Heterogeneity of synonymous substitution rates in the *Xenopus* frog genome

Quintin Lau[ID][1]*, **Takeshi Igawa**[2], **Hajime Ogino**[2], **Yukako Katsura**[3], **Toshimichi Ikemura**[4], **Yoko Satta**[1]

1 Department of Evolutionary Studies of Biosystems, Sokendai (The Graduate University for Advanced Studies), Hayama, Kanagawa, Japan, **2** Amphibian Research Center, Hiroshima University, Higashi-Hiroshima, Hiroshima, Japan, **3** Department of Cellular and Molecular Biology, Primate Research Institute, Kyoto University, Inuyama, Aichi, Japan, **4** Department of Bioscience, Nagahama Institute of Bio-Science and Technology, Nagahama, Japan

* quintin@soken.ac.jp

Abstract

With the increasing availability of high quality genomic data, there is opportunity to deeply explore the genealogical relationships of different gene loci between closely related species. In this study, we utilized genomes of *Xenopus laevis* (XLA, a tetraploid species with (L) and (S) sub-genomes) and *X. tropicalis* (XTR, a diploid species) to investigate whether synonymous substitution rates among orthologous or homoeologous genes displayed any heterogeneity. From over 1500 orthologous/homoeologous genes collected, we calculated proportion of synonymous substitutions between genomes/sub-genomes (k) and found variation within and between chromosomes. Within most chromosomes, we identified higher k with distance from the centromere, likely attributed to higher substitution rates and recombination in these regions. Using maximum likelihood methods, we identified further evidence supporting rate heterogeneity, and estimated species divergence times and ancestral population sizes. Estimated species divergence times (XLA.L-XLA.S: ~25.5 mya; XLA-XTR: ~33.0 mya) were slightly younger compared to a past study, attributed to consideration of population size in our study. Meanwhile, we found very large estimated population size in the ancestral populations of the two species ($N_A = 2.55 \times 10^6$). Local hybridization and population structure, which have not yet been well elucidated in frogs, may be a contributing factor to these possible large population sizes.

Introduction

In comparisons between closely related species, it is well considered that different loci within a genome have different genealogical relationships. For example, many studies have compared several loci in humans with the three great ape species and showed that a majority (46.6–58.5% of loci) support the closest relationship between humans and chimpanzees; meanwhile, other loci support closer genealogical relationships between humans and gorillas or chimpanzees and gorillas [1–3]. Such variation in genealogical relationships between different gene loci can be attributed to large ancestral population size and short duration of successive speciation; this large ancestral population size can sometimes be resulting from heterogeneity of mutation

rates among loci. With the increase in available genomic data, there is potential to explore these genealogical relationships in any species groups, with particular reference to investigating heterogeneity of mutation rates.

Such inconsistency of species genealogy with gene genealogy among loci has also been explored in anuran frogs, where different genealogies were explored among four lineages: *Silurana*, *Xenopus*, *Pipa*, and *Hymenochirus* [4]. It is generally accepted that *Silurana* and *Xenopus* form a clade, and this genealogy was confirmed to have the strongest statistical support following Bayesian analysis [4]. Indeed, revalidation of *Xenopus* topology and taxonomy [5] classifies *Silurana* as a subgenus within *Xenopus*, which includes *X. tropicalis*. Based on recently published genomic data, the common ancestor of *X. laevis* and *X. tropicalis* is estimated to have diverged ~48 mya [6]. In addition, *X. laevis* is a tetraploid species, comprised of (L) and (S) sub-genomes that diverged ~34 mya and hybridized ~17 mya, while *X. tropicalis* remained a diploid species [6]. We are interested in how this kind of species history is reflected in the genome, especially through demographic history of these species. However, to accurately know the demographic history of these species, we must first ascertain the extent of mutation rate heterogeneity.

One of the first studies to investigate synonymous substitution rates (equivalent to mutation rate, according to neutral theory) in amphibian genes was based on singular nuclear DNA sequences [7], including *c-myc*, *slug*, and tyrosinase precursor gene, in different anuran lineages (summarized in Table 1). Using more comprehensive genomic sequence data, a more accurate average synonymous substitution rate among over 8000 orthologous genes in *X. laevis* and *X. tropicalis* (identified using BLASTP followed by synteny agreement using BAC-FISH) was reported to be 3.0 to 3.2 x10^{-9} substitutions/site/year [6]. However, further details have yet to be reported, in relation to the range and variation of synonymous divergence. Thus, in this study we further examine the publically available *Xenopus* genomes to investigate whether there is synonymous substitution rate heterogeneity in specific genomic regions or chromosomes.

Materials and methods

Data collection and calculation of synonymous substitutions

Three *Xenopus* genomes/sub-genomes were downloaded from Xenbase (http://www.xenbase.org/, RRID:SCR_003280): (i) XTR: *Xenopus tropicalis* (v9.1 genome assembly), (ii) XLA.L: *Xenopus laevis* (v9.1 genome assembly) L sub-genome, and (iii) XLA.S: *X. laevis* S sub-genome. Genomes were aligned using Synteny Mapping and Analysis Program (SyMAP) v 4.2 [8] and genomic locations for all orthologs between genomes of the two species or homoeologs between sub-genomes were compiled; the full data is available on the XenOrtho database

Table 1. Summary of synonymous substitution rates estimated in Anuran frogs in previous studies.

Synonymous substitution rate (x10^{-9} substitutions/site/year)	Species	Gene (s)	Nucleotide length	Reference
0.92–1.53	Three lineages of *Eleutherodactylus*	*c-myc*	~1340 bp	[7]
1.03 (0.68–1.42)	*Xenopus tropicalis*∗ –*Xenopus laevis*	*slug*	798 bp	[7]
1.69 (1.14–2.45)	*Boophis xerophilus*–*Micrixalus fuscus*	*tyrosinase precursor* gene exon 1	~500 bp	[7]
3.35 (2.43–4.52)	*Aglyptodactylus madagascariensis*–*Fejervarya syhadrensis*	*tyrosinase precursor* gene exon 1	~500 bp	[7]
3.0†	*X. tropicalis*—*X. laevis*	8806 homoeologous genes from genome	unspecified	[6]
3.2#	*X. laevis* L–*X. laevis* S			

∗*Silurana tropicalis* now referred to as *Xenopus tropicalis*; based on Ks/2T whereby T is 48 mya† and 34# mya for divergence between *X. laevis*- *X. tropicalis* and *X. laevis* L–*X. laevis* S, respectively.

https://doi.org/10.1371/journal.pone.0236515.t001

(https://sites.google.com/view/xenorthodb/xenortho-db). A total of 1742 orthologous/homo-eologous genes that had annotations in all three genomes/sub-genomes were collated, and we then excluded genes with overlapping annotations and genes with orthologs or homoeologs across different chromosomes; finally we obtained a set of 1598 orthologous loci from the three entire genomes and sub-genomes (54–4509 bp per gene, total 875911 bp).

Geneious® 11.1.5 (https://www.geneious.com) was used to extract sequences for the 1598 orthologous/homoeologous genes, using annotated gene names as search queries. Next, coding sequences (CDS) from the three genomes were aligned and gaps were excluded. For each genome pair combination (XTR-XLA.L, XTR-XLA.S, and XLA.L-XLA.S), we calculated the number (d) and proportion ($k = d/L$) of synonymous substitutions, whereby total length (L) of synonymous sites was calculated using MEGA-X-CC [9] with Nei-Gojobori method and Jukes Cantor correction for multiple hits. These parameters (d and k) were calculated and collated separately for each chromosome.

Maximum likelihood estimates of X and Y

Using the number of synonymous substitutions (d) and length (L), we utilized a maximum likelihood method that takes into account rate heterogeneity (α) across loci and is based on a discrete gamma model by Yang [10]. $\alpha < 1$ is indicative of high heterogeneity, while infinitely large α denotes constant substitution rate or no heterogeneity. Before ML estimations, we first confirmed that d fits with a negative binomial distribution using fitdistrplus package implemented in R [11]. Then, we calculated maximum likelihood (ML) estimates of α with assumption of $X = 0$; then α, d, and L were used for ML estimates of $X = 4N_Ag\mu$ and $Y = 2\mu t$, whereby N_A represents ancestral population size, t represents divergence time of two species compared, g represents generation time, and μ represents mutation rate (synonymous substitution rate) per site per year. We calculated X and Y for two α values: infinitely large α, and α values that gave best ML estimates of X and Y. This was conducted independently for each chromosome, and also for all chromosomes collated together ('genome-wide'). We then calculated N_A and t assuming a generation time of 1 year and $\mu = 2.05 \times 10^{-9}$ substitutions/site/year; this synonymous substitution rate is calculated based on data by Session et al. [6] ($T^* = 0.0308 = T \times \mu$, where T is ~15 mya, the time of pseudogenisation = assumed time of allotetraploidization, and μ is synonymous substitution rate).

Statistical analyses

All statistical tests were conducted in R [11] or GraphPad Prism version 8.1.1 (GraphPad Software, La Jolla, CA, USA; www.graphpad.com). Within each of the three pairwise comparisons (XTR-XLA.L, XTR-XLA.S, and XLA.L-XLA.S), we used Analysis of Variance (ANOVA) test in R to determine if there are differences in proportion of synonymous substitutions (k) between chromosomes; Tukey multiple comparisons of means (95% confidence interval) was also implemented in R to identify which specific chromosomes are different to each other. To analyze whether there are changes in k (*or K*) within each chromosome, we performed the partial *F*-test using ANOVA command in R to compare linear [lm($W \sim Z$)] and quadratic [lm($W \sim Z + I (Z^2)$)] models of k (represented by Z) in relation to chromosome position (W).

Results and discussion

Synonymous divergence between and within chromosomes

We have utilized publically available genomic data from *X. laevis* and *X. tropicalis* to identify heterogeneity in the proportion of synonymous substitutions (k). Genome-wide mean k

between genes was lower in the pairwise comparison of XLA.L and XLA.S sub-genomes (k = 0.183 ± 0.043, median = 0.184) compared to inter-species comparisons (k = 0.227 ± 0.055 or 0.235 ± 0.056, median = 0.224 or 0.232) (Table 2). This is expected, since *X. tropicalis* and *X. laevis* diverged ~48 mya whereas the two *X. laevis* sub-genomes diverged ~34 mya [6], and this additional time allowed for increased accumulation of synonymous substitutions. Overall, the k values in our study are slightly lower than those previously estimated from genomic data (median XTR-XLT Ks = 0.286, XLA.L-XLA.S Ks = 0.218) [6]. This might be attributed to the different approaches in extraction of orthologous/homoeologous genes; our approach utilized alignment of synteny blocks, which could be more conservative and focused on single copy orthologs whilst avoiding paralogs and orthologs located on different chromosomes.

When we investigated k between each chromosome (i.e. at the inter-chromosomal level), there were no significant differences when comparing the *X.laevis* sub-genomes (Table 2, Fig 1A). However, when comparing either XLA.L or XLA.S with XTR, we found that k was higher in orthologs located in chromosome 10 (XTR10) relative to all other chromosomes (p < 0.0001, Table 2, Fig 1B and 1C). In addition, orthologs had higher divergence in chromosome 5 compared to chromosome 3 (p < 0.001, Table 2, Fig 1B and 1C). The significantly higher k in XTR10 when comparing XLA to XTR could be related to the dynamic fusion of chromosomes 9 and 10 in *X. laevis* [12], which we discuss later in this section.

Next, we examined intra-chromosomal divergence between XLA.L-XLA.S and XLA-XTR and found variation in k between certain chromosomes. In chromosomes 1–7, a non-linear relationship, whereby k increases with distance from the centromere (Fig 2A and S1 Fig), was confirmed by partial F-test, irrespective of the sub-genome (XLA.S, XLA.L, or XTR) used for chromosome location ($Pr(F)$ < 0.001, S1 Table). This intra-chromosomal heterogeneity suggests that recombination is more prominent with distance from the centromere, and supports the finding that recombinations and local synonymous substitutions are positively correlated [13, 14]. Furthermore, when we investigated CG composition in full genome data at each chromosome, we found higher CG composition towards the telomeres (S2 Fig). This tendency of higher CG (that is, potentially higher CpG methylation) may explain the higher synonymous substitution rates away from the centromere. While an association between GC content and recombination rate has not been demonstrated in amphibians to date, a positive correlation has been identified in many other species [15–17]. The presence of transposable elements (TEs) may also contribute to heterogeneity in substitution rates within a chromosome. In

Table 2. Proportion of synonymous substitutions (k = d/L) calculated from pairwise comparisons of the three *Xenopus* genomes/sub-genomes.

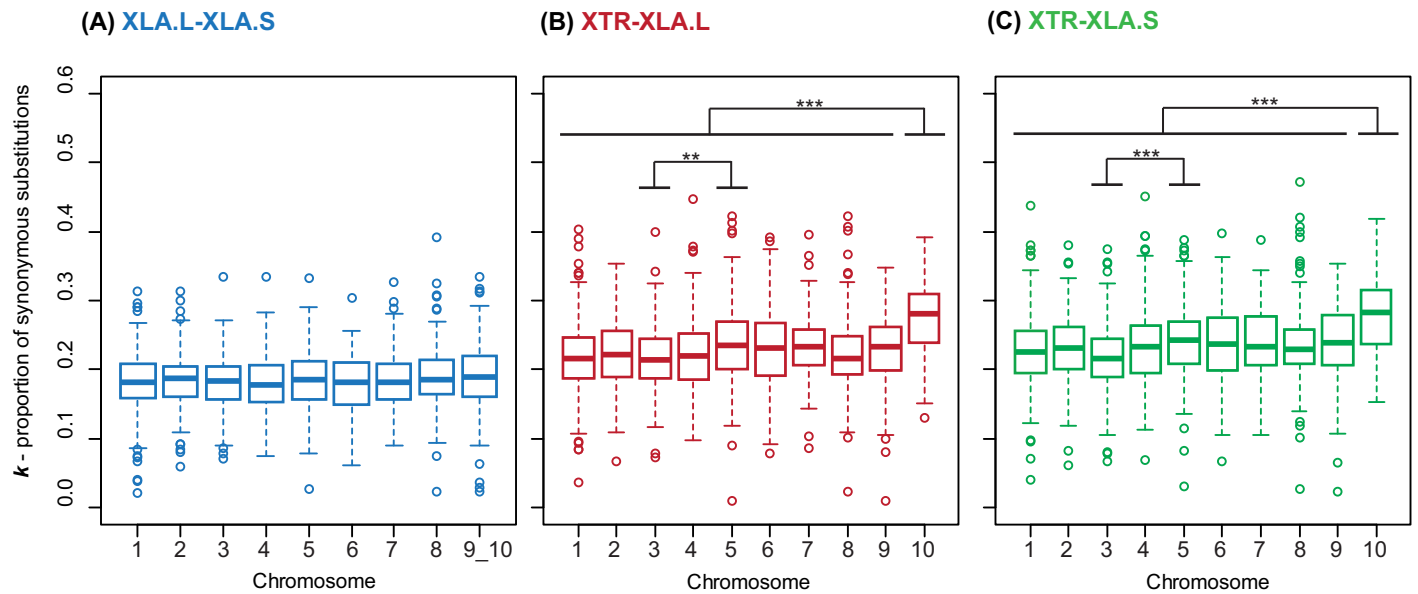| Chromosome	# genes	Proportion of synonymous substitutions, k = d/L[mean±s.d.]		
		XLA.L-XLA.S	XTR-XLA.L	XTR-XLA.S
Genome-wide	1596	0.183 ± 0.043	0.227 ± 0.055	0.235 ± 0.056
1	299	0.179 ± 0.043	0.218 ± 0.051	0.228 ± 0.053
2	187	0.184 ± 0.041	0.224 ± 0.049	0.232 ± 0.048
3	206	0.181 ± 0.037	0.217 ± 0.046	0.219 ± 0.048
4	185	0.181 ± 0.041	0.222 ± 0.056	0.232 ± 0.058
5	164	0.184 ± 0.044	0.238 ± 0.061	0.244 ± 0.056
6	148	0.178 ± 0.043	0.230 ± 0.056	0.235 ± 0.057
7	109	0.185 ± 0.043	0.233 ± 0.053	0.238 ± 0.051
8	122	0.190 ± 0.047	0.223 ± 0.059	0.237 ± 0.061
9	102	-	0.230 ± 0.058	0.239 ± 0.062
10	74	-	0.276 ± 0.053	0.279 ± 0.056
9_10	196	0.190 ± 0.050	-	-

Fig 1. Boxplots depicting proportion of synonymous substitutions (*k*) for each chromosome. Between (A) *X. laevis* L (XLA.L) and S (XLA.S) sub-genomes, (B) *X. tropicalis* (XTR) and XLA.L, and (C) XTR and XLA.S. **p < 0.01, ***p < 0.0001 (adjusted p-values from Tukey multiple comparisons of means).

Drosophila melanogaster, TEs tended to be accumulated in the proximal region (i.e. near centromere) of autosomes [18]; transposon density was also found to be negatively correlated with recombination rate in this species [19]. Moreover, a study in *X. tropicalis* found that frequency of TEs was negatively correlated with GC content [20], but chromosomal location of only specific transposons have been mapped in *X. laevis* [6]. Based on these previous studies, there may be a lower frequency of TEs in the distal parts of the *Xenopus* chromosomes along with high CG content, which may be related to high recombination rates.

Exceptional cases were observed in chromosomes 8, 9, and 10. In chromosome 8, *k* varied depending on the genome used as reference for location. When we used XTR or XLA.L genomes as references, the relation between *k* and location was non-linear (Fig 2B, $Pr(F) <$ 0.01), much like in chromosomes 1–7. However, when XLA.S was used, there was no bi-directional increase in *k* away from the centromere (Fig 2C, $Pr(F) = 0.08$–0.293). This is likely attributed to the chromosomal rearrangements and increased deletions previously reported in the S genome [6]; specifically, there is an inversion in the p-arm of XLA8S as well as homoeologous identity between the XLA8S p-arm and XLA8L q-arm (S3A Fig). We note that XLA3S also has chromosomal rearrangements but instead has a non-linear relation between *k* and location; this may be attributed to a lower degree of rearrangement and/or inversion being mainly limited to the q-arm of XLA3 (S3B Fig).

XTR-XLA and XLA.L-XLA.S intra-chromosomal examination of chromosomes 9 and 10 revealed contrasting relationships between *k* and chromosome location. The relationship between *k* and location was opposite between XTR9 and XTR10: *k* (XTR-XLA) increased with location in XTR9 (Fig 2D) but decreased with location in XTR10 (Fig 2E). Within the homoeologous XLA9_10, higher *k* in all pairwise genome comparisons was observed in the p-arm (Fig 2F and S1 Fig), which is predominantly equivalent to XTR10 based on cytogenetic mapping of chromosomes (S3C Fig); *k* then decreases distally in the q-arm of XLA9_10, which is mostly orthologous to XTR9 [6]. Chromosomes 9 and 10 probably fused sometime between the divergence of *X. laevis* with *X. tropicalis* (~48 mya) and the divergence of *Xenopus* (L) with *Xenopus*
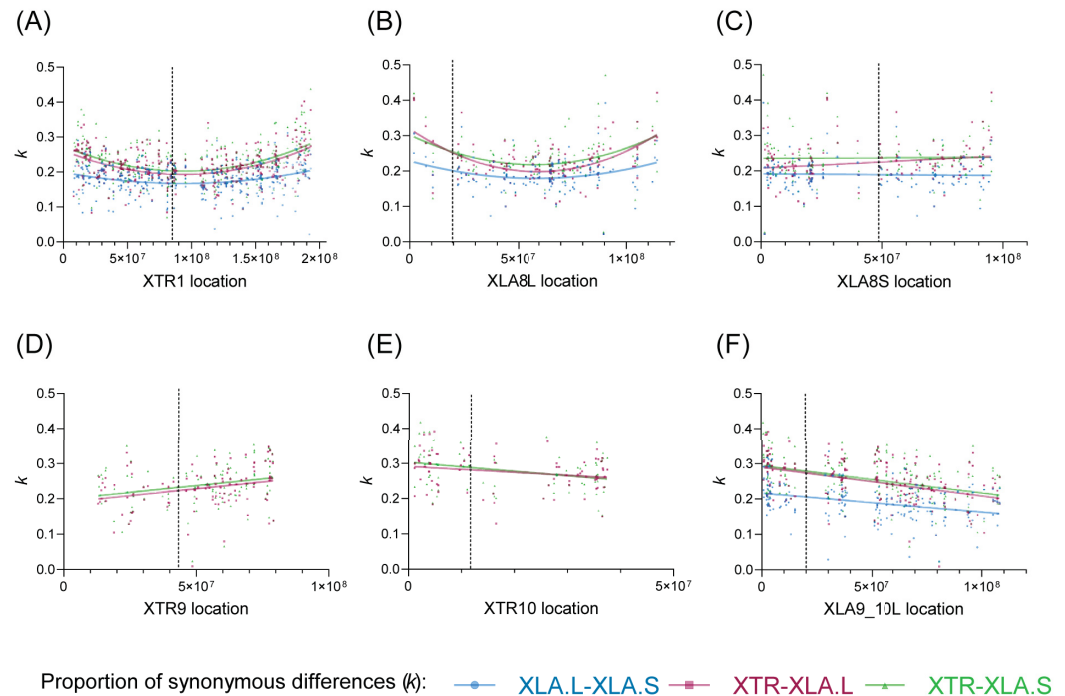
Fig 2. Relationship patterns between proportion of synonymous substitutions (*k*) and chromosome location.
Representative figures displaying different relationship patterns for each genome pair combination (XTR-XLA.L in red,
XTR-XLA.S in green, and XLA.L-XLA.S in blue). (A) Chromosome 1 (e.g. XTR1) showed a common non-linear relationship
between *k* and location. This higher *k* in distal parts of the chromosome was also seen in chromosomes 2 to 7, irrespective of
the genome used for location in the *x-axis* (S1 Fig). (B, C) Chromosome 8 showed contrasting patterns based on the genome
used for location reference: (B) XTR8 (S1 Fig) and XLA8L showed higher *k* in the distal part of chromosome, while (C) XLA8S
had no marked relationship between *k* and location, likely due to intra-chromosomal rearrangements [6]. (D) In XTR9, *k*
increased distally with location, while (E) it decreased in XTR10; (F) in the homoeologous XLA9_10, *k* decreased with location.
Centromere positions are indicated by dotted vertical line and estimated based on position of frog centromeric repeat 1 (Fcr1)
[21] or centromeric markers from *X. tropicalis* [22]. Full results are shown in S1 Fig.

https://doi.org/10.1371/journal.pone.0236515.g002

(S) (~34 mya). It seems possible that homoeologous XLA9_10 chromosome, when compared
to chromosomes 1–8, had a shorter amount of time for intrachromosomal recombination or
accumulation of substitutions away from the centromere. The fusion of chromosomes 9 and
10 in XLA may have also impacted on mutation and recombination within XLA9_10. In addi-
tion, it may also be possible that different chromosomes experienced high recombination and
mutation rates, while others could be under greater evolutionary constraint. For example in
birds, synonymous substitution rates were higher in microchromosomes compared to macro-
chromosomes [23]. While *Xenopus* frogs do not have microchromosomes, our finding of a
higher proportion of synonymous substitutions in XTR10 (Fig 1) could support that recombi-
nation rates differ between and within anuran chromosomes.

Estimation of ancestral population size and species divergence time

Among all pairwise genome comparisons across the ten chromosomes, maximum likelihood
estimates of α ranged from α = 14.9–41.8 (S2 Table) with an assumption of no ancestral poly-
morphism. Using these initial α values, we estimated X and Y again and sometimes found that
a smaller or larger α gave better ML estimates (S2 Table); these were subsequently used. ML
estimates of $X = 4N\mu$ and $Y = 2\mu t$ were significantly different (p < 0.001) when using an infi-
nitely large α (i.e. no rate heterogeneity) compared to using α with best ML estimates (Fig 3);
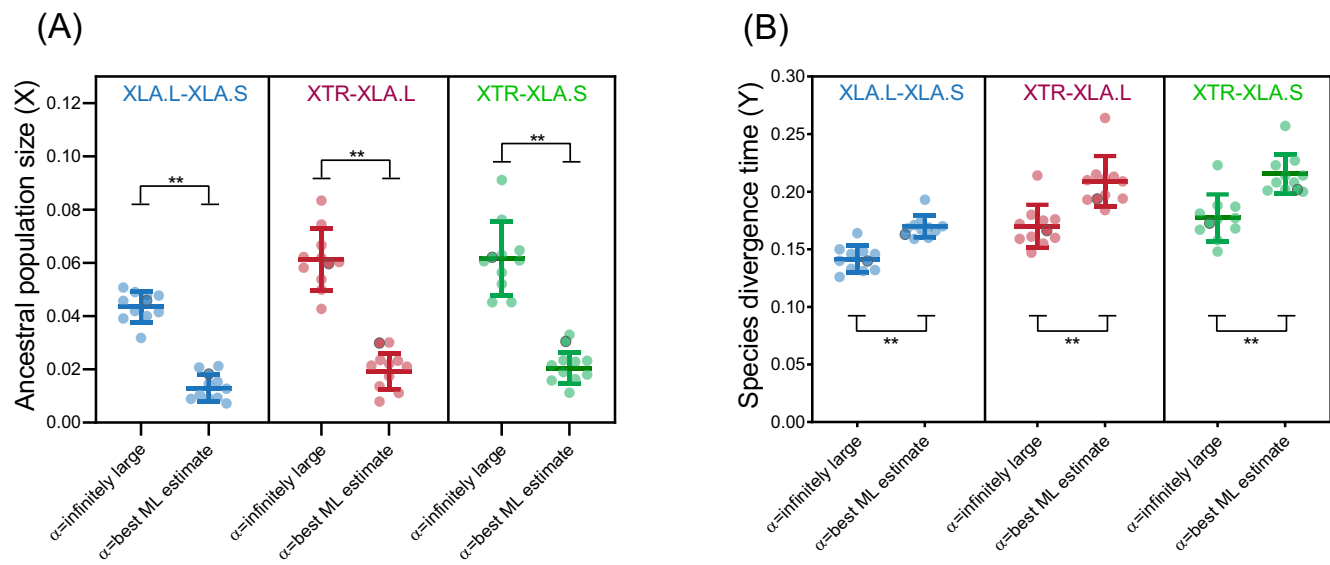
(A)

(B)

Fig 3. Best maximum likelihood estimates of $X = 4N\mu$ and $Y = 2\mu t$. Estimates were used to calculate ancestral population size and species divergence time, respectively (in ancestor of XLA-XTR or XLA.S-XLA.L). These estimates were significantly different compared to those using infinitely large α (**p < 0.0001, t-test). Line bars represent mean ± s.d.; each point represents each chromosome, with genome-wide results indicated by dark circles.

this provides evidence that some rate heterogeneity is present in the *Xenopus* genome. Using the α values with best ML estimates, we calculated species divergence time and ancestral population size.

Firstly, chi-square tests for all chromosomes showed no significant differences between observed synonymous divergence against a negative binomial distribution (all p > 0.01; S3 Table). The only exception was when data from all chromosomes were compiled (chi-square p-value < 0.01), likely due to intra-chromosomal variation, and thus genome-wide estimates were subsequently not conducted.

The estimated species divergence time of *t* = 38.7–47.0 mya (XLA.S-XLA.L, range among chromosomes) and *t* = 44.8–64.3 mya (XTR-XLA) (S2 Table) is older than the previous estimates (using genomic data) of 34 and 48 mya, respectively [6], but similar to some other studies. For example, based on mitochondrial DNA, the common ancestor of *Xenopus* frogs was dated back to 31.8–54.6 mya [24]. In addition, the divergence time between *Xenopus* and *Silurana* (based on DNA, morphology, and fossil calibration) was estimated to be 27–51 mya [25]. Furthermore, simple molecular clock estimates date the *X. laevis–X. tropicalis* divergence and genome duplication in *X laevis* at around 50 and 40 mya, respectively [26].

The slight discrepancies in *t* could be attributed to the past assumption that divergence of sequence is equivalent to species divergence time, but actually N_A can influence divergence time of sequences. Our study using a maximum likelihood approach has accounted for the impact of potentially large N_A on estimations of divergence time and substitution rate. From the ML estimates, we also calculated an ancestral population size of $N_A = 0.88–2.58 \times 10^6$ (XLA.S-XLA.L, range among chromosomes) and $N_A = 0.96–4.02 \times 10^6$ (XTR-XLA) (S2 Table). This is an unexpectedly large estimated population size, considering that it is comparable to that of fruitflies ($N_e = 2–7 \times 10^6$), and at least one order of magnitude higher than lizards and bony fish [27]. In extant humans, the estimated effective population is much lower at below 10^4 [28, 29].

Even though naturally-occurring species hybridization is restricted to a few *Xenopus* spp. (including *X. laevis* and excluding *X. tropicalis*) [30, 31], it seems possible that population structure and local hybridization may have contributed to a large estimated population size. Indeed, the hybridization of *X. laevis* (L) and (S) genomes as well as the fusion of chromosomes 9 and 10 and various intrachromosomal recombinations may have contributed to high population size estimations. Recently, there is emerging evidence supporting that hybridization may have an important role in adaptation [32–34].

The seemingly large population size is based on our large ML estimates of *X*, which is dependent on both population size and mutation rate. Therefore, high mutation rate may be contributive to the large *X* estimate. We briefly examined variation in *μ* using the ML estimates of $Y = 2μt$ and constant species divergence time (48 mya and 34 mya for divergence times of *X. laevis*—*X. tropicalis* and *X. laevis* L–*X. laevis* S, respectively [6]) and found that *μ* varied between chromosomes (1.85–2.59 and 2.10–2.59 x10^{-9} substitutions/site/year, respectively). However, no clear increase was evident, supporting that large population size instead of mutation rate more likely contributed to the large ML estimates of *X*. Future investigation of genomic sequences of other *Xenopus* spp. as well as *X. tropicalis* or *X. laevis* sub-populations will aid in elucidating the mechanisms behind these large estimations. Indeed, the *Xenopus* topology and taxonomy validated using morphological and molecular analyses (mitochondrial and nuclear genes) [5], with a large diversity of polyploidy and high number of independent polyploidization events, could be further explored using genomics. Moreover, genomic studies of other species within the sub-genus *Silurana*, including *X. mellotropicalis* and *X. epitropicalis* [35], will be important to further understand the relationship between chromosomal rearrangements and mutation rates before and after polyploidization.

Conclusions

Using available genome data, we have demonstrated the presence of synonymous substitution rate heterogeneity within *Xenopus* frogs at the inter- and intra-chromosomal level. We found that chromosome 10 had higher *k* compared to other chromosomes, as well as peculiar intrachromosomal patterns, likely related to the fusion of chromosomes 9 and 10 in *X. laevis*. Within most other chromosomes, we identified a pattern of higher *k* at both the distal and proximal ends of the chromosomes, and this may be caused by more frequent recombination and elevated mutation rates in these regions. In addition, maximum likelihood estimations provided additional evidence of rate heterogeneity across all chromosomes. Our estimated species divergence times are a little different to that of a previous study, possible because our approach has accounted for ancestral population size. Our study is one of the first to estimate ancestral population size in *Xenopus* frogs, which we found to be surprisingly high. This large population size may be attributed to hybridization and population structure and thus warrants further investigation to validate our findings.

Supporting information

S1 Fig. Relationship between proportion of synonymous substitutions (*k*) and location within each chromosome, for each genome pair combination (XTR-XLA.L in red, XTR-XLA.S in green, and XLA.L-XLA.S in blue). Representative plots are presented in Fig 2. Centromere locations (vertical dotted line) are based on [15, 16].
(DOCX)

S2 Fig. Frequency of CG dinucleotides along each chromosome based on whole genome data of *Xenopus tropicalis* and *X. laevis* L and S. Y-axes represent CG composition in 1Mb

sliding windows with a 100kb step on the chromosomes, X-axes represent chromosome locations (x1000000). The dinucleotide composition was computed using a batch-learning self-organizing map (BLSOM) program (Abe et al. 2003). The BLSOM program can be obtained from UNTROD, Inc. (y_wada@nagahama-i-bio.ac.jp).
(DOCX)

S3 Fig. Unique patterns observed between proportion of synonymous substitutions *k* and location within (A) chromosome 8, (B) chromosome 3 (normal pattern), and (C) chromosome 9/10 could be attributed to chromosomal rearrangements previously identified using cytogenetic mapping by Session et al. [5].
(DOCX)

S1 Table. Quadratic regression analyses of relationship between W = chromosome location and Z = proportion of synonymous substitutions (*k*). Partial F-test was used to compare two models: linear [lm($W{\sim}Z$)] and quadratic [lm($W{\sim}Z + I (Z{^{\wedge}}2)$)]. Values presented are significance probabilities associated with the F values, Pr(F), and were calculated independently for each pairwise comparison of genomes: XLA.L-XLA.S (blue), XTR-XLA.L (red), and XTR-XLA.S (green). Values with *Pr(F)* < 0.01 (in bold) significantly support a quadratic model.
(DOCX)

S2 Table. Maximum likelihood estimates and calculations of rate heterogeneity (α), ancestral population size (NA), and divergence time (t). Although chromosome 9_10 is homoeologous in XLA.L and XLA.S, we separated estimations to chromosome 9 or 10 based on location in XTR.
(DOCX)

S3 Table. Chi-square p-values following fitting observed data (synonymous difference) against the distribution 'nbinom' by maximum likelihood.
(DOCX)

Acknowledgments

Author Contributions

Conceptualization: Quintin Lau, Takeshi Igawa, Yoko Satta.

Formal analysis: Quintin Lau, Takeshi Igawa, Yoko Satta.

Funding acquisition: Hajime Ogino, Yoko Satta.

Investigation: Quintin Lau, Takeshi Igawa, Yukako Katsura, Toshimichi Ikemura.

Supervision: Hajime Ogino, Yoko Satta.

Writing – original draft: Quintin Lau.

Writing – review & editing: Takeshi Igawa, Hajime Ogino, Yukako Katsura, Toshimichi Ikemura, Yoko Satta.

References

1. Satta Y, Klein J, Takahata N. DNA archives and our nearest relative: The trichotomy problem revisited. Mol Phylogenet Evol. 2000; 14: 259–275. https://doi.org/10.1006/mpev.2000.0704 PMID: 10679159

2. O'hUigin C, Satta Y, Takahata N, Klein J. Contribution of homoplasy and of ancestral polymorphism to the evolution of genes in anthropoid primates. Mol Biol Evol. 2002; 19: 1501–1513. https://doi.org/10.1093/oxfordjournals.molbev.a004213 PMID: 12200478

3. Chen F-C, Li W-H. Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. Am J Hum Genet. 2001; 68: 444–456. https://doi.org/10.1086/318206 PMID: 11170892

4. Bewick AJ, Chain FJJ, Heled J, Evans BJ. The Pipid Root. Syst Biol. 2012; 61: 913–926. https://doi.org/10.1093/sysbio/sys039 PMID: 22438331

5. Evans BJ, Carter TF, Greenbaum E, Gvodík V, Kelley DB, McLaughlin PJ, et al. Genetics, morphology, advertisement calls, and historical records distinguish six new polyploid species of African clawed frog (Xenopus, Pipidae) from West and Central Africa. PLoS One. 2015; 10: e0142823. https://doi.org/10.1371/journal.pone.0142823 PMID: 26672747

6. Session AM, Uno Y, Kwon T, Chapman JA, Toyoda A, Takahashi S, et al. Genome evolution in the allo-tetraploid frog Xenopus laevis. Nature. 2016; 538: 336–343. https://doi.org/10.1038/nature19840 PMID: 27762356

7. Crawford AJ. Relative rates of nucleotide substitution in frogs. J Mol Evol. 2003; 57: 636–641. https://doi.org/10.1007/s00239-003-2513-7 PMID: 14745532

8. Soderlund C, Bomhoff M, Nelson WM. SyMAP v3.4: A turnkey synteny system with application to plant genomes. Nucleic Acids Res. 2011; 39: e68. https://doi.org/10.1093/nar/gkr123 PMID: 21398631

9. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018; 35: 1547–1549. https://doi.org/10.1093/molbev/msy096 PMID: 29722887

10. Yang Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. J Mol Evol. 1994; 39: 306–314. https://doi.org/10.1007/BF00160154 PMID: 7932792

11. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria, Austria; 2016. Available: https://www.r-project.org

12. Uno Y, Nishida C, Takagi C, Ueno N, Matsuda Y. Homoeologous chromosomes of Xenopus laevis are highly conserved after whole-genome duplication. Heredity (Edinb). 2013; 111: 430–436. https://doi.org/10.1038/hdy.2013.65 PMID: 23820579

13. Hellmann I, Ebersberger I, Ptak SE, Pääbo S, Przeworski M. A neutral explanation for the correlation of diversity with recombination rates in humans. Am J Hum Genet. 2003; 72: 1527–1535. https://doi.org/10.1086/375657 PMID: 12740762

14. Spencer CCA, Deloukas P, Hunt S, Mullikin J, Myers S, Silverman B, et al. The influence of recombination on human genetic diversity. PLoS Genet. 2006; 2: e148. https://doi.org/10.1371/journal.pgen.0020148 PMID: 17044736

15. Fang L, Ye J, Li N, Zhang Y, Li SG, Wong GKS, et al. Positive correlation between recombination rate and nucleotide diversity is shown under domestication selection in the chicken genome. Chinese Sci Bull. 2008; 53: 746–750. https://doi.org/10.1007/s11434-008-0159-y

16. Fullerton SM, Carvalho AB, Clark AG. Local rates of recombination are positively correlated with GC content in the human genome. Mol Biol Evol. 2001; 18: 1139–1142jen. https://doi.org/10.1093/oxfordjournals.molbev.a003886 PMID: 11371603

17. Jensen-Seaman MI, Furey TS, Payseur BA, Lu Y, Roskin KM, Chen CF, et al. Comparative recombination rates in the rat, mouse, and human genomes. Genome Res. 2004; 14: 528–538. https://doi.org/10.1101/gr.1970304 PMID: 15059993

18. Charlesworth B, Lapid A, Canada D. The distribution of transposable elements within and between chromosomes in a population of Drosophila melanogaster. II. Inferences on the nature of selection against elements. Genet Res. 1992; 60: 115–130. https://doi.org/10.1017/s0016672300030809 PMID: 1334900

19. Rizzon C, Marais G, Gouy M, Biémont C. Recombination rate and the distribution of transposable elements in the Drosophila melanogaster genome. Genome Res. 2002; 12: 400–407. https://doi.org/10.1101/gr.210802 PMID: 11875027

20. Shen JJ, Dushoff J, Bewick AJ, Chain FJJ, Evans BJ. Genomic dynamics of transposable elements in the Western clawed frog (Silurana tropicalis). Genome Biol Evol. 2013; 5: 998–1009. https://doi.org/10.1093/gbe/evt065 PMID: 23645600

21. Edwards NS, Murray AW. Identification of Xenopus CENP-A and an associated centromeric DNA repeat. Mol Biol Cell. 2005; 16: 1800–1810. https://doi.org/10.1091/mbc.e04-09-0788 PMID: 15673610

22. Khokha MK, Krylov V, Reilly MJ, Gall JG, Bhattacharya D, Cheung CYJ, et al. Rapid gynogenetic mapping of Xenopus tropicalis mutations to chromosomes. Dev Dyn. 2009; 238: 1398–1346. https://doi.org/10.1002/dvdy.21965 PMID: 19441086

23. Axelsson E, Webster MT, Smith NGC, Burt DW, Ellegren H. Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. Genome Res. 2005; 15: 120–125. https://doi.org/10.1101/gr.3021305 PMID: 15590944

24. Evans BJ, Kelley DB, Tinsley RC, Melnick DJ, Cannatella DC. A mitochondrial DNA phylogeny of African clawed frogs: Phylogeography and implications for polyploid evolution. Mol Phylogenet Evol. 2004; 33: 197–213. https://doi.org/10.1016/j.ympev.2004.04.018 PMID: 15324848

25. Cannatella D. Xenopus in space and time: fossils, node calibrations, tip-dating, and paleobiogeography. Cytogenet Genome Res. 2015; 145: 283–301. https://doi.org/10.1159/000438910 PMID: 26279165

26. Hellsten U, Khokha MK, Grammer TC, Harland RM, Richardson P, Rokhsar DS. Accelerated gene evolution and subfunctionalization in the pseudotetraploid frog Xenopus laevis. BMC Biol. 2007; 5: 31. https://doi.org/10.1186/1741-7007-5-31 PMID: 17651506

27. Nei M, Graur D. Extent of protein polymorphism and the neutral mutation theory. Evolutionary Biology. 1984. pp. 73–118. https://doi.org/10.1007/978-1-4615-6974-9_3

28. Takahata N. Allelic genealogy and human evolution. Mol Biol Evol. 1993; 10: 2–22. https://doi.org/10.1093/oxfordjournals.molbev.a039995 PMID: 8450756

29. Tenesa A, Navarro P, Hayes BJ, Duffy DL, Clarke GM, Goddard ME, et al. Recent human effective population size estimated from linkage disequilibrium. Genome Res. 2007; 17: 520–526. https://doi.org/10.1101/gr.6023607 PMID: 17351134

30. Furman BLS, Cauret CMS, Colby GA, John Measey G, Evans BJ. Limited genomic consequences of hybridization between two African clawed frogs, Xenopus gilli and X. laevis (Anura: Pipidae). Sci Rep. 2017; 7: 1091. https://doi.org/10.1038/s41598-017-01104-9 PMID: 28439068

31. Fischer WJ, Koch WA, Elepfandt A. Sympatry and hybridization between the clawed frogs Xenopus laevis laevis and Xenopus muelleri (Pipidae). J Zool. 2000; 252: 99–107. https://doi.org/10.1017/S0952836900009110

32. Chen C, Pfennig KS. Female toads engaging in adaptive hybridization prefer high-quality heterospecifics as mates. Science. 2020; 367: 1377–1379. https://doi.org/10.1126/science.aaz5109 PMID: 32193328

33. Oziolor EM, Reid NM, Yair S, Lee KM, VerPloeg SG, Bruns PC, et al. Adaptive introgression enables evolutionary rescue from extreme environmental pollution. Science. 2019; 364: 455–457. https://doi.org/10.1126/science.aav4155 PMID: 31048485

34. Jones MR, Scott Mills L, Alves PC, Callahan CM, Alves JM, Lafferty DJR, et al. Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. Science. 2018; https://doi.org/10.1126/science.aar5273 PMID: 29930138

35. Knytl M, Smolík O, Kubíčková S, Tlapáková T, Evans BJ, Krylov V. Chromosome divergence during evolution of the tetraploid clawed frogs, Xenopus mellotropicalis and Xenopus epitropicalis as revealed by Zoo-FISH. PLoS One. 2017; https://doi.org/10.1371/journal.pone.0177087 PMID: 28545147