De novo mutations in domestic cat are consistent with an effect of reproductive longevity on both the rate and spectrum of mutations

Richard J. Wang[1], Muthuswamy Raveendran[2,3], R. Alan Harris[2,3], William J. Murphy[4], Leslie A. Lyons[5], Jeffrey Rogers[2,3], and Matthew W. Hahn[1,6]

[1] Department of Biology, Indiana University, Bloomington, IN
[2] Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX
[3] Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX
[4] Veterinary Integrative Biosciences, Texas A&M University, College Station, TX
[5] Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, University of Missouri, Columbia, MO
[6] Department of Computer Science, Indiana University, Bloomington, IN

Abstract

The mutation rate is a fundamental evolutionary parameter with direct and appreciable effects on the health and function of individuals. Here, we examine this important parameter in the domestic cat, a beloved companion animal as well as a valuable biomedical model. We estimate a mutation rate of 0.86×10^{-8} per bp per generation for the domestic cat (at an average age of 3.8 years). We find evidence for a strong paternal age effect, with more mutations transmitted by older sires. Our analyses suggest that the cat and the human have accrued similar numbers of mutations in the germline before reaching sexual maturity. The per-generation mutation rate in the cat is slightly lower than what has been observed in humans, but consistent with the shorter generation time in the cat. Using a model of reproductive longevity, which takes into account differences in the reproductive age and time to sexual maturity, we are able to explain much of the difference in per-generation rates between species. We further apply our reproductive longevity model in a novel analysis of mutation spectra and find that the spectrum for the cat resembles the human mutation spectrum at a younger age of reproduction. Together, these results implicate changes in life-history as a driver of mutation rate evolution between species. As the first direct observation of the paternal age effect outside of primates, our results also suggest a phenomenon that may be universal among mammals.

Introduction

Mutation is a fundamental force in evolution, the primary source of genetic novelty, and often a fitness burden to its bearers in the form of inherited disease. Germline mutations, and the processes that govern the occurrence and rate of these heritable genetic changes, have long been of interest to evolutionary biologists (e.g. Muller, Crow). Approaches for studying the mutation rate have changed drastically with changes in sequencing technologies. The oldest estimates of mutation rates predate the availability of molecular data and were based on the incidence of rare dominant disease (Haldane 1935; reviewed in Nachman 2004). With low-cost whole-genome sequencing, direct estimates of the mutation rate can now be made by comparison of parents and offspring. These estimates of the per-generation mutation rate in a growing number of species have enabled an increased understanding of how this fundamental parameter evolves.

When determining the number of mutations an individual inherits from its parents, the estimated rate necessarily spans a generation. Since generation times vary among species, direct estimates of the mutation rate across species are difficult to compare. Nevertheless, current data suggest at least a twofold range of mutation rates among primates (Chintalapati and Moorjani 2020). In almost all primate species studied to date, paternal age has an effect on the number of inherited mutations (Kong et al. 2012; Venn et al. 2014; Goldmann et al. 2016; Thomas et al. 2018; Besenbacher et al. 2019; Bergeron et al. 2020; Wang et al. 2020). Because males transmit so many more mutations to their offspring than females, the paternal age effect dominates in comparisons of the total number of inherited mutations. Together, the paternal age effect and male mutation bias obscure the mechanisms driving mutation rate evolution: have per-generation mutation rates changed due to differences in the quality of DNA repair or are they due to changes in life-history? Variation in the age of reproduction across species, coupled with

parental age effects (Rahbari et al. 2016; Jónsson et al. 2017), would suggest that per-generation rates are expected to vary across species simply because of differences in life-history.

One approach that attempts to appropriately compare mutation rates between species is to take the calendar age of the parents at conception into account (e.g. Besenbacher et al. 2019). Given a known relationship between parental age and the number of mutations inherited by offspring, this approach can predict differences in the mutation rate between species solely due to differences in the age at reproduction for parents from each species. Deviations from this prediction imply changes apart from life-history, be they cellular, molecular, or environmental. This approach assumes that mutations accumulate at a constant rate across the lifespan of an individual (cf. Gao et al. 2019), and is therefore equivalent to comparing per-year estimates of the mutation rate (Besenbacher et al. 2019; Chintalapati and Moorjani 2020). However, germline mutations appear to accumulate at different rates across different life stages (Rahbari et al. 2016; Scally 2016; Harland et al. 2017; Jónsson et al. 2018; Sasani et al. 2019; Jonsson et al. 2021). Furthermore, the estimated relationship between parental age and the number of inherited mutations inferred from trio studies can only be predictive to the age of puberty: the number of germline mutations cannot be directly observed via transmission before parents reach reproductive age. Therefore, a model of "total longevity" (i.e., calendar age) in comparisons between species may not fully account for differences in mutation accumulation across different life stages. This is especially true when the age at puberty differs between species.

Differences in the rate of mutation accumulation across life stages have motivated an alternative model that splits the process into two regimes: before and after the onset of sexual maturity (e.g. Thomas and Hahn 2014; Amster and Sella 2016; Gao et al. 2016). The disproportionate contribution of mutations from fathers—and the biology of spermatogenesis,

which requires continuous division of the male germline after puberty (Crow 2000)—strongly suggest that reproductive age marks an important change in the rate of mutation accumulation. A starting point for comparisons between species is therefore to take the reproductive age of parents into account. Like the total longevity model described above, this "reproductive longevity" model (Thomas et al. 2018) assumes a common relationship among species between parental age and the number of mutations inherited by offspring. The key distinction for this model is that age-related mutation accumulation only begins with the onset of puberty; the number of mutations before sexual maturity is assumed to be independent of the age at which puberty is reached. In this model, differences in the per-generation mutation rate between species are best explained by reproductive longevity: that is, the length of time for mutation accumulation after puberty. Even when the per-generation mutation rate has evolved between species, this model can help to delineate the stage(s) at which changes have occurred (e.g. Wang et al. 2020).

Most pedigree studies of mutation have, to date, been focused on humans and other primates (e.g. Kong et al. 2012; Venn et al. 2014; Besenbacher et al. 2019). To better understand how mutation rates evolve, a broader collection of species will need to be considered. Distinguishing between models for the evolution of mutation rates will also require species with a larger span of generation times and greater differences in the age of puberty. In this study, we present a direct estimate of the mutation rate from the domestic cat (*Felis catus*). Cats are an important companion animal for many households and a common target for interventions by veterinary medicine. Recent studies of genetic variation in the cat and the development of feline genomic tools (e.g., Genova et al. 2018; Bredemeyer et al. 2020; Buckley et al. 2020) have advanced the domestic cat as a biomedical model. Reliable mutation rates would also help to

accurately estimate the evolutionary history of domesticated and wild cat species (Figueiró et al. 2017; Li et al. 2019).

In this study, we consider models for the evolution of mutation rates in the domestic cat. The difference in lifespan between cats and longer-lived primates, like humans, provide a strong contrast for testing models of mutation accumulation. In addition to considering a model of mutation rates across species, we perform a novel analysis contrasting the mutation spectrum across species. As with the total number of inherited mutations, the mutation spectrum also changes with the age of parents (Jónsson et al. 2017). Therefore, differences in mutation spectrum between species with very different ages at reproduction may also be explained by changes in life-histories. Our results suggest that a model of reproductive longevity explains most of the difference in both the rate and spectrum of mutations between the cat and the human.

Results

Mutation rate in the domestic cat

We sequenced individuals from 11 domestic cat trios to a median of 41× coverage using Illumina short-read sequencing. Of the 22 total individuals sequenced, 19 were part of a larger pedigree, while 3 were from a standalone trio (Fig. 1). DNA from this standalone trio was collected from cultured fibroblasts, while all others were collected from blood samples. After filtering sites based on mapping quality and coverage, we retained an average 1.8 Gb out of the 2.5 Gb reference, from which we identified candidate *de novo* mutations in each trio. Briefly, we considered sites on autosomes where both parents were homozygous for the reference allele and the offspring was heterozygous for an alternate allele not observed elsewhere in the dataset (see Methods). We applied a set of stringent filters to the initial list of candidates and found 233 single nucleotide *de novo* mutations across the 11 trios (Supplemental Data 1). The structure of the pedigree allowed us to trace mutations to a third generation in four grandchildren, and we found transmission in 40 out of the 78 potential instances, close to the expected 50% transmission rate. From this set of transmitted mutations, we determined that 70% had a paternal origin, consistent with a male mutation bias (Wilson Sayres and Makova 2011; Ségurel et al. 2014).

To estimate the per-generation mutation rate, we divided the number of mutations identified in each trio by the observed genome size and corrected for the false negative rate from our stringent set of filters. We estimated this latter value as the "site callability," by calculating the fraction of sites that pass our filters from a sample of sites across the genome in each trio (see Methods; Besenbacher et al. 2019; Wang et al. 2020). This approach produced stable estimates

of the mutation rate as the filter stringency was increased (Fig. S1), providing confidence in our estimate. We found the mean per-generation mutation rate in cat to be 0.86×10^{-8} per bp for parents at an average age of 3.8 years across sexes, compared to 1.20×10^{-8} per bp in humans with an average age of 30.1 years (Jónsson et al. 2017); Table 1 shows the rate estimated for each cat trio separately. Notably, the mutation rate estimated for the trio using DNA isolated from cultured fibroblasts was not notably different from those using DNA from blood samples.

Assuming the average parental age in our sample (3.8 years) is representative of the average age of reproduction in the cat, we estimate a per-year mutation rate of 2.2×10^{-9} per bp. This is much higher than the human rate of 0.43×10^{-9} per bp per year (Jónsson et al. 2017), or the per-year rate from any reported primate (Besenbacher et al. 2019; Campbell et al. 2020; Wang et al. 2020; Wu et al. 2020). This higher rate is driven by the similar number of mutations at puberty, but a much shorter generation time in the cat.

The mutation spectrum in the cat largely resembles the spectrum found in the human and most other primates (Fig. S2). The biggest difference was a significantly lower percentage of A>G transitions in the cat, 19%, compared to the 27% found in humans (χ^2 test, $p < 0.005$; human data from Jónsson et al. 2017). This was offset by a slightly higher percentage of C>T transitions, 48% in cat vs. 42% in human, leading to an overall transition / transversion ratio of 1.99 in cat. As in humans, C>T transitions at CpG sites accounted for a substantial fraction of all observed mutations (21%). We estimate the mutation rate at CpG sites in the cat to be 1.76×10^{-7} per bp per generation.

Testing the reproductive longevity model of the mutation rate

Direct estimates of the mutation rate have consistently demonstrated a strong parental age effect in primates, with more mutations inherited by the offspring of older parents. To consider whether such an effect exists in the domestic cat, we performed a Poisson regression on the number of mutations (adjusted by observable genome size) with parental age, and found a strong paternal age effect on mutation rate ($p < 0.01$). Each additional year of paternal age at conception leads to an estimated 3.1 (95% CI: [0.8, 5.4]) additional mutations in the offspring. This estimate for the strength of the paternal age effect in cat is not significantly different from the per-year effect observed in humans (Kong et al. 2012; Francioli et al. 2015; Jónsson et al. 2017). In contrast, we found no significant effect of maternal age on the mutation rate (Fig. S3), though much larger sample sizes were needed to detect a significant maternal age effect in humans (Goldmann et al. 2016; Jónsson et al. 2017).

The accumulation of mutations with age post-puberty among species is predicted by both the total longevity and reproductive longevity models (Fig. 2a). The key difference in predictions between the two models is in the number of mutations at sexual maturity. In the total longevity model, the cat mutation rate at sexual maturity is predicted by extending the regression fit of the parental age effect found in humans to the age at sexual maturity in the cat. In contrast, the reproductive longevity model assumes a similar number of mutations occur before puberty across species, with the parental age effect only acting post-puberty (Fig. 2a). We compared the predictions made by these two models to the data from our cat trios: Figure 2b shows the increase in per-generation mutation rates with years of paternal age after puberty (assuming a puberty age of 0.5 y in cat, 13 y in human; Tsutsui et al. 2004; Parent et al. 2003). The

reproductive longevity model provides a significantly better fit to the mutation rates observed in the cat (*t*-test of residuals, $p = 0.0002$).

The per-generation mutation rate estimated at the age of puberty in cats and humans are highly similar: we estimate 0.59×10^{-8} per bp (95% CI: [0.39, 0.78]) in cat vs. 0.67×10^{-8} per bp ([0.65, 0.69]) in human (Jónsson et al. 2017). This constant number between species who shared a common ancestor approximately 85 million years ago (Tarver et al. 2016) is consistent with a conserved pre-puberty developmental program among mammals, as has previously been noted in owl monkeys (Thomas et al. 2018). Note that the number of pre-puberty mutations was slightly lower in rhesus macaques (Wang et al. 2020).

A reproductive longevity model of the mutation spectrum

Parental age affects not only the number of *de novo* mutations inherited by offspring, but also their composition (Carlson et al. 2020). In humans, the proportion of C>T transitions decreases with parental age while the proportion of C>G transversions increases with parental age (Goldmann et al. 2016; Jónsson et al. 2017). The dependence of the mutation spectrum on age suggests multiple underlying mutation processes whose relationships differ with age (Goldmann et al. 2019). Since differences in life history appear to explain much of the variation in per-generation mutation rates between cat and human, we were interested in knowing whether a model of reproductive longevity could be used to explain differences in the mutation spectrum. That is, does the mutation spectrum in the cat resemble the mutation spectrum of a human that has had the same amount of post-pubertal time for mutation accumulation? Or, alternatively, are there different mutational processes in cats that explain the difference in mutation spectrum?

To build a reproductive longevity model of the mutation spectrum, we used mutation data from a large human dataset (Jónsson et al. 2017). We modeled the accumulation of mutations for each mutation class as independent and fit the number of mutations in each class with a Poisson regression. For each class of mutations, $c \in$ {A>C, A>G, A>T, C>A, C>G, C>T}, the number of mutations $N_{c,i}$ in individual i was modeled as

$$N_{c,i} \sim \mathrm{Poisson}(\mu_{c,\mathrm{i}})$$

where

$$\mu_{c,i} = \beta_0^c + \beta_1^c X_i$$

is the per-generation rate, β_0^c and β_1^c are the regression coefficients for the mutation class, and X_i is the age at conception for parents of the individual. From this regression, we predicted the spectrum for a given age by dividing the predicted number of mutations in each class by the total number of mutations predicted across classes. Figure 3a shows how the predicted mutation spectrum changes with parental age relative to the average spectrum observed in the human dataset (with an average age of 30.1; Jónsson et al. 2017). Under this model, the mutation spectrum for species with different parental ages at conception can vary dramatically from the one observed in humans without any change to the mutational process between species.

We compared the fit of the observed cat mutation spectrum (at an average parental age of 3.8 y) to the predicted spectrum under both the total longevity and reproductive longevity models. The predicted mutation spectrum for the cat is expected to depart more from the average human spectrum under the total longevity model than under the reproductive longevity model because there is a larger difference in the ages considered (calendar age vs. reproductive age; Fig. 3a). We found that the predicted spectrum under the reproductive longevity model provided a better fit to the observed cat mutation spectrum (Fig. 3b; root mean square error 3.6% vs.

4.5%). Under the better-fitting reproductive longevity model, the predicted proportion of C>T transitions, 45.1%, closely matched the one found in the cat spectrum, 47.6%. The most notable difference between this prediction and the observed spectrum was the proportion of A>G transitions, which accounted for 25.9% of mutations in the prediction, but only 18.8% of observed mutations in the cat. While the overall fit of the spectrum predicted by the total longevity model was worse, its prediction for the proportion of A>G transitions and A>T transversions was much closer to the observed spectrum in the cat (predicted vs. observed for A>G and A>T respectively: 21.7% vs. 18.9% and 9.9% vs. 11.6%). Together, these results suggest that, while the reproductive longevity model captures much of the mutation spectrum's dependence on age, multiple mutation processes likely underlie differences in mutation spectra between species.

Discussion

We find a mutation rate and spectrum for the domestic cat that is variable throughout its lifespan and consistent with a model of reproductive longevity. The estimated per-generation rate, at 0.86×10^{-8} per bp, is lower than that found in the human but, because of the shorter lifespan of the cat and similar number of mutations pre-puberty between species, corresponds to a higher per-year rate. Though studies in mice have hinted at a paternal age effect (Walter et al. 1998; see also Uchimura et al. 2015), this is the first demonstration of its existence from direct estimates in mammals outside of primates (Figure 2b). Our results suggest that the paternal age effect on mutation rate may be universal to all mammals. As far as we are aware, a paternal age effect of similar magnitude has now been found across every mammalian species examined (Kong et al. 2012; Venn et al. 2014; Goldmann et al. 2016; Thomas et al. 2018; Besenbacher et al. 2019; Bergeron et al. 2020; Wang et al. 2020; see Wu et al. 2020 for an exception in the baboon). Differences in the magnitude of the parental age effect are not the only way in which the mutation rate could evolve. Changes to the number of mutations before reaching puberty, as suggested in macaques (Wang et al. 2020), will affect the per-generation rate without any change in the parental age effect. However, we found no significant difference in the number of mutations before puberty between cats and humans (Figure 2b).

Pedigree studies have underscored the difference between mutation rates measured within species and substitution rates measured between species (Chintalapati and Moorjani 2020). In humans, the difference between these two rates has called into question assumptions about how mutation rates translate into substitution rates (Scally and Durbin 2012; Ségurel et al. 2014; Thomas and Hahn 2014; Amster and Sella 2016; Gao et al. 2016). Connecting the two will involve a better understanding how mutation rates vary across a lifespan, especially pre- and

post-puberty. This dichotomy is the major difference between the total longevity and reproductive longevity models (Figure 2a). Determining whether rates of mutation accumulation or the number of mutations at puberty differ among closely related primates (e.g. Besenbacher et al. 2019) is quite difficult because of their highly similar ages at puberty and lifespans. Interestingly, our estimate of the per-generation mutation rate in the cat is nearly double the estimate in the wolf (*Canis lupus*), another member of the order Carnivora, at 0.45×10^{-8} per bp (Koch et al. 2019). Given the relatively similar age at reproduction for the wolves used in this study (mean: 3.1 y), and the consistency of our estimates with a reproductive longevity model, we are not sure where the discrepancy in these estimates lies. Nevertheless, distinguishing between modes of evolution in the mutation rate will require more species with varied life histories.

The spectrum of mutations in the cat was remarkably similar to those found in most primates, with a preponderance of C>T and A>G transitions (Kong et al. 2012; Venn et al. 2014; Thomas et al. 2018; Besenbacher et al. 2019; Bergeron et al. 2020; Wang et al. 2020; Wu et al. 2020; see Campbell et al. 2020 for an exception in the mouse lemur). We showed that differences in the mutation spectrum between the cat and the human can in large part be explained by differences in age according to a model of reproductive longevity. Our results indicate that, like mutation rate itself, the mutation spectrum should be treated as a function-valued trait that changes with parental age. Mutation spectra have typically been reported and compared with little regard for the reproductive age of individuals in the sample. But just as in the case for mutation rates, meaningful comparisons must take reproductive age and longevity into account. This applies not only to comparisons of spectra between species and populations (Harris and Pritchard 2017), but also to inferences about spectra from different stages of the

germline (Rahbari et al. 2016; Bae et al. 2018). Distinct spectra from different developmental stages have been used to imply that a unique set of mutational processes underlie each stage. However, a single set of processes pervading all stages is equally capable of producing significant differences in spectra. Distinguishing between these competing hypotheses requires the comparison of models for mutation spectra that consider reproductive age.

In addition to sequencing blood from individuals in the larger pedigrees, we sequenced DNA taken from cell lines developed from a single trio. Researchers routinely sequence blood samples for studies of *de novo* mutations to avoid sampling a small number of cells from a single clonal lineage that may contain a somatic mutation at high frequency. Furthermore, cell lines that have been passaged for a long period can accumulate mutations at a very high rate: in the 1000 human genomes project (Altshuler et al. 2010), 90-95% of mutations initially detected were cell line-specific. Notably, the estimate here from a single trio of cultured fibroblasts was not distinguishable from the blood-derived estimates. However, the collected fibroblasts proceeded through only two or three passages before DNA was collected, which may have mitigated any error-prone behavior.

Finally, our results have important implications for multiple aspects of feline evolution and veterinary care. Accurate estimates of mutation rates are important for reconstructing the evolutionary history of cats. By combining genomic data from domesticated cats with data from the wildcat (*Felis silvestris lybica*), our mutation rate estimates will allow for more refined estimates of the timing of cat domestication (Driscoll et al. 2007; Ottoni et al. 2017) and felid evolution in general (Figueiró et al. 2017; Li et al. 2019). Cat breeding, and the management of captive felids, has altered population dynamics over the past few thousand years. In the wild, the normal life expectancy for cats has been estimated at ~5 years of age. Both the breeding of

domestic cat breeds and species survival plans of wild felids allow breeding in older animals, advancing the reproductive age for these populations. The results presented here suggest that advanced parental age may lead to the production of offspring with parental-age related maladies and diseases, as in humans (Hassold and Hunt 2001; Malaspina et al. 2002; Kong et al. 2012). Therefore, population managers may need to become more vigilant to the health problems in offspring associated with advanced paternal age.

Methods

Samples and sequencing

Archived whole blood samples were provided from 19 domestic cats in an extended pedigree taken from a cross-bred colony of cat breeds maintained for biomedical disease model characterization. The disease models did not compromise growth or reproductive maturity of the cats (Lyons et al. 2004; Lyons et al. 2016; Cogné et al. 2020). All animal care and use was conducted in accordance with policies and guidelines approved by the IACUC of the University of Missouri, protocol 8313. In addition, cultured fibroblasts from a trio of three domestic cats were isolated at the National Cancer Institute, Poolesville animal colony.

Genomic DNA was isolated from samples for whole genome sequencing. Standard PCR-free libraries were prepared using KAPA Hyper PCR-free library reagents (KAPA Biosystems). Total genomic DNA was sheared into fragments of approximately 200-600 bp and purified using AMPure XP beads. Sheared DNA molecules were subjected to double size selection with different ratios of AMPure XP beads. This was followed up with DNA end-repair and 3'-adenylation before the ligation of barcoded adapters. Library quality was evaluated by fragment analysis and qPCR assay. These WGS libraries were sequenced on an Illumina HiSeq X, producing 150 bp paired-end reads.

Mapping and variant calling

Sequenced reads were aligned with BWA-MEM v. 0.7.12-r1039 (Li 2013) to the domestic cat reference genome Felis_catus_9.0 (Buckley et al. 2020). Picard MarkDuplicates v. 1.105 (Broad Institute 2019) was used to identify and mark duplicate reads from the BAM files.

We used GATK version 3.6 (Van der Auwera et al. 2013) to call variants using best practices. HaplotypeCaller was used to generate gVCF files for each sample and joint genotype calling across samples was performed with GenotypeGVCFs. We applied GATK hard filters: `(SNPs: "QD < 2.0 || FS > 60.0 || MQ < 40.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0"; Indels: "QD < 2.0 || FS > 200.0 || ReadPosRankSum < -20.0")` and removed calls that failed.

Filters for candidate mutations

An initial set of candidate mutations was identified as Mendelian violations in each trio. Specifically, we looked for violations where both parents were reference homozygous and the offspring was heterozygous for an alternate allele. As this is the most common type of genotyping error (Wang et al. 2021), we then applied the following stringent filters to this initial set of candidates to get a set of high-confidence candidates:

- Read depth at the candidate site must be between 20 and 60 for every individual in the trio. Sites with too few reads may be sampling errors while sites with too many reads may be problematic repetitive regions (Li 2014).
- High genotype quality (GATK) in all individuals, (GQ > 70).
- Candidate mutations must be present on both the forward and reverse strand in the offspring (ADF, ADR > 0).
- Candidate mutation must not be present in any reads from either parent (AD = 0).
- Candidate mutation must not be present in any other unrelated sample.
- Candidate mutation must not have low allelic depth in the offspring (Allelic Balance > 0.35).

We assessed the sensitivity of our mutation rate estimates across a range of stringency criteria and found them to be in good agreement across reasonable filter limits (see Fig. S1).

Estimating site callability

To convert the raw number of candidate mutations into an estimate of the mutation rate, we applied existing strategies that considered differences in coverage and filtering (see Besenbacher et al. 2019; Wang et al. 2020). In brief, the number of identified mutations was divided by the total callable sites: a product of the number of sites covered by the appropriate sequencing depth and the estimated probability that a site would be called correctly given that it was a true de novo mutation. The mutation rate is then calculated as:

$$\mu_{s,i} = \frac{N_{\mathrm{mut},i}}{2 * \sum_x C_i(x)},$$

where $\mu_{s,i}$ is the per-base mutation rate for trio i, $N_{\mathrm{mut},i}$ is the number of *de novo* mutations in trio i, and $C_i(x)$ is the callability of site x in that trio. This strategy assumes the ability to call each individual in the trio correctly is independent, allowing us to estimate $C_i(x)$ as:

$$C_i(x) = C_c(x)C_p(x)C_m(x)$$

where C_c, C_p, and C_m are the probability of calling the child, father, and mother correctly for trio i. These values are estimated by applying the same set of stringent filters to high-confidence calls from each trio. For heterozygous variants in the child,

$$C_c(x) = \frac{N_{\mathrm{het,filtered}}}{N_{\mathrm{het,all}}}$$

where $N_{\mathrm{het,all}}$ is the number of variants where one parent is homozygous reference and the other parent is homozygous alternate, leading to high confidence in the child heterozygote call, and $N_{\mathrm{het,filtered}}$ is the set of such calls that pass our child-specific candidate mutation filters. The parental callability, $C_p(x)$ and $C_m(x)$, were estimated in a similar manner, by calculating the proportion of remaining sites in each after the application of the stringent mutation filters.

Reproductive longevity model of the mutation spectrum

We assume the relationship between parental age and the number of mutations in the offspring is independent for each parent (i.e. no interaction between maternal and paternal age). Let c be the class for a given mutation, where $c \in$ { A>C, A>G, A>T, C>A, C>G, C>T }. Let $N_{c,i}$ be the number of mutations of class c in individual i from a given parent. We model

$$N_{c,i} \sim \text{Poisson}(\mu_{c,\text{i}})$$

where $\mu_{c,i}$ is the per-generation rate of class c mutations in individual i, and is fit to a Poisson regression using an identity link with

$$\mu_{c,i} = \beta_0^c + \beta_1^c X_i$$

where X_i is the age at conception of the given parent for individual i. Let $N'_c(X')$ be the predicted number of mutations of class c at a given parental age X'. The predicted set of mutations from a parent with age X' is then

$$N_{\text{parent}}(X') = \langle N'_{\text{A>C}}(X'), N'_{\text{A>G}}(X'), N'_{\text{A>T}}(X'), N'_{\text{C>A}}(X'), N'_{\text{C>G}}(X'), N'_{\text{C>T}}(X') \rangle$$

The spectrum for an individual with parents having ages X'_{mother} and X'_{father} can then be calculated as

$$f(X'_{\text{mother}}, X'_{\text{father}}) = \frac{N_{\text{mother}}(X'_{\text{mother}}) + N_{\text{father}}(X'_{\text{father}})}{\sum N_{\text{mother}}(X'_{\text{mother}}) + \sum N_{\text{father}}(X'_{\text{father}})}.$$

This model is initially fit with the mutations and parental ages from a large human dataset (Jónsson et al. 2017). From this fit, the cat mutation spectrum under the total longevity model was then predicted at the average parental age of individuals in our sample; i.e., as $f(X'_{\text{mother}} = 3.8 \text{ years}, X'_{\text{father}} = 3.8 \text{ years})$. Under the reproductive longevity model, we shifted the average parental age in our sample by the time to puberty in humans (13 years), predicting a cat mutation

spectrum for our sample that resembles a human spectrum at 16.8 years ($X'_{\text{mother}} = X'_{\text{father}} = 16.8\ \text{years}$). To compare the two predictions, we calculated the root mean square error to the observed cat spectrum across all mutation classes.

Acknowledgements

This work was funded by the Precision Health Initiative of Indiana University with support for LAL from the University of Missouri, Gilbreath McLorn Endowment.

References

Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Peltonen L, et al. 2010. Integrating common and rare genetic variation in diverse human populations. *Nature* 467:52–58.

Amster G, Sella G. 2016. Life history effects on the molecular clock of autosomes and sex chromosomes. *PNAS* 113:1588–1593.

Bae T, Tomasini L, Mariani J, Zhou B, Roychowdhury T, Franjic D, Pletikos M, Pattni R, Chen B-J, Venturini E, et al. 2018. Different mutational rates and mechanisms in human cells at pre-gastrulation and neurogenesis. *Science* 359:550–555.

Bergeron LA, Besenbacher S, Bakker J, Zheng J, Li P, Pacheco G, Sinding M-HS, Kamilari M, Gilbert MTP, Schierup MH, et al. 2020. The germline mutational process in rhesus macaque and its implications for phylogenetic dating. *bioRxiv*:2020.06.22.164178.

Besenbacher S, Hvilsom C, Marques-Bonet T, Mailund T, Schierup MH. 2019. Direct estimation of mutations in great apes reconciles phylogenetic dating. *Nature Ecology & Evolution* 3:286–292.

Bredemeyer KR, Harris AJ, Li G, Zhao L, Foley NM, Roelke-Parker M, O'Brien SJ, Lyons LA, Warren WC, Murphy WJ. 2020. Ultracontinuous single haplotype genome assemblies for the domestic cat (*Felis catus*) and Asian leopard cat (*Prionailurus bengalensis*). *Journal of Heredity*.

Broad Institute. 2019. Picard Toolkit. Github Repository. http://broadinstitute.github.io/picard/: Broad Institute Available from: http://broadinstitute.github.io/picard/

Buckley RM, Davis BW, Brashear WA, Farias FHG, Kuroki K, Graves T, Hillier LW, Kremitzki M, Li G, Middleton RP, et al. 2020. A new domestic cat genome assembly based on long sequence reads empowers feline genomic medicine and identifies a novel gene for dwarfism. *PLOS Genetics* 16:e1008926.

Campbell CR, Tiley GP, Poelstra JW, Hunnicutt KE, Larsen PA, Lee H-J, Thorne JL, Reis M dos, Yoder AD. 2020. Pedigree-based measurement of the *de novo* mutation rate in the gray mouse lemur reveals a high mutation rate, few mutations in CpG sites, and a weak sex bias. *bioRxiv*:724880.

Carlson J, DeWitt WS, Harris K. 2020. Inferring evolutionary dynamics of mutation rates through the lens of mutation spectrum variation. *Current Opinion in Genetics & Development* 62:50–57.

Chintalapati M, Moorjani P. 2020. Evolution of the mutation rate across primates. *Current Opinion in Genetics & Development* 62:58–64.

Cogné B, Latypova X, Senaratne LDS, Martin L, Koboldt DC, Kellaris G, Fievet L, Le Meur G, Caldari D, Debray D, et al. 2020. Mutations in the Kinesin-2 motor KIF3B cause an autosomal-dominant ciliopathy. *The American Journal of Human Genetics* 106:893–904.

Crow JF. 2000. The origins, patterns and implications of human spontaneous mutation. *Nature Reviews Genetics* 1:40–47.

Driscoll CA, Menotti-Raymond M, Roca AL, Hupe K, Johnson WE, Geffen E, Harley EH, Delibes M, Pontier D, Kitchener AC, et al. 2007. The near eastern origin of cat domestication. *Science* 317:519–523.

Figueiró HV, Li G, Trindade FJ, Assis J, Pais F, Fernandes G, Santos SHD, Hughes GM, Komissarov A, Antunes A, et al. 2017. Genome-wide signatures of complex introgression and adaptive evolution in the big cats. *Science Advances* 3:e1700299.

Francioli LC, Polak PP, Koren A, Menelaou A, Chun S, Renkens I, Genome of the Netherlands Consortium, van Duijn CM, Swertz M, Wijmenga C, et al. 2015. Genome-wide patterns and properties of *de novo* mutations in humans. *Nature Genetics* 47:822–826.

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. *PNAS* 116:9491–9500.

Gao Z, Wyman MJ, Sella G, Przeworski M. 2016. Interpreting the dependence of mutation rates on age and time. *PLOS Biology* 14:e1002355.

Genova F, Longeri M, Lyons LA, Bagnato A, Gandolfi B, Aberdein D, Alves PC, Barsh GS, Beale HC, Bergström TF, et al. 2018. First genome-wide CNV mapping in *Felis catus* using next generation sequencing data. *BMC Genomics* 19:895.

Goldmann JM, Veltman JA, Gilissen C. 2019. *De novo* mutations reflect development and aging of the human germline. *Trends in Genetics* 35:828–839.

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. *Nature Genetics* 48:935–939.

Haldane JBS. 1935. The rate of spontaneous mutation of a human gene. *Journ. of Genetics* 31:317.

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*:079863.

Harris K, Pritchard JK. 2017. Rapid evolution of the human mutation spectrum. *eLife* 6:e24284.

Hassold T, Hunt P. 2001. To err (meiotically) is human: The genesis of human aneuploidy. *Nature Reviews Genetics* 2:280–291.

Jonsson H, Magnusdottir E, Eggertsson HP, Stefansson OA, Arnadottir GA, Eiriksson O, Zink F, Helgason EA, Jonsdottir I, Gylfason A, et al. 2021. Differences between germline genomes of monozygotic twins. *Nature Genetics* 53:27–34.

Jónsson H, Sulem P, Arnadottir GA, Pálsson G, Eggertsson HP, Kristmundsdottir S, Zink F, Kehr B, Hjorleifsson KE, Jensson BÖ, et al. 2018. Multiple transmissions of de novo mutations in families. *Nat Genet* 50:1674–1680.

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.

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.

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

Li G, Figueiró HV, Eizirik E, Murphy WJ. 2019. Recombination-aware phylogenomics reveals the structured genomic landscape of hybridizing cat species. *Molecular Biology and Evolution* 36:2111–2126.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *Q-bio, arXiv:1303.3997* [Internet]. Available from: http://arxiv.org/abs/1303.3997

Li H. 2014. Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics* 30:2843–2851.

Lyons LA, Biller DS, Erdman CA, Lipinski MJ, Young AE, Roe BA, Qin B, Grahn RA. 2004. Feline polycystic kidney disease mutation identified in PKD1. *JASN* 15:2548–2555.

Lyons LA, Creighton EK, Alhaddad H, Beale HC, Grahn RA, Rah H, Maggs DJ, Helps CR, Gandolfi B. 2016. Whole genome sequencing in cats, identifies new models for blindness in AIPL1 and somite segmentation in HES7. *BMC Genomics* 17:265.

Malaspina D, Brown A, Goetz D, Alia-Klein N, Harkavy-Friedman J, Harlap S, Fennig S. 2002. Schizophrenia risk and paternal age: A potential role for *de novo* mutations in schizophrenia vulnerability genes. *CNS Spectrums* 7:26–29.

Nachman MW. 2004. Haldane and the first estimates of the human mutation rate. *J Genet* 83:231–233.

Ottoni C, Van Neer W, De Cupere B, Daligault J, Guimaraes S, Peters J, Spassov N, Prendergast ME, Boivin N, Morales-Muñiz A, et al. 2017. The palaeogenetics of cat dispersal in the ancient world. *Nature Ecology & Evolution* 1:1–7.

Parent A-S, Teilmann G, Juul A, Skakkebaek NE, Toppari J, Bourguignon J-P. 2003. The timing of normal puberty and the age limits of sexual precocity: variations around the world, secular trends, and changes after migration. *Endocrine Reviews* 24:668–693.

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. *Nature Genetics* 48:126–133.

Sasani TA, Pedersen BS, Gao Z, Baird L, Przeworski M, Jorde LB, Quinlan AR. 2019. Large, three-generation human families reveal post-zygotic mosaicism and variability in germline mutation accumulation. *eLife* 8:e46922.

Scally A. 2016. Mutation rates and the evolution of germline structure. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371:20150137.

Scally A, Durbin R. 2012. Revising the human mutation rate: implications for understanding human evolution. *Nat Rev Genet* 13:745–753.

Ségurel L, Wyman MJ, Przeworski M. 2014. Determinants of mutation rate variation in the human germline. *Annual Review of Genomics and Human Genetics* 15:47–70.

Tarver JE, dos Reis M, Mirarab S, Moran RJ, Parker S, O'Reilly JE, King BL, O'Connell MJ, Asher RJ, Warnow T, et al. 2016. The interrelationships of placental mammals and the limits of phylogenetic inference. *Genome Biology and Evolution* 8:330–344.

Thomas GWC, Hahn MW. 2014. The human mutation rate is increasing, even as it slows. *Mol Biol Evol* 31:253–257.

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. *Current Biology* 28:3193–3197.

Tsutsui T, Nakagawa K, Hirano T, Nagakubo K, Shinomiya M, Yamamoto K, Hori T. 2004. Breeding season in female cats acclimated under a natural photoperiod and interval until puberty. *Journal of Veterinary Medical Science* 66:1129–1132.

Uchimura A, Higuchi M, Minakuchi Y, Ohno M, Toyoda A, Fujiyama A, Miura I, Wakana S, Nishino J, Yagi T. 2015. Germline mutation rates and the long-term phenotypic effects of mutation accumulation in wild-type laboratory mice and mutator mice. *Genome Res.* 25:1125–1134.

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Angel G del, Levy□Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, et al. 2013. From FastQ data to high-confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics* 43:11.10.1-11.10.33.

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

Walter CA, Intano GW, McCarrey JR, McMahan CA, Walter RB. 1998. Mutation frequency declines during spermatogenesis in young mice but increases in old mice. *Proc Natl Acad Sci U S A* 95:10015–10019.

Wang RJ, Radivojac P, Hahn MW. 2021. Distinct error rates for reference and nonreference genotypes estimated by pedigree analysis. *Genetics* 217:1–10.

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

Wilson Sayres MA, Makova KD. 2011. Genome analyses substantiate male mutation bias in many species. *BioEssays* 33:938–945.

Wu FL, Strand AI, Cox LA, Ober C, Wall JD, Moorjani P, Przeworski M. 2020. A comparison of humans and baboons suggests germline mutation rates do not track cell divisions. *PLOS Biology* 18:e3000838.

Figures and Tables

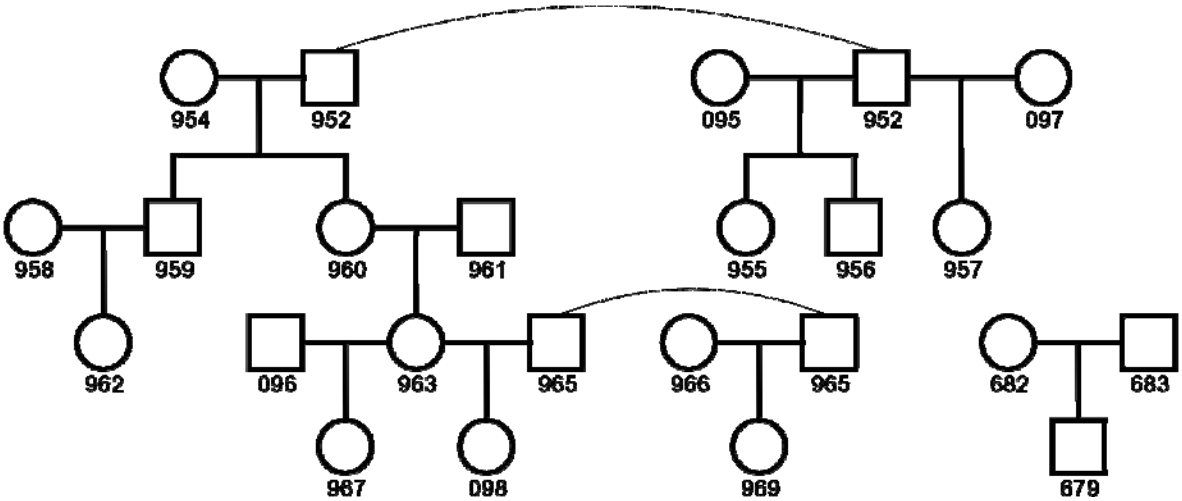

Figure 1. Pedigree of sequenced individuals

The 22 total individuals sequenced form 11 separate trios for mutation analysis. Whole blood was sampled from the 19 individuals in the larger pedigree, while cultured fibroblasts were sampled from the 3 individuals in the standalone trio. Dashed lines connect identical individuals in the pedigree.

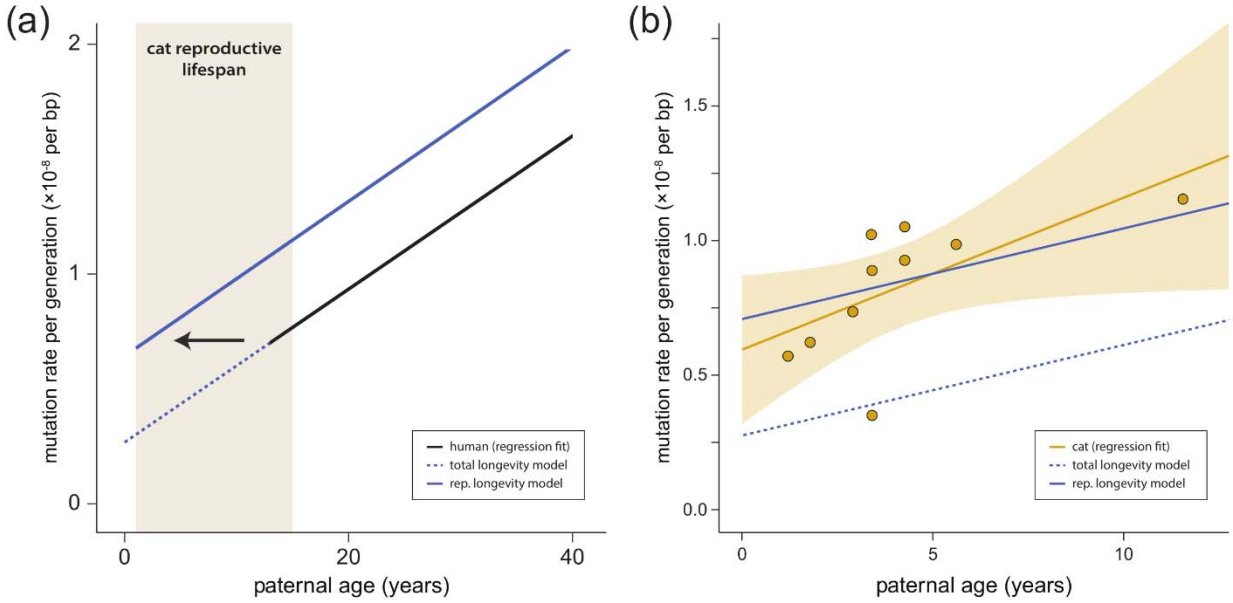

Figure 2. Mutation rates increase with paternal age.
(a) The predicted per-generation mutation rate in the domestic cat based on human data (regression fit) using the reproductive longevity model and the total longevity model. (b) The per-generation mutation rate increases with paternal age in the domestic cat; each point represents mutation data from one trio. The reproductive longevity model provides a much better fit to the observed pattern with age (*t*-test of residuals; $p = 0.0002$).

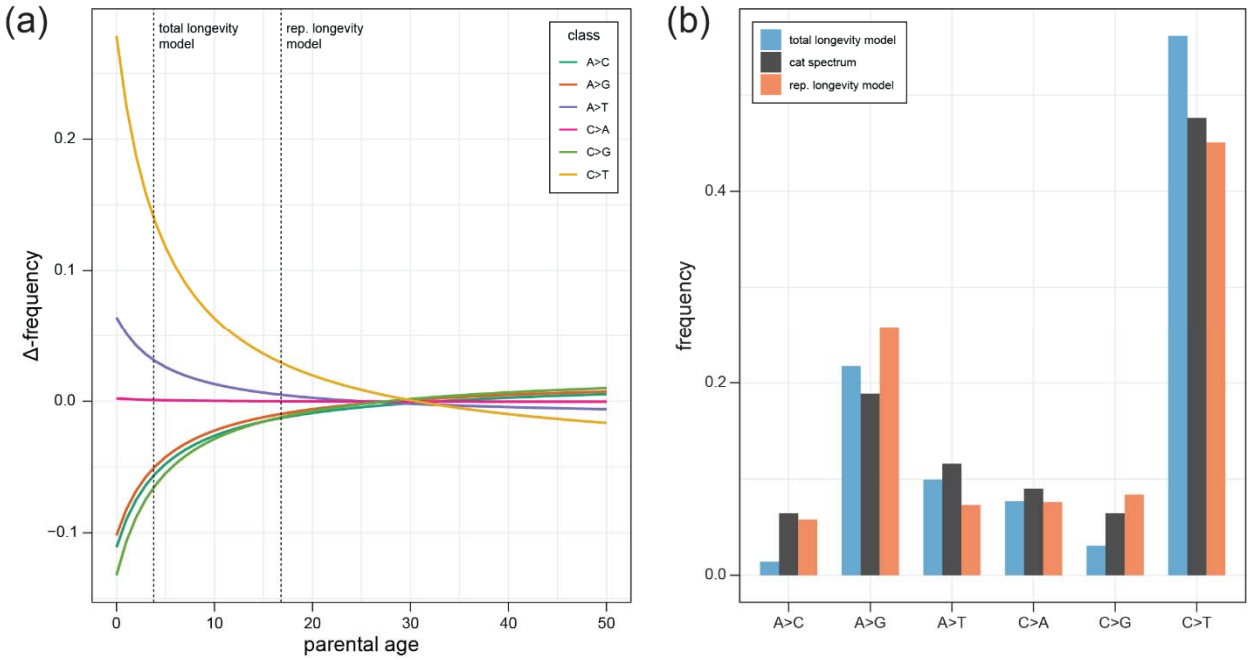

Figure 3. Longevity models of the mutation spectrum.
(a) The predicted change in each mutation class as a function of average parental age (colored lines). The predictions come from fitting observed human data using a Poisson regression (see main text). The vertical lines indicate the spectrum that would be predicted under a total longevity model and a reproductive longevity model.
(b) The observed cat mutation spectrum (black bars) compared to the predicted frequency of each mutation class under a total longevity model and a reproductive longevity model.

Table 1. Mutation counts and the per generation mutation rate

Proband ID[1]	Age at conception (y)		Mutations	Mean depth[2]	Callability	Haploid Size (Mb)	Rate ($\times 10^{-8}$)
	Paternal	Maternal					
679	-	-	22	47.8	0.644	1493	1.14
955	3.9	4.4	21	38.8	0.714	1655	0.89
956	3.9	4.4	7	37.0	0.620	1611	0.35
957	12.0	1.4	29	42.7	0.751	1674	1.15
959	4.8	3.3	20	39.5	0.671	1609	0.93
960	4.8	3.3	24	38.2	0.695	1643	1.05
962	3.4	2.1	20	42.0	0.723	1880	0.74
963	2.3	2.8	17	40.6	0.711	1923	0.62
967	1.7	2.7	17	39.4	0.719	2070	0.57
969	3.9	1.6	27	39.7	0.676	1952	1.02
098	6.1	3.1	29	42.5	0.729	2018	0.99

[1] ID of offspring for each independent trio in Figure 1
[2] Mean read depth across individuals of the trio with offspring as proband

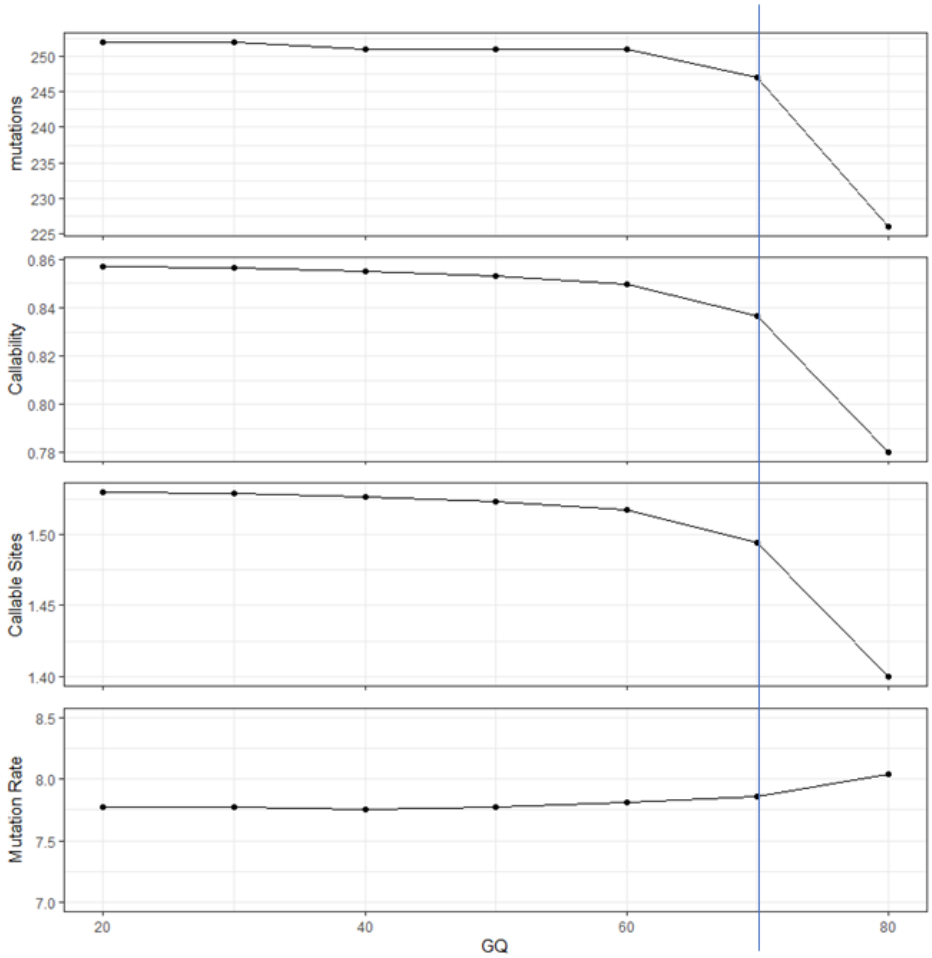

Figure S1. Mutation rate estimation with increasing stringency of genotype quality (GQ) filter.

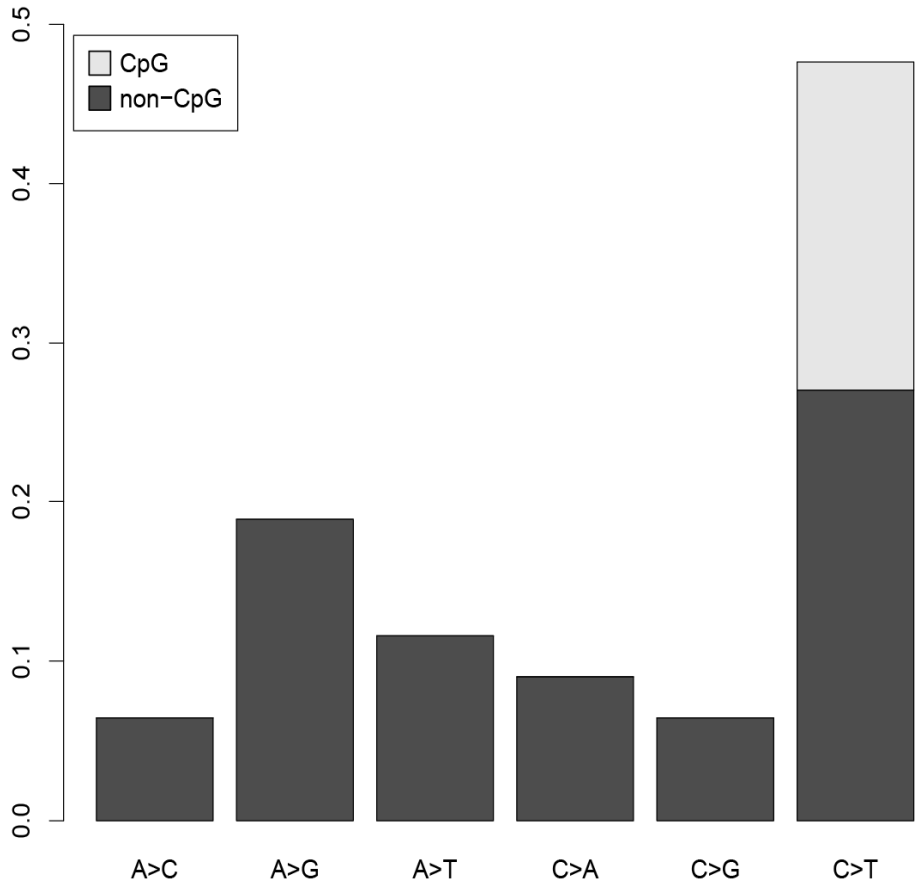

Figure S2. Cat mutation spectrum

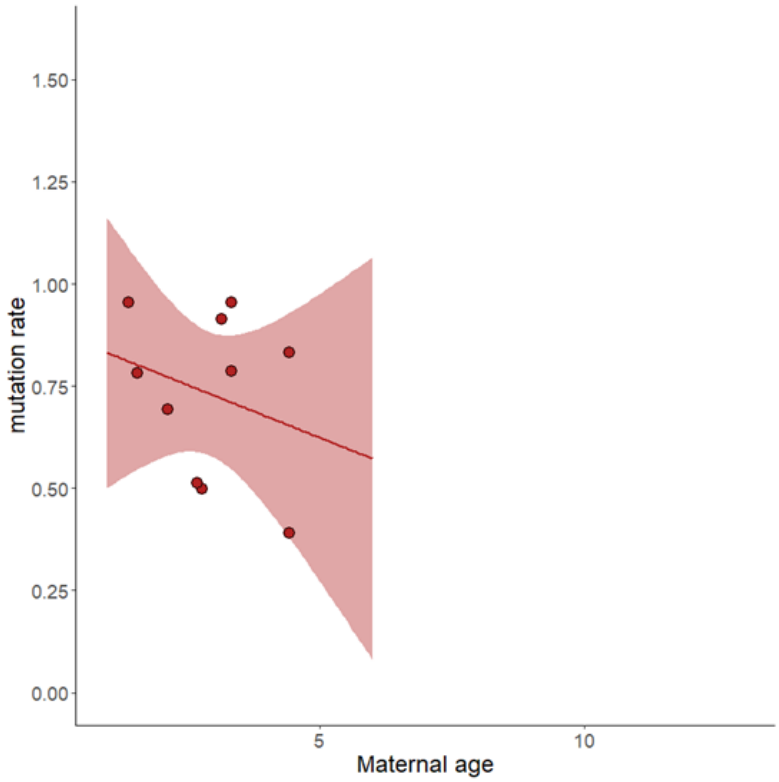

Figure S3. Mutation rate with maternal age