTAATGTCGA
GTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCCACCCCTGCAAT
CTTATCCCTCAATCTTATCCCTCAATCTTATCCCACAATCTTATCCCACAATATTAACCCGCAACATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCACTGAACTCACCCACCTTAATGTC
GAGTGACTCCTA
CCCTAACACATCGTGCTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCCACCCCTGCAAT
CTTATCCCTCAATCTTATCCCTCAATCTTATCCCTCAATCTTATCCCTCAATCTTATCCCTCAATCTTATCCCTCAATCTTATCCCACA
ATCTTATCCCACAATCTTATCCCACAATAATTAACCCGCAACATTATCCCGCAATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTCCCTCCTGTAACTCATACCCACCCCTGCAAT
CTTATCCCTCAATCTTATTCCACAATCTTATCCCACAATCTTATCCCACAATCTTATCCCACAATATTAACCCACAATATTAACCCACAATATTAACCCACAATATTAACCCACAATATTAACCCTCAACATTATCCCGC
AATCTTTCACTACGTACTTACTATTTCCTTCACCCCCACCACTGAACTCACCCACCTTAATGTCGAGTGACTCCTA
CCCTAACACATCGTGCTTTAAAACACGCCTTAACCTCTACAGGTTCTCTTTCAGATTATGTGTTATTTTACCCGCACATCTATACTATGCAGATTAACTTCTCAGGTTCCTCCTGTAACTCATAATT
[Repeat 4: 18 - 19 bp x 35 units + 3' partial unit, total 665 bp]
ATTCTTCATACTTACTCTT
ATTCTTCTTACTTACTCTT
ATTCTTCTTACTTACTCTT

```
ATTCTTCTTACTTACTCTT
ATTCTTCTTACTTACTCTT
ATTCTTCATACTTACTCTT
ATTCTTCATACTTACTCTT
ATTCTTCATACTTACTCTT
ATTCTTCATACTTACTCTT
ATTCTTCATACTTACTCTT
ATTCTTCATACTTACTCTT
ATTCTTCATACTTACTCTT
ATTCTTTATACTTACTCTT
ATTCTTTATACTTACTCTT
ATTCTTTATACTTACTCTT
ATTCTTTATACTTACTCTT
ATTTCACACTCTTATTCT
ATTTTACATACTTATTCT
ATTTTTACATACTTATTCT
ATTTTTACATACTTATTCT
ATTTTTACATACTTATTCT
ATTTTTACATACTATTCT
ATTTTTACATACTTATTCT
ATTTTTACATACTTATTCT
ATTTTACATACTTATTCT
ATTTTTACATACTATTCT
ATTTTTACATACTTATTCT
ATTTTACATACTTATTCT
ATTTTTACATACTTATTCT
ATTTTTACATACTTACTCT
ATTTTACATACTTACTCT
ATTTTTACATACTTATTCT
ATTTTTTACGTACTTATTCT
ATTTTTTACGTACTTATTCT
ATTTTTTACGTACTTATTTT
ATTCTTA
```

TTCTGTGGACTTCTTTACCTCTACTGAACTTTCTCATCTTAAGTCACAAGACATCTGCCCCAATATATCGTGCTTTTAAACACGCTTATTCTATGAAACTTATTAGGACCACGAAGATTCTCCACAGATCTTTCTGCT
TGCTCCAAAATTTATAACACAAATTATGGTTGATAGATTATTTTTGTCGCTTGTTCATTCATGAGATAGTTAGTTAAAACCTGACTCAACCAGATGCCTTACCTAGTCCGCTTGGTCCTCTTGTTCCACTTGTTCTA
C

[Repeat 5: 9 bp x 16 units + 3' partial repeat unit, total 151 bp]
```
TTGGCCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGATCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGGTCCGC
TTGTCC
```

ATAAACACCTCTCCCAAATCCTCAATGAAAAACTTCATTTCTCCGACCCTCTATCAATTTCTAGTCTTCGTTGGTGCCTTTTTCTCTACCACTCTACTTATTTTCTTTCTAACTATTGCATCTCCACTAATTAAATATC
TAACTTGCAAACTATTCTACAAACTGGCTTATAAACTAATACGCAAACTAATCCGTGAACTAATCCGTAAACTTGTTCATAAACTGATTCCTAAGCTACGCCGTATTCTAAACCCTATTCTCATCCTTGCTTTTATCCC
CCCTCTT

[Repeat 6: 12 bp x 9 units + 3' partial repeat unit, total 113 bp]
```
AACCCCGATGTT
AACCCCGATGTT
AACCCCAATGAT
AACCCCAATGAT
AACCCCAATGTT
AACCCCGATGTT
AACCCCGATGTT
AACCCCGATGTT
AACCCCGATGTT
AACCCC
```

TAATCCTATTCCTAACCCCGGGCCTGAGCCTGATCTAAATATTCCTCTCCCCGATGACAATCCTCTCCTAAATTTTATTCTACATGTTTTTAACGGCCTTCTCGACGATATTATCCTCGAATTCCTCCTTGATTGAGA
TGACTCTTCCAGCTCAAGCTCAGAAGACCTAGAAGTCGAACTAGAAAATATTATAAACGTCTTCGGCCCTCAAGTACCAACAGACTCCGGGAACGAATCCGGAAGTGCTTCAGAAGACAATCCTCACCCTCATGCTC
CTGCAAATAACTCTCAATCTGACTCAGATTACGAATCTGATGCTAGCCCTTACCGCCACTTCATTCCCCGCTAACACGTTTTCTTTTGCTTTTTCCCCACTAACACGTTTTCTTTTGCTTTTTCCCCACTAACACGTT
TTCTTTTACTTTTCCCACCCGCATAAATTTTTTTTATTAAAAATTTTCTTAGCGCATCA
11910

C. Control reagion of Breviceps mossambicus (voucher 14-001)
The total length of this control region is 6618 bp. The three major repetitive elements are shown.

```
                                             - putative TAS -
1                                                                                                        ***       ***
AAATACAACATTCTTTAATCCTTGTCTTTATTTACTCAACAACACAACACCTTCACACTCTGTATGTTTAATAACCATTAGTTTCATTAACTCTAGGATATCATTACCACATACATTTATTAATCATATATATGTACTAG
TACATAATATGTACTATTTATATTCTGTATGTTTAATAACCATTAGTTTCATTAACTCTAGGATATCATTACCACATACATTTATTAATCATATATGTACTAGTACATAATATGTACATGTACTATTTATGTTTTTAAG
TACATTAGT
```

[Repeat 1: 18 bp x 33 units 594 bp]
```
  - putative TAS -
    ***        ***
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
ATTATGCACTAGTACATT
```

```
ATCGGTTTATAAACATGAATCATATATGTGTTTAATAACCATTAATTTATTAACCCCAAGCATATCTTTACCTTTACACTGTGTAATAATCATGAGTTTATGCGCTAACACATAAACGCTTAATAACCAAATATCATTTTT
CGTTTAACATTAAATATCATACTGCGTTTAACAAACGTATATTATGATACGTTAAATATCATTCATCATATTTATGTAAGTAACATTAATTTATTAACCAAAGAAAATCTTAGTACATTTATGTCTAAATAAACATTTAT
GCAACATTTTATACTTGATTACACATTAAATTATCATCCATTGAAATATTATCACCAAATACAACCCAACTCAAACACCCCACATAAAAATATTAACCCCTAGACATTCATGTGTACGCCCATACTATCACAAACTAGAT
GTACCTCTCTGTCATACCACATGAATATCACCTTCCAATATTAGTCCACTCCATGGTCTTACTAAACATCACAAACCAAAATAATCCTATGGCTAAAAAATCATAAAATAATAATTCAACAAATTCGTTATCGAACATAA
AATATACTCCACTTCAATGTCTTATTGATCCCATAATTAGTCAATATAACATTCAAATTACATAAACCATAAATTTGACTTTCAAATTAACACCATAACAAACATCCATGGTCTGTGATTTTAGGAAGAAATTTTAAGT
GTTATCGGTCTAGCACCTTAATCCACATACTGAAAACCACATGAATATGATTATTAATTAATTATTCATAATAAACCTGTAGACTTTGTTCTAGGGACTTATGTATCATAATCCAGCGGAGCCGTTGGTAAAGTTTCGC
GATAAATCCTTACAGCATTGCTACCTATTTTCATTGCGAAACTTCACCAATATATCGAGATCAATCACGTGTGCTCCACACTTCCAGAATTGCATATCATATAGGGTATAATTACTCTTCAGTTTCTGAGGCTTCTGG
TAAAGTGCAAACTAGCCTCATTCATCTCTCTAATTACCCATCAATTGATTTCCCAGGGGGGCACCCTGCGTGTTCCATATAATACATTTTTAGTCCGTGATCGCGGCATTAGGCCCATCGGAAAACTGCTTCTTTTTTT
TTTTATACCTTTGACTCCGCGCTCAAGAAACTTGAGCCATGGTAAACTACTCATTATTCGTAAGGACAGCTGGGTCAAGCTTTCTCTCGTCTGCTGCGTCAAAGGATTTTGGGCACCAGCGCACTCATCTCGGTAC
TCCCCTATTTCGTCCGCTCTTCGGGGGTCTGCTACGGTTTCTAGTCAATCAACCCCTAATGCCCCCTAACATAAAAGTAAGAACTAATCACACATAGTGGCGTGTCAATTTTAACTAATGCTTGTTAATCATACAAAC
CTTATTGAGACACAATCCTTCAACAGACTTTTCAATAGACAAGAAT
```

-putative CSB1-
```
***        ***
ATGGAAAGAAACATTTCTATATGAAATTAAACCTCCTCCAACCCCCAACCCCCAACGGGCATTCATAAACATATCAGTTTAATAAACATGACATTCATTGTTTTTTAAAGCATTAACTTATTGCTCGACGAGCATTACA
CTAATGCATTACGAGCATAATACTATTGTATATTCCTCCCCCTCTCACCAAATTTTTAAAAAAAATTTACATTATCATTTATTTTTAAAATTTCGGCCGA
```

-- putative CSB2---   -- putative CSB3--
```
AACCCCCCCCTTACCCCCCCCAAACATAGAGTTTCCTCACCTTTTTAACAACCCCCCCGGAGTTAAAAAGACTATTGATTCTTCTCTGCTAATTTAGCTTCTGTTTTATACCCTGCTGCAATATATTTATTACATTAAA
TAATTTATCACAATATATGCAACATATGTGTTTGTATATCAATATTTAAAATTACTCACGTCACCAAACGTAATCAAATGTAATCAACTTTAATTATAAATATTTGCTATGATACTCAAAATATACCTCATACCCCCACA
AATATTTGAAATATTCTTATCATGAACACCTCAAATACATTACATATACGTGTCTTCTTCTACAT
```

[Repeat 2: 49 - 83 bp x 12 units + 3 partial unit) 854 bp]
```
TACTCGAAAAATAACGCCTAAACATCACACACGTATATATATTACCACCTACAAATAATAGTTTATCCCACGC
TACTTCCACATGTATTGACATACAAGACGTATACATGTATACATGTATATACATGTGTATACATG
TACCCTGCGCTACTTCCACATGTATTGACATACAATATATATAGTATATACATGTGTATACATGTGCCCTGCGC
TACTTCCACATGTATTGACATACAAGACGTATACATGTATACATGTATATACATGTGTATACATG
TACCCTGCGCTACTTCCACATGTATTGACATACAAGACGTATACATGTATACATGTATATACATGTGCCCTGCGC
TACTTCCACATGTATTGACATACAATATATATAGTATATACATGTGTATACATG
TACCCTGCGCTACTTCCACATGTATTGACATACAATACGTATACATGTATACATGTATATACATGTGTATACATG
```

TACCCTGCGCTACTTCCACATGTATTGACATACAATATATATAGTATATACATGTGTATACATGTGCCCTGCGC
TACTTCCACATGTATTGACATACAAGACGTATACATGTATACATGTATA
TACATGTGTCCTGCGCTACTTCCACATGTATTGACATACAATATATATAGTATATACATGTGTATACATG
TACCCTGCGCTACTTCCACATGTATTGACATACAATACGTATACATGTATACATGTATATACATGTGTATACATG
TACCCTGCGCTACTTCCACATGTATTGACATACAAGACGTATACATGTATACATGTATACATGTGTATACATATTCCCTGCGC
TACTTCCACATAATCGATATAC

TTCTACACATATTACCCACACCACCTCCACAAACTAAATCAAATATCTACTTGTGGGCAAAG

[Repeat 3: 49 -83 bp x 12 units 2780 bp]
TCCAATAGACAATAAACTATGAAGTTAAAATCTTCATGCTCATAACTACCCTACTTGTGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCCCATAAAATAGATCTAAAAGCAATTCTACTTTCAACGGTTCAACACGAACCTTTATATGCAGGCATGTTCAAAATACACATTCCACAAACTAGATCCAAA
TACTCCCATGTGAGCACAA
TCCAATAGACAATAAACTATGAAGTTAAAATCTTCATGCTCATAACTACCCTACTTGTGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCTCACTATTCCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCTCACTATTCCCCATGTGAGCACAA
TCCAATAGACAATAAACTATGAAGTTAAAATCTTCATGCTCATACTACCCTACTTGTGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCCCATAAAATAGATCTAAAAGCAATTCTACTTTCAACGGTTCAACACGAACCTTTATATGCAGGCATGTTCAAAATACACATTCCACAAACTAGATCCAAA
TACTCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAGTTAAAATCTTCATGCTCATAACTACCCTACTTGTGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCTCACTATTCCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAGTTAAAACCCTTCATGCTCATAACTACCCCTACTTGTGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCTCACTATTCCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAGTTAAAATCTTCATGCTCATAACTACCCTACTTATGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCCCATAAAATAGATCTAAAAGCAATTCTACTTTCAACGGTTCAACACGAACCTTTATATGCAGGCATGTTCAAAATACACATTCCACAAACTAGATCCAAA
TACTCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAGTTAAAATCTTCATACTCATAACTACCCTACTTGTGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCTCACTATTCCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAGTTAAAACCCTTCATGCTCATAACTACCCCTACTTGTGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCTCACTATTCCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAGTTAAAATCTTCATGCTCATAACTACCCTACTTATGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCCCATAAAATAGATCTAAAAGCAATTCTACTTTCAACGGTTCAACACGAACCTTTATATGCAGGCATGTTCAAAATACACATTCCACAAACTAGATCCAAA
TACTCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAGTTAAAATCTTCATGCTCATAACTACCCTACTTGTGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCTCACTATTCCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAATTAAAATCTTCATGCTCATAACTACCCTACTTGTGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCTCACTATTCCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAGTTAAAATCTTCATGCTCATAACTACCCCACTTATGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCCCATAAAATAGATCTAAAAGCAATTCTACTTTCAACGGTTCAACACGAACCTTTATATGCAGGCATGTTCAAAATACACATTCCACAAACTAGATCCAAA
TACTCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAGTTAAAATCTTCATGCTCATAACTACCCTACTTGTGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCTCACTATTCCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAGTTAAAATCTTCATGCTCATAACTACCCTACTTATGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCCCATAAAATAGATCTAAAAGCAATTCTACTTTCAACGGTTCAACACGAACCTTTATATGCAGGCATGTTCAAAATACACATTCCACAAACTAGATCCAAA
TACTCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAGTTAAAATCTTCATGCTCATAACTACCCTACTTGTGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCTCACTATTCCCCATGTGAGCACAG
TCCAATAGACAATAAACTATGAAATTAAAATCTTCATGCTCATAACTACCCCTACTTGTGGGCAAAG
TCCAATAGACAATAAACTATGAAGTTAAAACCTTCATGCCCATAAAATAGATCTAAAAGCAATTCTACTTTCAACGATTCAACACGAACCTTTATATGCAGGCATGTTCAAAATACACATTCCACAAACTAGATCCACA
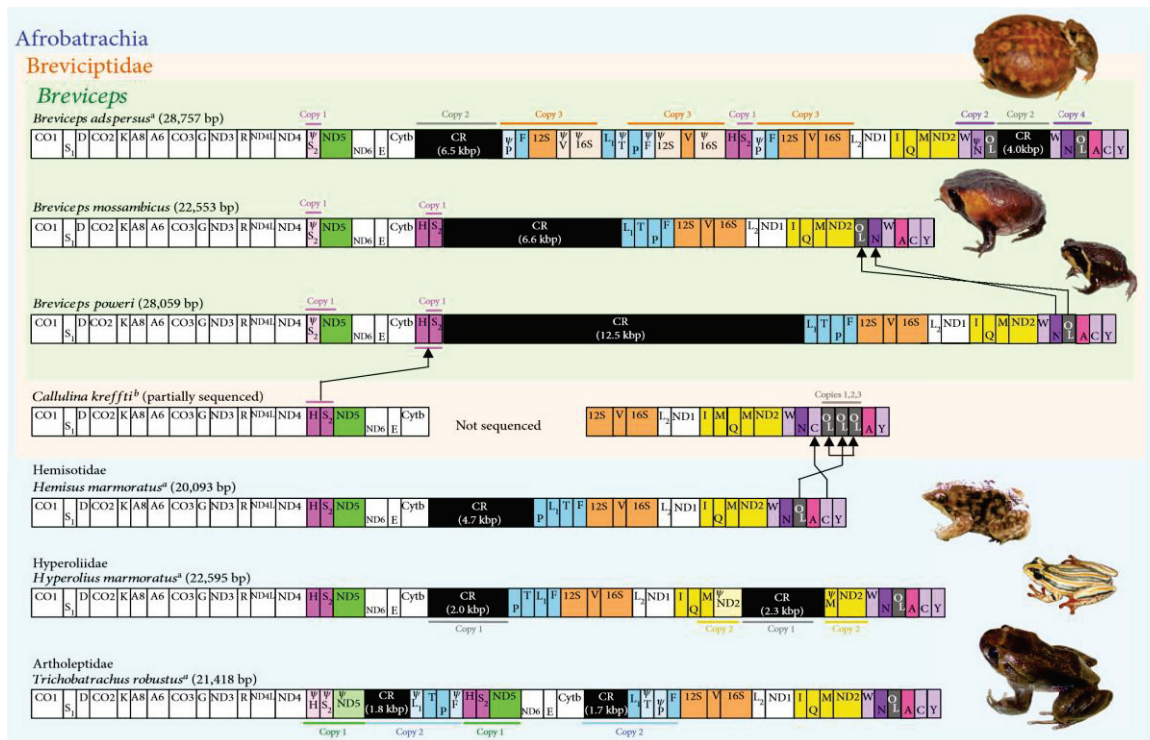ATA

TTCTTCCTCACTAAATGCTATGCCCAC
6618

Figure 2. PCR fragments of Breviceps poweri and B. mossambicus control
regions.
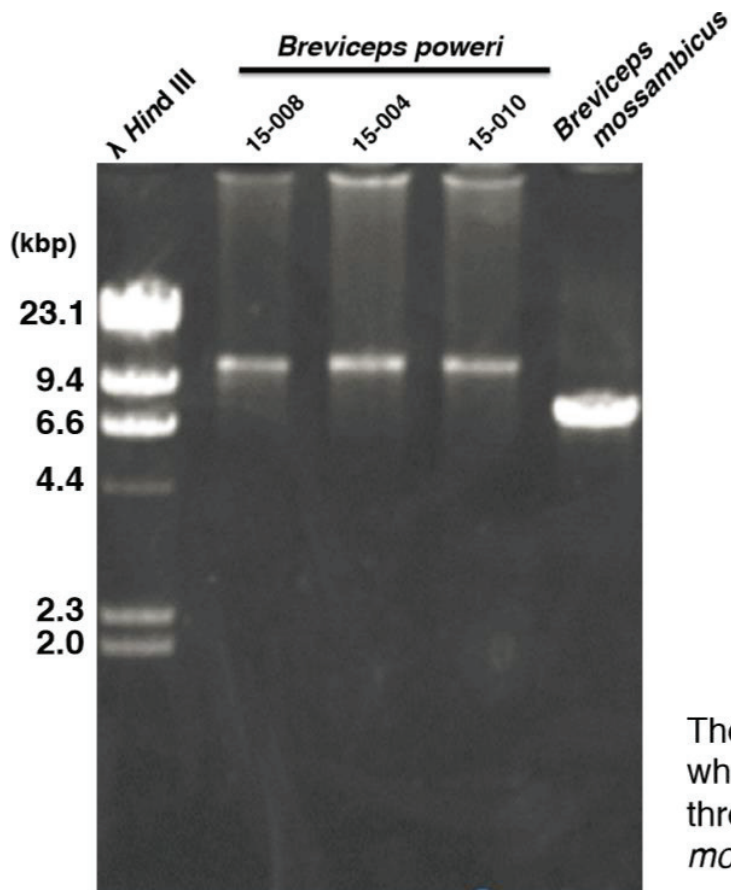
The PCR fragments containing the
whole control regions amplified from
three *Breviceps poweri* and one *B.
mossanbicus* specimens are shown.

Figure. 3

Mitochondrial genome organization of afrobatrachians and other anurans. The mitochondrial (mt) genome organization of *Breviceps mossambicus* and *B. poweri* determined in this study is compared with that of other afrobatrachians, neobatrachians, and vertebrates reported in previous studies (Kurabayashi and Sumida (2013) and Zhang et al. (2013)). Genes, pseudogenes, control regions (CRs), and *light*-strand replication origins (O$_L$) are shown in boxes. The *heavy*- and *light*-strand encoded genes are denoted above and below each gene box, respectively. The boxes do not reflect the actual sizes of the genes and CRs. The single-letter amino acid codes designate the corresponding transfer RNA genes (*trn*s). L1, L2, S1, and S2 indicate *trn*s for Leu (UUR), Leu (CUN), Ser (UCN), and Ser (AGY), respectively. "ψ" shows a pseudogene. Other gene abbreviations are as follows. 12S and 16S: 12S and 16S ribosomal RNAs; CO1–3: cytochrome c oxidase subunits 1–3; Cytb: cytochrome apoenzyme *b*; ND1–6 and 4L: NADH dehydrogenase subunits 1–6 and 4L. The genes, pseudogenes, O$_L$, and CRs with duplications and/or rearrangements in afrobatrachians are colored. "Copy" with number shows the duplicated regions within a species. Closed arrows between species indicate the rearranged genes and the presumed evolutionary direction of the translocations. The photos of afrobatrachian species are also provided (excluding *Callulina kreffti*).

Fig.4
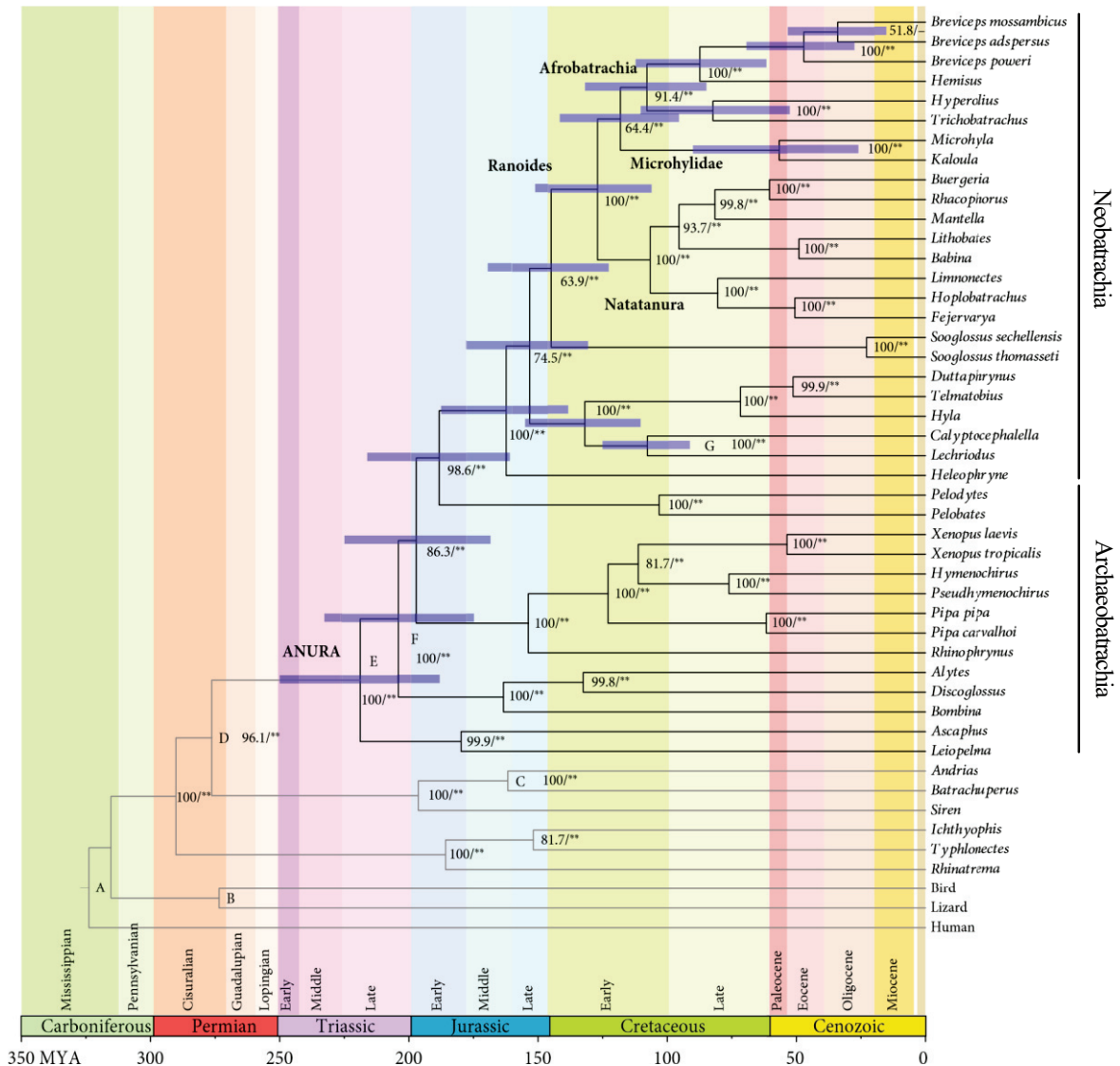
Time tree of anurans. A phylogenetic tree reflecting the divergence ages estimated using a Bayesian relaxed dating method with the 15,093 bp nucleotide data. The tree topology of amphibians is the same as that of the resultant ML and BI trees. Bold branches indicate the lineages leading to the extant anurans. Horizontal blue bars on each node indicate 95% credibility intervals of estimated divergence age. The bootstrap probability (BP) of ML and Bayesian postprobabilities (BPP) are also shown on the right side of each node (BP value/*,** > 95 and 99 BPPs), and the calibration points used in the dating analysis are indicated on the corresponding nodes (A to G). The scale of the horizontal

Fig.5
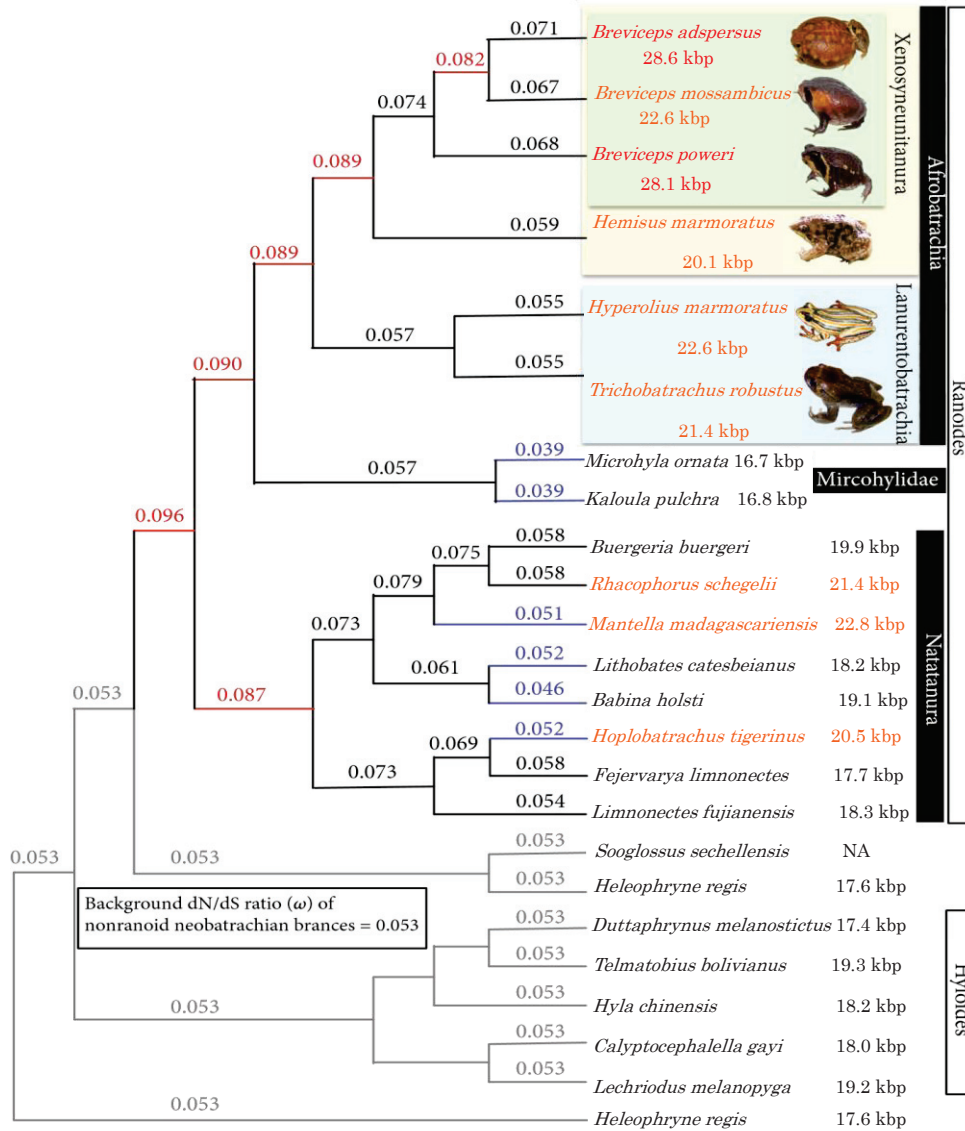
Changes in the dN/dS ratio (ω) among ranoid lineages. The estimated ω values of neobatrachian branches are shown (based on model 4 in Table 2). The tree topology is the same as those of the ML and BI trees reconstructed in this study. The constant ω (0.053) of nonranoid neobatrachian lineages was regarded as the background value. The estimated ω is shown on each ranoid branch. A high ω indicates the relaxation of purifying pressure. The branches for which ω values are lower (<0.053) and 1.5 times higher (>0.08) than the background are shown in blue and red colors, respectively. The frog taxa having mtDNAs exceeding >20kbp and 28kbp are also highlighted by orange and red colors, respectively.