Research

Extensive variation in germline de novo mutations in *Poecilia reticulata*

Yuying Lin,[1] Iulia Darolti,[2] Wouter van der Bijl,[1] Jake Morris,[3] and Judith E. Mank[1]

[1]*Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada;* [2]*Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland;* [3]*School of Biological Science, University of Bristol, Bristol BS8 1TQ, United Kingdom*

The rate of germline mutation is fundamental to evolutionary processes, as it generates the variation upon which selection acts. The guppy, *Poecilia reticulata*, is a model of rapid adaptation, however the relative contribution of standing genetic variation versus de novo mutation (DNM) to evolution in this species remains unclear. Here, we use pedigree-based approaches to quantify and characterize germline DNMs in three large guppy families. Our results suggest germline mutation rate in the guppy varies substantially across individuals and families. Most DNMs are shared across multiple siblings, suggesting they arose during early embryonic development. DNMs are randomly distributed throughout the genome, and male-biased mutation rate is low, as would be expected from the short guppy generation time. Overall, our study shows remarkable variation in germline mutation rate and provides insights into rapid evolution of guppies.

[Supplemental material is available for this article.]

Mutations underlie much of the diversity of life on earth. Moreover, the timing, rate, and spectra of germline mutations are fundamental aspects of molecular evolution and the capacity of organisms to adapt (Lynch et al. 2016). The rate of mutations is a product of several processes. De novo mutations (DNMs) arise from errors that DNA damage repair mechanisms fail to correct (Gao et al. 2019; Seplyarskiy et al. 2021; De Manuel et al. 2022), and so the fidelity of replication and the efficacy of repair mechanisms are major contributors to the DNM rate. Although all cell divisions, somatic and germline, experience DNMs, only germline mutations can be passed on to the next generation, and the rate of DNMs is far higher in somatic versus germline tissue (Milholland et al. 2017), likely representing different investment in costly repair mechanisms.

Accurate mutation rate estimates are critical for phylogenetic and systematic studies. However, because mutations are rare events, it can be difficult to identify their timing and rate of occurrence. Recent sequencing-based approaches to study germline mutation have used pedigrees to characterize germline DNMs, and revealed the influence of many factors, including generation time (Francioli et al. 2015; Kong et al. 2012; Carlson et al. 2018; Wang et al. 2022; Bergeron et al. 2023), life history (Bergeron et al. 2023), and number of cell divisions (Ellegren 2007). However, these factors do not fully explain DNM variation across species (Wu et al. 2020; Campbell et al. 2021; De Manuel et al. 2022), and recent work in humans suggests stochastic biological processes and, to a lesser extent, family-specific effects (Goldmann et al. 2021) may be important. Together, these findings indicate that a complex interplay of multiple factors contributes to the variation in the number of DNMs in families.

The guppy, *Poecilia reticulata*, is a model system for rapid ecological adaptation (Reznick et al. 1990, 1997; Gordon et al. 2015; Whiting et al. 2022) and the swift and convergent patterns of phenotypic adaptation in guppies can arise via two possible genetic mechanisms in the absence of hybridization (Stern 2013; Lee and Coop 2017). It is possible that DNMs underlie the rapid rates of phenotypic adaptation observed in guppies. However, functional mutations are more often deleterious than adaptive (Yoder and Tiley 2021), and a high germline rate of DNM fostering rapid adaptation could ultimately prove more of a mutational burden than an adaptive boon. Additionally, because somatic mutation rates are strongly anticorrelated with life span (Cagan et al. 2022), which is relatively short in guppies, on average 2–3 yr in captivity (Reznick et al. 2006), and because poikilothermic organisms experience lower rates of DNA damage (Adelman et al. 1988), we might expect low mutation rates in guppies. Consistent with this, DNM rates in fish are slightly lower than homeothermic vertebrates (Feng et al. 2017; Bergeron et al. 2023).

Alternatively, rapid adaptation could be a product of selection on standing genetic variation (Barrett and Schluter 2008), and consistent with this, natural populations of guppies show extensive genetic and phenotypic polymorphism (Almeida et al. 2021; Whiting et al. 2021, 2022; Lin et al. 2022). The role of DNMs versus standing genetic variation in the rapid adaptation of guppies has important implications to the locus and nature of evolution in this key ecological model, as well as what type of molecular signature we might expect to detect. Selection on adaptive DNMs will produce a signature of hard sweeps and would most often be associated with a few loci of large effect (Pritchard et al. 2010; Matuszewski et al. 2015). In contrast, selection acting on standing variation is associated with soft sweeps, and will more often result in fixing many alleles of small effect (Hermisson and Pennings 2017).

The wealth of ecological data available for guppies, coupled with the short generation time (3 mo) and ease of rearing controlled pedigrees in the laboratory, makes them an ideal system

Corresponding author: linyuying@zoology.ubc.ca
Article published online before print. Article, supplemental material, and publication date are at https://www.genome.org/cgi/doi/10.1101/gr.277936.123.

for studying rates of DNMs, and the potential role of these rates in adaptation. Here, we sequenced three large *P. reticulata* families, reconstructed a high-quality genome from an individual collected from Quare River in Trinidad, where our laboratory stock guppies derive, and applied stringent parent-offspring trio-based variant filtering criteria to accurately identify and quantify the number of DNMs, providing insights into the role of germline mutations in the rapid evolution of this species. Moreover, our large families, each with ten offspring, allow us to estimate the timing of mutations across germline development, revealing the mosaic nature of the germline genome through developmental mutations (Goldmann et al. 2019).

Results

Identifying DNMs from three guppy pedigrees

We resequenced 36 individuals from three guppy pedigrees, each including unrelated parents, five sons, and five daughters (Supplemental Figs. S1–S4), with individual mean sequencing depth ≈ 25× (Table 1; Supplemental Table S1). To limit the impact of genetic differences between the Ensembl female reference genome and the guppy populations in this study, we reconstructed a high-quality pseudogenome with assembly contiguity metric auN = 26897868 and the number of Ns per 100 kb = 2949, from a 10x assembly of a female individual taken from the Quare River in Trinidad (see Almeida et al. 2021 for details), the original population from which our laboratory stock population derives (for details, see Methods).

Germline mutations are rare events, and as such are notoriously difficult to differentiate from the various errors introduced from sequencing, read alignments, genotyping, and variant filtering. Thus, we applied step-by-step stringent filtering to minimize false discovery rates (Supplemental Fig. S5; Supplemental Table S2). First, we discarded all multiply mapped reads, retaining only uniquely mapped reads for further analysis. Second, we genotyped each parent-offspring trio separately, instead of all offspring in each family together, using two independent genotype callers, GATK v.4.2.6.1 (Van der Auwera and O'Connor 2020) and BCFtools v.1.16 (Li et al. 2009), and used the intersection of inferred DNMs from both methods. We called sites with Mendelian violations, where the parents are homozygous for the same allele and the child is heterozygous, as putative DNMs. By applying the same individual-level and site-level filtering criteria, and intersecting the results from the two genotype callers, we obtained an average number of 5.20, 24.90, and 4.80 DNMs across each pedigree (Fam1, Fam2, Fam3) (Table 1; Supplemental Table S1). To calculate the germline mutation rate in each trio, we determined the denominator, the callable genome size, of each individual (see Methods). After removing repetitive regions and restricting the coverage > ½× or <2× of the individual average sequencing depth, we retained an approximate callable genome size of ≈430 Mb out of the 720 Mb reference genome across individuals (Table 1; Supplemental Tables S1, S3).

We used the Integrative Genomics Viewer (IGV) (Thorvaldsdóttir et al. 2013) to visualize read alignments and remove false positive DNMs, defined as those missing in the genotypes because of local realignments and those in highly variable genomic regions (see Methods). On average, 17% of initial DNMs were removed as false positives (Supplemental Tables S4–S6). To detect the false negative rate by simulation, we inserted 1000 artificial DNMs to our read data. Our pipeline yielded an 89.5% detection rate, suggesting a 10.5% potential false negative rate, representing a good balance between Type I and Type II error.

Distribution and variation of DNMs among families

After stringent variant filtering processes (see Methods), we observed a large variation in DNM rate across guppy pedigrees, with an average DNM rate of 0.60×10^{-8}, 2.90×10^{-8}, and 0.56×10^{-8} per nucleotide per generation in Fam1, Fam2, and Fam3, respectively (Table 1; Supplemental Table S1; Fig. 1A). The average mutation rate in Fam2 is significantly different from those in Fam1 (Wilcoxon rank-sum test, $P = 2 \times 10^{-4}$) and Fam3 (Wilcoxon rank-sum test, $P = 2 \times 10^{-4}$) (Table 1; Fig. 1A).

Even though the proportion of DNMs in coding sequence was far lower than those in noncoding sequence (Fig. 1B), this was nonsignificant after correcting for vast differences in callable genome size between these categories (Fisher's exact test, $P = 0.38$, Fig. 1B). We observed significantly more transitions associated with CpG sites, including C > T transitions (χ^2 test, adjusted $P = 3.71 \times 10^{-5}$, Fig. 1C; Supplemental Table S7) and the reverse complement G > A transitions (χ^2 test, adjusted $P = 1.19 \times 10^{-4}$, Fig. 1C; Supplemental Table S7), consistent with findings in other species (Goldmann et al. 2016; Jónsson et al. 2017; Kessler et al. 2020). Because DNMs are rare, some types of mutations, including G > C, G > T, T > A, T > C, T > G, are absent by chance in different families in our study. However, there is no significant difference in mutation spectra across families (χ^2 test, $P = 0.20$, Fig. 1C).

Theoretically, DNMs can occur at any stage of germline development as well as through the process of gametogenesis itself. DNMs occurring early in germline development are shared between a greater proportion of daughter cells compared to those occurring later or in gametogenesis itself, and are therefore more likely to be shared among siblings. DNMs occurring later, or in gametogenesis itself, would result in a small proportion of affected gametes and would not be expected to be shared across multiple siblings. Importantly, averaging across all our families, less than half of DNMs are observed in single offspring (Figs. 1D, 2; Supplemental Tables S1, S4–S6, S8). This high proportion of DNMs shared by at least two siblings suggests that many DNMs result from replication errors during germline development (Goldmann et al. 2019). Importantly, Fam2 has a higher proportion of shared DNMs than either Fam1 (χ^2 test, $P = 1.83 \times 10^{-3}$) or Fam3 (χ^2 test, $P = 8.70 \times 10^{-2}$), suggesting that the higher overall

Table 1. Summary statistics of DNMs in three guppy pedigrees

Pedigree	Coverage[a]	Mapping Rate[a]	Callable Genome Size[a]	No. of DNMs (range)	μ, 10^{-8} per site per generation (range)
Fam1	26.42×	99.36%	432,137,349 bp	5.20 (2.00–9.00)	0.60 (0.23–1.27)
Fam2	24.57×	99.37%	428,029,972 bp	24.90 (21.00–31.00)	2.90 (2.45–3.62)
Fam3	26.10×	99.31%	429,364,825 bp	4.80 (2.00–6.00)	0.56 (0.24–0.71)

[a]All data are averaged across offspring in each pedigree; details can be found in Supplemental Tables S1, S6–S9.

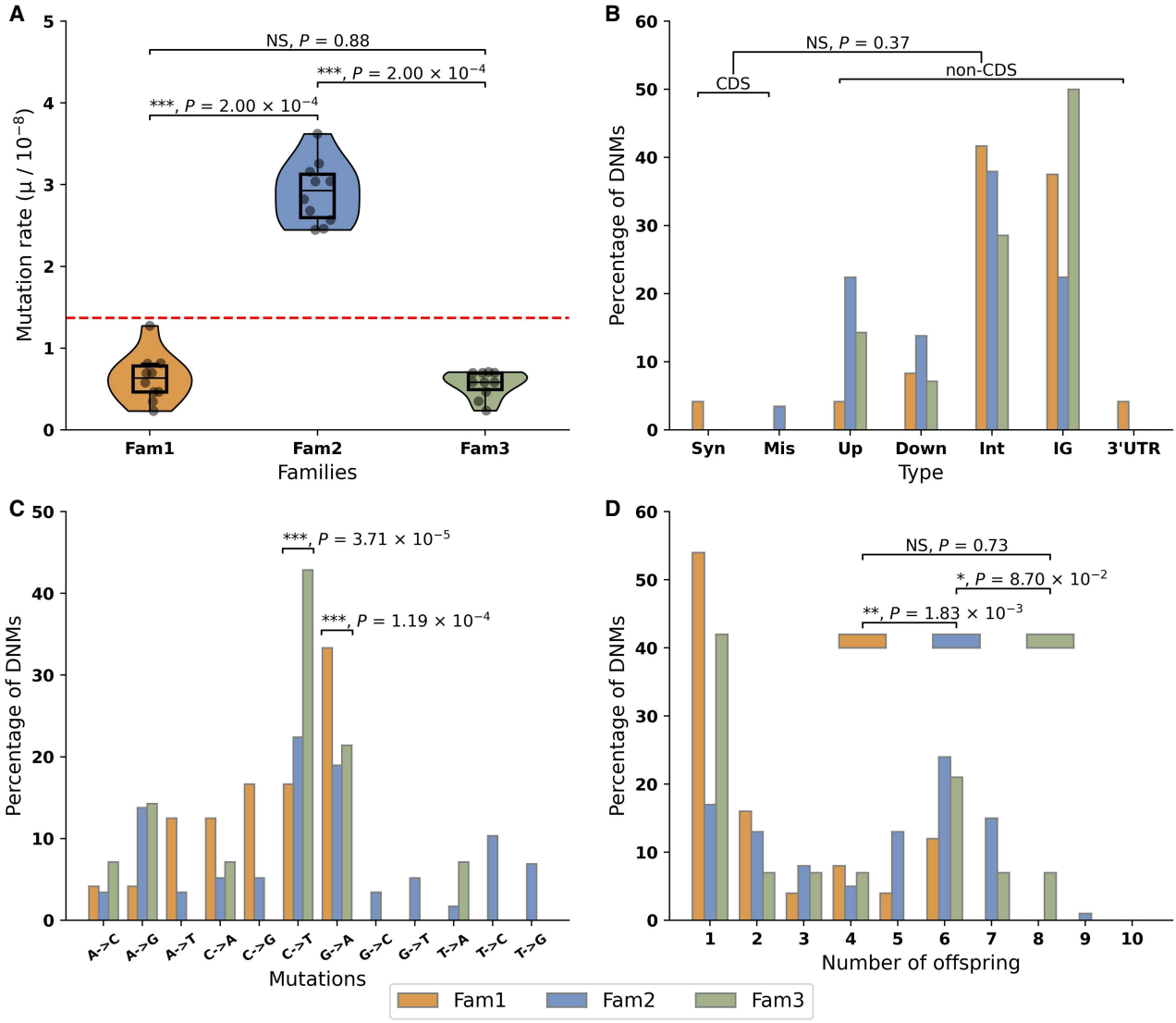

Figure 1. Summary statistics of DNMs in three guppy pedigrees. (*A*) Mutation rate in three families. Red line represents the average mutation rate across three families. Asterisks indicate statistical difference among families (Wilcoxon signed-rank test, [***] $P<0.001$; NS: nonsignificant). Orange: Fam1; blue: Fam2; green: Fam3. (*B*) Functional distribution of DNMs. There was no statistical difference (Fisher's exact test, NS: nonsignificant) in the proportion of DNMs between coding sequences (CDS), including synonymous mutations (Syn) and missense mutations (Mis), and noncoding sequence (nonCDS), including upstream (Up), downstream (Down), intron (Int), intergenic (IG), and 3′ UTR variant (3′ UTR) once correcting for callable genome size. (*C*) Mutation spectrum. Asterisks on the *top* of each category indicate significant difference to other categories (χ^2 test, [***] $P<0.001$). (*D*) Shared mutations among siblings: 45.73%, 72.76%, and 57.14% of DNMs are shared by at least two siblings in Fam1, Fam2, and Fam3, respectively, likely reflecting mutations occurring during early germline development. DNMs observed in only a single offspring represent 54.17%, 17.24%, and 42.86% of all DNMs in Fam1, Fam 2, and Fam3, respectively, and these most likely represent mutations after the onset of gametogenesis (Goldmann et al. 2019). The relative proportion of DNMs in a single offspring and DNMs shared by multiple siblings in Fam2 is significantly different from other two families (χ^2 test, [**] $P<0.01$; [*] $P<0.1$; NS: nonsignificant).

DNM rate in this family is largely the product of more early developmental mutations.

The greater number of cell divisions associated with spermatogenesis is expected in many species to result in a male-mutation bias (Ellegren 2007), although we might expect male-biased mutation to be relatively low in guppies because of their short overall generation time, typically 3 mo (Reznick et al. 1997). To estimate male mutation bias in our guppy families, we phased each individual and matched offspring haplotype blocks bearing DNMs to each parent haplotype to determine parent of origin, omitting DNMs with ambiguous phasing. Our data show a consistent male-to-female mutation rate ratio across all three guppy families of ≈1.5 (Supplemental Tables S4–S6, S9).

Germline mutation rates can vary across the genome, depending on contextual genetic background (Carlson et al. 2018). The DNMs we observe are statistically random in terms of genomic distribution (Fig. 2, χ^2 test, see Supplemental Table S3).

Discussion

Germline DNMs are passed on from parent to offspring, and represent a major source of genetic variation upon which evolution

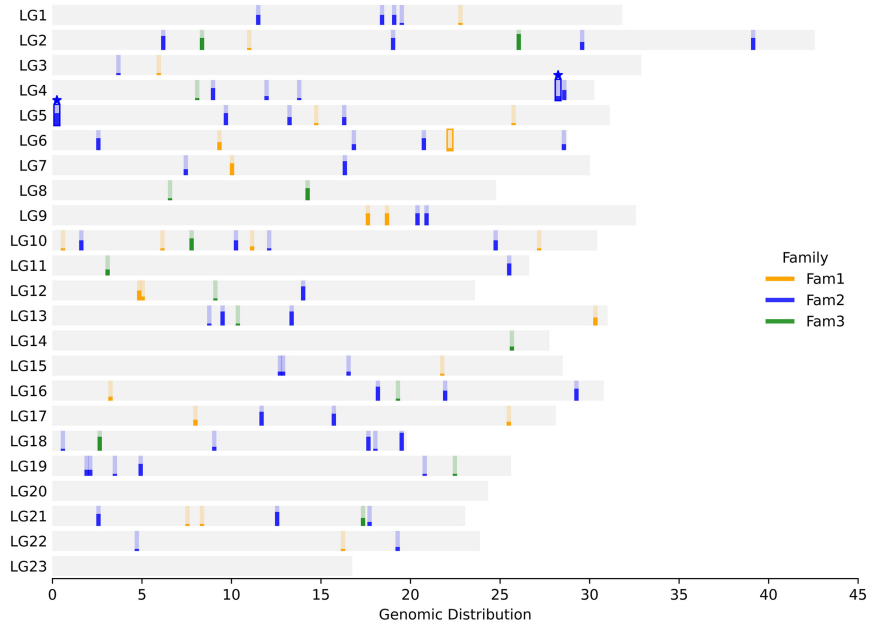

Figure 2. Genomic distribution of DNMs in three guppy families for each chromosomal linkage group (LG). The height of each bar represents the number of offspring sharing the same DNM in the family. DNMs with an outline represent those in coding regions. Stars (★) indicate nonsynonymous mutations. No chromosome shows significant deviation from the expected number of DNMs based on its proportion of callable sites (χ^2 test, see details in Supplemental Table S6). Fam1 (orange), Fam2 (blue), and Fam3 (green).

acts. As such, the rate of germline DNMs is critical for defining the adaptive potential and mutational burden of a population. Here we assessed three large pedigrees for DNMs in the guppy, a long-standing model of rapid ecological adaptation. Guppies can rapidly adapt to shifting ecologies (Reznick et al. 1990), and this raises important questions about whether selection acts in these cases on DNMs or on standing genetic variation. The role of DNMs in guppy adaptation has important implications to the molecular signatures we might expect to detect, namely hard sweeps associated with recent DNMs versus soft sweeps acting on standing genetic variation (Hermisson and Pennings 2005, 2017; Harris et al. 2018).

Overall, our estimate of the germline DNM rate, 1.35×10^{-8}, is roughly comparable to estimates of other vertebrate species (Fig. 3; Supplemental Table S10), suggesting that rapid DNMs may not be the primary locus of selection in rapidly adapting guppy populations. Indeed, relatively few sites in the guppy genome show evidence of the hard sweeps associated with recent DNMs (Fraser et al. 2015), and patterns are more consistent with soft sweeps on standing genetic variation (van der Zee et al. 2022).

Our study design allows us to examine variation in the number of germline mutations across guppy individuals, and we observe a remarkable range in DNM rates across our families (Table 1). There is evidence of variation in germline mutation in some mammals, either as a result of increased paternal age (Kong et al. 2012; Carlson et al. 2018; Wang et al. 2022) or stochastic processes (Goldmann et al. 2021), however the degree of variation observed in these species is far less than we observe in guppies. Our data come from captive reared guppies kept under controlled, consistent conditions, suggesting that the differences in DNM rate that we observe are intrinsic, rather than the result of environmental variation. Moreover, the short generation time (Reznick et al. 1997) and life span (Reznick et al. 2006) in guppies guarantees that the parents in our families varied very little in overall age, suggesting that paternal age may have little effect in this variation we observe.

More importantly in the context of guppies, the wide variation in DNM rate suggests that some lineages, with higher mutation rates, might in fact show more potential for adaption by DNMs than standing variation. Although overall, our results suggest that the rapid adaptation often observed in guppy populations (Reznick et al. 1990, 1997; Gordon et al. 2015; Whiting et al. 2022) is largely the product of selection on standing genetic variation, the extensive variation in DNM rate we observe suggests that if heritable, DNM ratio might vary extensively in natural populations with concurrent variation in adaptive potential. The role of heritable variation in DNM rate and adaptive potential remains an important area for further exploration.

It is worth noting that alternative bioinformatic pipelines have yielded DNM rate estimates that vary by a factor of two for the same parent-offspring trios (Bergeron et al. 2022). This variation represents differences in the balance between stringent filtering to remove false positives, and the minimization of false negatives. Although we were stringent in our filtering and inclusion criteria, our simulations reveal that the pipeline yielded an 89.50% detection rate, suggesting a 10.5% potential false negative rate. Because we used the same pipeline across all families, any error from false negatives or false positives should affect all families similarly, allowing for direct comparisons.

Moreover, compared to other poikilothermic species, the absolute value of the mutation rate in the guppy is similar to whole genome pedigree-based estimates in poikilotherms, and also similar to rates seen in some endothermic primates (Fig. 3; Supplemental Table S10). The high DNM ratio in Fam2 largely explains the extensive range of DNM rate observed here (Fig. 1A; Supplemental Tables S4–S6), and without this family, our estimates would be similar to other fish species (Fig. 3). It might be expected that the lower metabolic rates in poikilothermic species would produce less oxidative damage to the DNA molecule (Adelman et al. 1988; Martin and Palumbit 1993). However, poikilothermy does not itself explain DNM rate variation across vertebrates (Bergeron et al. 2023), suggesting that the role of homeostasis in DNM rate is small.

Our large pedigrees, each comprised of ten offspring, make it possible to assess DNM sharing among siblings, and thus assess the degree of mosaicism in the germline genome (Goldmann et al. 2019). Guppies are live-bearing fish and females store sperm (Evans et al. 2002). Spermatogenesis starts at 3 mo of age (Evans et al. 2002) and continues throughout adulthood (Billard and Escaffre 1969). Oogenesis begins at 3 mo of age and continues asynchronously thereafter (Droller and Roth 1966). Our results suggest that the developmental timing of DNMs in the germline is highly variable (Fig. 1D). If a DNM occurs early in germline development, it will be passed on to a greater proportion of daughter cells than those that occur later in development, and thus it is

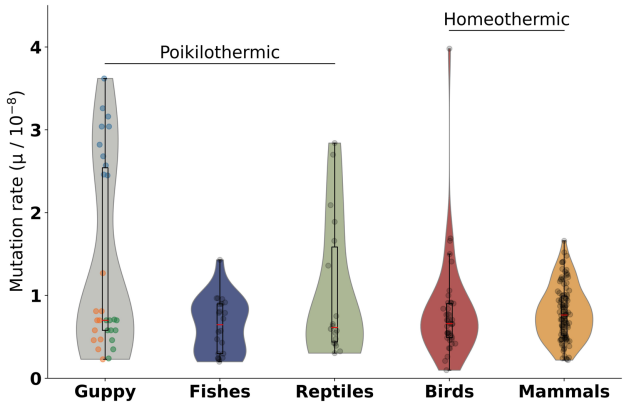

Figure 3. Distribution of trio DNMs in different animals. Distribution of trio DNM rate across different groups of poikilothermic animals, including fishes and reptiles, and homeothermic animals, including birds and mammals. Each data point represents a DNM rate in a parent-offspring trio. Three families in the guppy categories: Fam1 (orange), Fam2 (blue), and Fam3 (green). More details can be found in Supplemental Table S10. Data points from Awadalla et al. (2010); Bergeron et al. (2022, 2023); Besenbacher et al. (2015); Campbell et al. (2021); Conrad et al. (2011); Feng et al. (2017); Harland et al. (2017); Jónsson et al. (2017); Koch et al. (2019); Kong et al. (2012); Lindsay et al. (2019); Martin et al. (2020); Pfeifer (2017); Rahbari et al. (2016); Roach et al. (2010); Smeds et al. (2016); Tatsumoto et al. (2017); Thomas et al. (2018); Venn et al. (2014); Wang et al. (2020, 2022); Wu et al. (2020).

more likely to be transmitted to the next generations and shared by multiple siblings. Although the number of shared mutations varied among our three families, ranging from 50% to 73%, it was nonetheless the majority of DNMs, indicating most inherited DNMs occur during germline development instead of after the onset of gametogenesis. This is consistent with studies in smaller human families which nonetheless revealed substantial sharing of DNMs among siblings and variation in the proportion among families (Rahbari et al. 2016).

We also assessed the male mutation bias in DNM generation. Male mutation bias occurs from the increased number of cell divisions in spermatogenesis compared to oogenesis (Ellegren 2007). This male-biased mutation rate scales with increasing parental age (Kong et al. 2012; Francioli et al. 2015; Carlson et al. 2018), and the relatively short life span of guppies might suggest a minimal male mutation bias. Indeed, our estimate of ≈1.5 is on par with that of other short-lived species (Wilson Sayres and Makova 2011).

Overall, our results indicate that the developmental timing, rate, and spectrum of DNMs varies substantially across families in guppies, suggesting that mutation rate as a trait can vary across at the level of individuals, families, and populations and might be able to evolve under different conditions (Lynch et al. 2016). Our results also suggest that comparisons across species based on just one or a few parent-offspring trios may experience high error because of intraspecific variation. Taken together, our results present an important advance in our understanding of the variation in germline DNMs, and their contribution to adaption.

Methods

Samples and DNA sequencing

We selected three families from a previous laboratory experiment (Morris et al. 2020), each of which contains one father, one mother, five sons, and five daughters. During the experiment setup, all the fish were kept in the same flow-through aquarium rack, and experienced the same feeding and temperature regimes in the laboratory, as well as the same water within the cell flow system. All the fathers and mothers were paired when they were 4 mo old. We sampled offspring when they were sexually mature, which is roughly 4 mo of age, from two consecutive clutches. In total, 36 samples from three pedigrees were collected. DNA was extracted from whole head tissues with Qiagen DNeasy Blood & Tissue Kit, following the manufacturer protocol. We prepared the shotgun library with IDT dual-index adaptor. After DNA quality control, all the samples were sequenced individually on the Illumina NovaSeq 4000 platform.

Reference genome reconstruction

Divergence between the individual of the Ensembl reference genome v.99 (Künstner et al. 2016) and our laboratory population might trigger false positive DNM calls and complicate downstream DNM detection. Our laboratory population originated from a high predation population of the Quare River in Trinidad. We, therefore, selected the most contiguous 10x Genomics Linked Read female genome assembly from the Quare high predation population, presented in Almeida et al. (2021).

We customized our reference genome first by assembling the genome using LongRanger v2.1.1 (10x Genomics), removing redundant short sequences for which 95% of contigs overlapped with longer ones. We further improved the assembly using a *k*-mer based approach, ARKS + LINKS v1.0.4 (Coombe et al. 2018) and anchored scaffolds into chromosomes using RAGTAG v2.1.0 (Alonge et al. 2019, 2022). Genome completeness was assessed by QUAST v5.1.0 (Mikheenko et al. 2018) and BUSCO v 5.3.2 (Simão et al. 2015), see details in Supplemental Tables S11 and S12.

We performed de novo prediction and modeling of transposable element (TE) families using RepeatModeler2 (Flynn et al. 2020) with default parameters in both the Ensembl reference genome and the reconstructed genome. Repetitive sequences were then identified by RepeatMasker v.4.1.1 (Smit 1996) using databases outputted by RepeatModeler2 (Flynn et al. 2020). In total, 27.04% of sequences were identified as repetitive sequences and thus were removed (see details in Supplemental Table S13).

Alignment, genotyping, and SNP filtering for DNM detection

We used FastQC v0.11 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Trimmomatic v0.36 (Bolger et al. 2014) to remove adapter sequences and low-quality reads. After quality control, we recovered ~25× average read depth per sample. High-quality reads were aligned to the reconstructed female genome, using BWA-MEM v0.7.15 (Li and Durbin 2009) with default parameters. We filled in mate coordinates and mate-related flags, sorted alignment by coordinates, and marked PCR duplicates with Picard (v2.0.1; https://broadinstitute.github.io/picard/).

To reduce the chance of removing true DNMs in each individual, we conducted parent-offspring trio-based genotyping instead of the joint genotyping with all offspring in a family. We called genotypes with two independent haplotype-aware software, GATK4 and BCFtools, using only those sites called by both approaches (see details in Supplemental Table S2).

First, we applied hard filtering to the raw SNP data set in trios following the GATK best practice "QD < 2.0, FS > 60.0, SOR > 3.0, MQ < 40.0, MQRankSum < −12.5 and ReadPosRankSum < −8.0" for raw GATK4 genotype data set and "MQBZ < −3 || RPBZ < −3 || RPBZ > 3 || FORMAT/SP > 32 || SCBZ > 3 || TYPE =

"INDEL"" for raw BCFtools genotype data set. We selected biallelic SNPs in which genotypes are homozygous in parents but heterozygous in offspring with Mendelian violation quality score > 30 using GATK "SelectVariants", and restricted SNPs with genotype quality > 30 and with no genotype-level missing data included for further analysis. The difficulty in definitively mapping reads from repetitive regions makes it impossible to assess DNMs in these regions, therefore we also excluded any SNPs that fall in the genomic regions identified as low complexity by RepeatModeler2 (Flynn et al. 2020) and RepeatMasker v.4.1.1 (Smit 1996) (results can be found on 3.7 Repetitive Genomic Regions Identification).

To further limit the false DNM discovery rate, we filtered SNPs at the individual level as well. We removed all sites in which coverage < ½ × or >2× of individual average sequencing depth. For each trio, all three family members were required to pass these criteria for a site to be retained. Variant callers such as GATK4 and BCFtools determine individual genotypes based on calculated genotype likelihood. This in fact affects DNM calls, as the false discovery rate can arise when heterozygous loci of parents are miscalled as homozygous and from SNPs resulting from somatic mosaicism. Thus, each parental genotype was required to have alternative allele depth to be zero (AD = 0). For heterozygous loci in the offspring, we performed a two-sided binomial test, under the null hypothesis of 0.5 relative frequency of reads supporting reference allele and alternative alleles, with a cutoff *P*-value of 0.05, following methods in Wu et al. (2020) (Supplemental Table S2), removing low-level somatic mosaicism.

SNPs that are in genomic regions with complex read alignments would be problematic for DNM detection and usually indicate misalignments. Following methods in Wang et al. (2022), we further filtered any DNMs located in genomic regions in which >50% of aligned reads contain gaps or are multiply mapped to other genomic locations.

Finally, we visualized each inferred DNM in IGV, an interactive tool for the visual exploration of genomic data. We removed inferred DNMs in which the DNM was observed in the parent, but were removed during the realignment step in GATK4 and BCFtools, which can exclude alternative alleles.

Kinship coefficient analysis

To verify parentage in our pedigrees, we first called genotypes across each family using the GATK4 "HaplotypeCaller" with the following parameters: "--linked-de-bruijn-graph true --min-pruning 0 --recover-all-dangling-branches true" in "-ERC GVCF" mode. Then, we genotyped across individuals in each pedigree using GATK4 "CombineGVCFs" and "GenotypeGVCFs" with default parameters. Pairwise kinship coefficients within each family among the six parents were calculated using KING (Manichaikul et al. 2010) with default parameters.

Mutation rate and callable genome size estimation

To calculate the mutation rate, we estimated the callable genome size using GATK4 "HaplotypeCaller" with the following parameters: "--linked-de-bruijn-graph true --min-pruning 0 --recover-all-dangling-branches true" in "-ERC BP_SOLUTION" mode in each individual. For each individual, we removed sites with coverage that deviated from average individual sequencing read depth, and sites located in the repetitive genomic regions predicted as described above for filtering DNM sites. Lastly, SNPs that were heterozygous in either parent were excluded from the callable genome size estimation. The final callable genome size was intersected across each trio (Supplemental Table S3).

Mutation rate calculation followed the equation below, where μ is the mutation rate, M_i is the *i*th DNM, and C_i is the *i*th intersected callable genome loci of each trio.

$$\mu = \frac{\sum_{i=1}^{x} M_i}{2\sum_{i=1}^{n} C_i}$$

Genotype phasing and male mutation bias estimation

To understand paternal and maternal effects on DNMs, we first phased each individual using WhatsHap in read-based phasing mode with default parameters (Martin et al. 2016). Phasing blocks in offspring bearing DNMs were then matched back to the parent's phasing block to determine DNM phasing results. DNMs with ambiguous phasing were left unphased. Finally, we visualized read alignment from trios using IGV to check the linkage of DNMs and adjacent variants and validate genotype phasing results and inferred the parent of origin for phased DNMs.

Mutational spectra identification

Based on SnpEff (Cingolani et al. 2012) and the guppy reference genome v.99 (Künstner et al. 2016), we were able to annotate each variant and characterize the mutational spectra of DNMs. To test whether the number of DNMs located in coding regions are significantly different than expected based on the coding callable genome, we counted the total number of loci in coding regions and noncoding regions in callable genome size, and used Fisher's exact test.

Functional annotation

For each DNM, we identified the gene, RNA transcript, protein with annotation information from guppy reference genome v.99 (Künstner et al. 2016). We performed GO enrichment analysis using DAVID (Huang et al. 2009a,b) for the genes with a DNM in at least one parent-offspring trio, and recovered no overrepresentation of any GO term.

Simulation of DNM detection pipeline

To determine the false negative rate of our DNM detection pipeline, we used a simulation approach by inserting artificial mutations to the aligned reads of the offspring using BAMSurgeon (Ewing et al. 2015). We randomly simulated 1000 DNMs and restricted the insertion sites to the region of callable genomic region as described above (in Methods subsection "Mutation rate and callable genome size estimation") and nonvariable sites in which less than half of reads contain indels and gaps, as described in previous studies (Wu et al. 2020; Campbell et al. 2021). Then, we went through the full pipeline (Supplemental Fig. S5) to detect the number of DNMs and determined the proportion of simulated DNMs that we detected.

Data access

All raw sequence data of offspring and parents generated in this study have been submitted to the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/) under accession numbers PRJNA970282 and PRJNA858015, respectively. The codes used for processing the data are publicly available as Supplemental Code and at GitHub (https://github.com/Lin-Yuying/GuppyGermlineDNMs).

Competing interest statement

The authors declare no competing interests.

Acknowledgments

This work was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) and a Canada 150 Research Chair (to J.E.M.) and a Doctoral Scholarship from the China Scholarship Council (to Y.L., No. 201906040216). We thank Tom Booker for helpful suggestions on the simulation and Pedro Almeida for the Quare reference genome assembly. We also thank members of the Mank Lab for helpful feedback through the course of this project, as well as constructive comments on a previous version of this manuscript. We thank three anonymous reviewers and Ahmet Denli for improving our manuscript.

Author contributions: J.E.M. and Y.L. conceived the study and designed the experiments and analysis. I.D., J.M., and J.E.M. collected the data. Y.L. and I.D. performed DNA extractions. Y.L., W.v.d.B., and I.D. performed the data analysis. J.E.M. and Y.L. wrote the manuscript and all authors contributed to revisions.

References

Adelman R, Saul RL, Ames BN. 1988. Oxidative damage to DNA: relation to species metabolic rate and life span (aging/thymidine glycol/evolution). *Proc Natl Acad Sci* **85:** 2706–2708. doi:10.1073/pnas.85.8.2706

Almeida P, Sandkam BA, Morris J, Darolti I, Breden F, Mank JE. 2021. Divergence and remarkable diversity of the Y chromosome in guppies. *Mol Biol Evol* **38:** 619–633. doi:10.1093/molbev/msaa257

Alonge M, Soyk S, Ramakrishnan S, Wang X, Goodwin S, Sedlazeck FJ, Lippman ZB, Schatz MC. 2019. RaGOO: fast and accurate reference-guided scaffolding of draft genomes. *Genome Biol* **20:** 224. doi:10.1186/s13059-019-1829-6

Alonge M, Lebeigle L, Kirsche M, Jenike K, Ou S, Aganezov S, Wang X, Lippman ZB, Schatz MC, Soyk S. 2022. Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol* **23:** 258. doi:10.1186/s13059-022-02823-7

Awadalla P, Gauthier J, Myers RA, Casals F, Hamdan FF, Griffing AR, Côté M, Henrion E, Spiegelman D, Tarabeux J, et al. 2010. Direct measure of the de novo mutation rate in autism and schizophrenia cohorts. *Am J Hum Genet* **87:** 316–324. doi:10.1016/j.ajhg.2010.07.019

Barrett RD, Schluter D. 2008. Adaptation from standing genetic variation. *Trends Ecol Evol* **23:** 38–44. doi:10.1016/j.tree.2007.09.008

Bergeron LA, Besenbacher S, Turner TN, Versoza CJ, Wang RJ, Price AL, Armstrong E, Riera M, Carlson J, Chen HY, et al. 2022. The mutationathon highlights the importance of reaching standardization in estimates of pedigree-based germline mutation rates. *eLife* **11:** e73577. doi:10.7554/eLife.73577

Bergeron LA, Besenbacher S, Zheng J, Li P, Bertelsen MF, Quintard B, Hoffman JI, Li Z, St. Leger J, Shao C, et al. 2023. Evolution of the germline mutation rate across vertebrates. *Nature* **615:** 285–291. doi:10.1038/s41586-023-05752-y

Besenbacher S, Liu S, Izarzugaza JMG, Grove J, Belling K, Bork-Jensen J, Huang S, Als TD, Li S, Yadav R, et al. 2015. Novel variation and *de novo* mutation rates in population-wide *de novo* assembled Danish trios. *Nat Commun* **6:** 5969. doi:10.1038/ncomms6969

Billard R, Escaffre A-M. 1969. La spermatogenèse de Poecilia reticulata. I-estimation du nombre de générations goniales et rendement de la spermatogenèse. *Ann Biol Anim Biochim Biophys* **9:** 251–271. doi:10.1051/rnd:19690208

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30:** 2114–2120. doi:10.1093/bioinformatics/btu170

Cagan A, Baez-Ortega A, Brzozowska N, Abascal F, Coorens THH, Sanders MA, Lawson ARJ, Harvey LMR, Bhosle S, Jones D, et al. 2022. Somatic mutation rates scale with lifespan across mammals. *Nature* **604:** 517–524. doi:10.1038/s41586-022-04618-z

Campbell CR, Tiley GP, Poelstra JW, Hunnicutt KE, Larsen PA, Lee HJ, Thorne JL, dos Reis M, Yoder AD. 2021. Pedigree-based and phylogenetic methods support surprising patterns of mutation rate and spectrum in the gray mouse lemur. *Heredity (Edinb)* **127:** 233–244. doi:10.1038/s41437-021-00446-5

Carlson J, Locke AE, Flickinger M, Zawistowski M, Levy S, Myers RM, Boehnke M, Kang HM, Scott LJ, Li JZ, et al. 2018. Extremely rare variants reveal patterns of germline mutation rate heterogeneity in humans. *Nat Commun* **9:** 3573. doi:10.1038/s41467-018-05936-5

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6:** 80–92. doi:10.4161/fly.19695

Conrad DF, Keebler JEM, DePristo MA, Lindsay SJ, Zhang Y, Casals F, Idaghdour Y, Hartl CL, Torroja C, Garimella KV, et al. 2011. Variation in genome-wide mutation rates within and between human families. *Nat Genet* **43:** 712–714. doi:10.1038/ng.862

Coombe L, Zhang J, Vandervalk BP, Chu J, Jackman SD, Birol I, Warren RL. 2018. ARKS: chromosome-scale scaffolding of human genome drafts with linked read kmers. *BMC Bioinformatics* **19:** 234. doi:10.1101/306902

De Manuel M, Wu FL, Przeworski M. 2022. A paternal bias in germline mutation is widespread in amniotes and can arise independently of cell division numbers. *eLife* **11:** e80008. doi:10.7554/eLife.80008

Droller MJ, Roth TF. 1966. An electron microscope study of yolk formation during oogenesis in *Lebistes reticulatus* guppyi. *J Cell Biol* **28:** 209–232. doi:10.1083/jcb.28.2.209

Ellegren H. 2007. Characteristics, causes and evolutionary consequences of male-biased mutation. *Proc Biol Sci* **274:** 1–10. doi:10.1098/rspb.2006.3720

Evans JP, Pitcher TE, Magurran AE. 2002. The ontogeny of courtship, colour and sperm production in male guppies. *J Fish Biol* **60:** 495–498. doi:10.1006/jfbi.2001.1849

Ewing AD, Houlahan KE, Hu Y, Ellrott K, Caloian C, Yamaguchi TN, Bare JC, P'Ng C, Waggott D, Sabelnykova VY, et al. 2015. Combining tumor genome simulation with crowdsourcing to benchmark somatic single-nucleotide-variant detection. *Nat Methods* **12:** 623–630. doi:10.1038/nmeth.3407

Feng C, Pettersson M, Lamichhaney S, Rubin CJ, Rafati N, Casini M, Folkvord A, Andersson L. 2017. Moderate nucleotide diversity in the Atlantic herring is associated with a low mutation rate. *eLife* **6:** e23907. doi:10.7554/eLife.23907

Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci* **117:** 9451–9457. doi:10.1073/pnas.1921046117

Francioli LC, Polak PP, Koren A, Menelaou A, Chun S, Renkens I, Van Duijn CM, Swertz M, Wijmenga C, Van Ommen G, et al. 2015. Genome-wide patterns and properties of *de novo* mutations in humans. *Nat Genet* **47:** 822–826. doi:10.1038/ng.3292

Fraser BA, Künstner A, Reznick DN, Dreyer C, Weigel D. 2015. Population genomics of natural and experimental populations of guppies (*Poecilia reticulata*). *Mol Ecol* **24:** 389–408. doi:10.1111/mec.13022

Gao Z, Moorjani P, Sasani TA, Pedersen BS, Quinlan AR, Jorde LB, Amster G, Przeworski M. 2019. Overlooked roles of DNA damage and maternal age in generating human germline mutations. *Proc Natl Acad Sci* **116:** 9491–9500. doi:10.1073/pnas.1901259116

Goldmann JM, Wong WSW, Pinelli M, Farrah T, Bodian D, Stittrich AB, Glusman G, Vissers LELM, Hoischen A, Roach JC, et al. 2016. Parent-of-origin-specific signatures of *de novo* mutations. *Nat Genet* **48:** 935–939. doi:10.1038/ng.3597

Goldmann JM, Veltman JA, Gilissen C. 2019. *De novo* mutations reflect development and aging of the human germline. *Trends Genet* **35:** 828–839. doi:10.1016/j.tig.2019.08.005

Goldmann JM, Hampstead JE, Wong WSW, Wilfert AB, Turner TN, Jonker MA, Bernier R, Huynen MA, Eichler EE, Veltman JA, et al. 2021. Differences in the number of *de novo* mutations between individuals are because of small family-specific effects and stochasticity. *Genome Res* **31:** 1513–1518. doi:10.1101/gr.271809.120

Gordon SP, Reznick D, Arendt JD, Roughton A, Ontiveros Hernandez MN, Bentzen P, López-Sepulcre A. 2015. Selection analysis on the rapid evolution of a secondary sexual trait. *Proc Biol Sci* **282:** 20151244. doi:10.1098/rspb.2015.1244

Harland C, Charlier C, Karim L, Cambisano N, Deckers M, Mni M, Mullaart E, Coppieters W, Georges M. 2017. Frequency of mosaicism points towards mutation-prone early cleavage cell divisions in cattle. bioRxiv doi:10.1101/079863

Harris AM, Garud NR, Degiorgio M. 2018. Detection and classification of hard and soft sweeps from unphased genotypes by multilocus genotype identity. *Genetics* **210:** 1429–1452. doi:10.1534/genetics.118.301502

Hermisson J, Pennings PS. 2005. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* **169:** 2335–2352. doi:10.1534/genetics.104.036947

Hermisson J, Pennings PS. 2017. Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. *Methods Ecol Evol* **8:** 700–716. doi:10.1111/2041-210X.12808

Huang DW, Sherman BT, Lempicki RA. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37:** 1–13. doi:10.1093/nar/gkn923

Huang DW, Sherman BT, Lempicki RA. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4:** 44–57. doi:10.1038/nprot.2008.211

Jónsson H, Sulem P, Kehr B, Kristmundsdottir S, Zink F, Hjartarson E, Hardarson MT, Hjorleifsson KE, Eggertsson HP, Gudjonsson SA, et al. 2017. Parental influence on human germline *de novo* mutations in 1,548 trios from Iceland. *Nature* **549:** 519–522. doi:10.1038/nature24018

Kessler MD, Loesch DP, Perry JA, Heard-Costa NL, Taliun D, Cade BE, Wang H, Daya M, Ziniti J, Datta S, et al. 2020. *De novo* mutations across 1,465 diverse genomes reveal mutational insights and reductions in the Amish founder population. *Proc Natl Acad Sci* **117:** 2560–2569. doi:10.1073/pnas.1902766117

Koch EM, Schweizer RM, Schweizer TM, Stahler DR, Smith DW, Wayne RK, Novembre J. 2019. De novo mutation rate estimation in wolves of known pedigree. *Mol Biol Evol* **36:** 2536–2547. doi:10.1093/molbev/msz159

Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, Gudjonsson SA, Sigurdsson A, Jonasdottir A, Jonasdottir A, et al. 2012. Rate of *de novo* mutations and the importance of father's age to disease risk. *Nature* **488:** 471–475. doi:10.1038/nature11396

Künstner A, Hoffmann M, Fraser BA, Kottler VA, Sharma E, Weigel D, Dreyer C. 2016. The genome of the Trinidadian guppy, *Poecilia reticulata*, and variation in the Guanapo population. *PLoS One* **11:** e0169087. doi:10.1371/journal.pone.0169087

Lee KM, Coop G. 2017. Distinguishing among modes of convergent adaptation using population genomic data. *Genetics* **207:** 1591–1619. doi:10.1534/genetics.117.300417

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25:** 1754–1760. doi:10.1093/bioinformatics/btp324

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25:** 2078–2079. doi:10.1093/bioinformatics/btp352

Lin Y, Darolti I, Furman BLS, Almeida P, Sandkam BA, Breden F, Wright AE, Mank JE. 2022. Gene duplication to the Y chromosome in Trinidadian guppies. *Mol Ecol* **31:** 1853–1863. doi:10.1111/mec.16355

Lindsay SJ, Rahbari R, Kaplanis J, Keane T, Hurles ME. 2019. Similarities and differences in patterns of germline mutation between mice and humans. *Nat Commun* **10:** 4053. doi:10.1038/s41467-019-12023-w

Lynch M, Ackerman MS, Gout JF, Long H, Sung W, Thomas WK, Foster PL. 2016. Genetic drift, selection and the evolution of the mutation rate. *Nat Rev Genet* **17:** 704–714. doi:10.1038/nrg.2016.104

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M, Barrett J. 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26:** 2867–2873. doi:10.1093/bioinformatics/btq559

Martin AP, Palumbit SR. 1993. Body size, metabolic rate, generation time, and the molecular clock. *Proc Natl Acad Sci* **90:** 4087–4091. doi:10.1073/pnas.90.9.4087

Martin M, Patterson M, Garg S, Fischer SO, Pisanti N, Klau GW, Schöenhuth A, Marschall T. 2016. WhatsHap: fast and accurate read-based phasing. bioRxiv doi:10.1101/085050

Martin SH, Singh KS, Gordon IJ, Omufwoko KS, Collins S, Warren IA, Munby H, Brattström O, Traut W, Martins DJ, et al. 2020. Whole-chromosome hitchhiking driven by a male-killing endosymbiont. *PLoS Biol* **18:** e3000610. doi:10.1371/journal.pbio.3000610

Matuszewski S, Hermisson J, Kopp M. 2015. Catch me if you can: adaptation from standing genetic variation to a moving phenotypic optimum. *Genetics* **200:** 1255–1274. doi:10.1534/genetics.115.178574

Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* **34:** i142–i150. doi:10.1093/bioinformatics/bty266

Milholland B, Dong X, Zhang L, Hao X, Suh Y, Vijg J. 2017. Differences between germline and somatic mutation rates in humans and mice. *Nat Commun* **8:** 15183. doi:10.1038/ncomms15183

Morris J, Darolti I, van der Bijl W, Mank JE. 2020. High-resolution characterization of male ornamentation and re-evaluation of sex linkage in guppies. *Proc R Soc B: Biol Sci* **287:** 20201677. doi:10.1098/rspb.2020.1677

Pfeifer SP. 2017. Direct estimate of the spontaneous germ line mutation rate in African green monkeys. *Evolution* **71:** 2858–2870. doi:10.1111/evo.13383

Pritchard JK, Pickrell JK, Coop G. 2010. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr Biol* **20:** R208–R215. doi:10.1016/j.cub.2009.11.055

Rahbari R, Wuster A, Lindsay SJ, Hardwick RJ, Alexandrov LB, Al Turki S, Dominiczak A, Morris A, Porteous D, Smith B, et al. 2016. Timing, rates and spectra of human germline mutation. *Nat Genet* **48:** 126–133. doi:10.1038/ng.3469

Reznick DA, Bryga H, Endler JA. 1990. Experimentally induced life-history evolution in a natural population. *Nature* **346:** 357–359. doi:10.1038/346357a0

Reznick DN, Shaw FH, Rodd HF, Shaw RG. 1997. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science* **275:** 1934–1937. doi:10.1126/science.275.5308.1934

Reznick D, Bryant M, Holmes D. 2006. The evolution of senescence and post-reproductive lifespan in guppies (*Poecilia reticulata*). *PLoS Biol* **4:** e7. doi:10.1371/journal.pbio.0040007

Roach JC, Glusman G, Smit AFA, Huff CD, Hubley R, Shannon PT, Rowen L, Pant KP, Goodman N, Bamshad M, et al. 2010. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* **328:** 636–639. doi:10.1126/science.1186802

Seplyarskiy VB, Soldatov RA, Koch E, McGinty RJ, Goldmann JM, Hernandez RD, Barnes K, Correa A, Burchard EG, Ellinor PT, et al. 2021. Population sequencing data reveal a compendium of mutational processes in the human germ line. *Science* **373:** 1030–1035. doi:10.1126/science.aba7408

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31:** 3210–3212. doi:10.1093/bioinformatics/btv351

Smeds L, Qvarnström A, Ellegren H. 2016. Direct estimate of the rate of germline mutation in a bird. *Genome Res* **26:** 1211–1218. doi:10.1101/gr.204669.116

Smit AFA. 1996. The origin of interspersed repeats in the human genome. *Curr Opin Genet Dev* **6:** 743–748. doi:10.1016/S0959-437X(96)80030-X

Stern DL. 2013. The genetic causes of convergent evolution. *Nat Rev Genet* **14:** 751–764. doi:10.1038/nrg3483

Tatsumoto S, Go Y, Fukuta K, Noguchi H, Hayakawa T, Tomonaga M, Hirai H, Matsuzawa T, Agata K, Fujiyama A. 2017. Direct estimation of *de novo* mutation rates in a chimpanzee parent-offspring trio by ultra-deep whole genome sequencing. *Sci Rep* **7:** 13561. doi:10.1038/s41598-017-13919-7

Thomas GWC, Wang RJ, Puri A, Harris RA, Raveendran M, Hughes DST, Murali SC, Williams LE, Doddapaneni H, Muzny DM, et al. 2018. Reproductive longevity predicts mutation rates in primates. *Curr Biol* **28:** 3193–3197.e5. doi:10.1016/j.cub.2018.08.050

Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief Bioinformatics* **14:** 178–192. doi:10.1093/bib/bbs017

Van der Auwera G, O'Connor B. 2020. *Genomics in the cloud: using Docker, GATK, and WDL in Terra* (1st ed.). O'Reilly Media, Sebastopol, CA.

van der Zee MJ, Whiting JR, Paris JR, Bassar RD, Travis J, Weigel D, Reznick DN, Fraser BA. 2022. Rapid genomic convergent evolution in experimental populations of Trinidadian guppies (*Poecilia reticulata*). *Evol Lett* **6:** 149–161. doi:10.1002/evl3.272

Venn O, Turner I, Mathieson I, de Groot N, Bontrop R, McVean G. 2014. Strong male bias drives germline mutation in chimpanzees. *Science* **344:** 1272–1275. doi:10.1126/science.344.6189.1272

Wang RJ, Thomas GWC, Raveendran M, Harris RA, Doddapaneni H, Muzny DM, Capitanio JP, Radivojac P, Rogers J, Hah MW. 2020. Paternal age in rhesus macaques is positively associated with germline mutation accumulation but not with measures of offspring sociability. *Genome Res* **30:** 826–834. doi:10.1101/gr.255174.119

Wang RJ, Raveendran M, Harris RA, Murphy WJ, Lyons LA, Rogers J, Hahn MW. 2022. *De novo* mutations in domestic cat are consistent with an effect of reproductive longevity on both the rate and spectrum of mutations. *Mol Biol Evol* **39:** msac147. doi:10.1093/molbev/msac147

Whiting JR, Paris JR, van der Zee MJ, Parsons PJ, Weigel D, Fraser BA. 2021. Drainage-structuring of ancestral variation and a common functional pathway shape limited genomic convergence in natural high- and low-predation guppies. *PLoS Genet* **17:** e1009566. doi:10.1371/journal.pgen.1009566

Whiting JR, Paris JR, Parsons PJ, Matthews S, Reynoso Y, Hughes KA, Reznick D, Fraser BA. 2022. On the genetic architecture of rapidly adapting and convergent life history traits in guppies. *Heredity (Edinb)* **128:** 250–260. doi:10.1038/s41437-022-00512-6

Wilson Sayres MA, Makova KD. 2011. Genome analyses substantiate male mutation bias in many species. *Bioessays* **33:** 938–945. doi:10.1002/bies.201100091

Wu FL, Przeworski M, Moorjani P, Przeworski M, Strand AI, Cox LA, Cox LA, Ober C, Wall JD, Strand AI, et al. 2020. A comparison of humans and baboons suggests germline mutation rates do not track cell divisions. *PLoS Biol* **18:** e3000838. doi:10.1371/JOURNAL.PBIO.3000838

Yoder AD, Tiley GP. 2021. The challenge and promise of estimating the *de novo* mutation rate from whole-genome comparisons among closely related individuals. *Mol Ecol* **30:** 6087–6100. doi:10.1111/mec.16007

Received March 30, 2023; accepted in revised form July 7, 2023.

Extensive variation in germline de novo mutations in *Poecilia reticulata*

Yuying Lin, Iulia Darolti, Wouter van der Bijl, et al.

Genome Res. published online July 13, 2023
Access the most recent version at doi:10.1101/gr.277936.123