# A mitochondrial genetic divergence proxy predicts the reproductive compatibility of mammalian hybrids

Richard Allen[1*], Hannah Ryan[1*], Brian W. Davis[2], Charlotte King[9,3], Laurent Frantz[1,7], Evan Irving-Pease[1], Ross Barnett[1], Anna Linderholm[1,5], Liisa Loog[1,6], James Haile[1], Ophélie Lebrasseur[1,10], Mark White[3], Andrew C. Kitchener[4,8], William J. Murphy[2], Greger Larson[1]

[1]Palaeogenomics & Bio-Archaeology Research Network, Research Laboratory for Archaeology and the History of Art, University of Oxford, Oxford OX1 3QY, UK
[2]Veterinary Integrative Biosciences, Texas A&M University, College Station TX 77843 USA
[3]Department of Archaeology, Durham University, Science Site, Durham, UK DH1 3LE
[4]Department of Natural Sciences, National Museums Scotland, Chambers Street, Edinburgh EH1 IJF, UK
[5] Department of Anthropology, Texas A&M University, College Station, TX 77843-4352, USA
[6]Department of Genetics , University of Cambridge, Downing Street, Cambridge CB2 3EH, United Kingdom
[7] School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK
[8]Institute of Geography, School of Geosciences, University of Edinburgh, Drummond Street, Edinburgh EH9 3PX, UK
[9]Department of Anatomy, University of Otago, Great King Street, Dunedin 9016, NZ
[10]Department of Archaeology, Classics and Egyptology, 12-14 Abercromby Square, University of Liverpool, L69 7WZ

* These authors contributed equally to this study.

**Corresponding authors**: Greger Larson: greger.larson@arch.ox.ac.uk, William J. Murphy: wmurphy@cvm.tamu.edu

## Abstract

Despite the significant genetic barriers that arise during the process of reproductive isolation between species, numerous hybrid offspring of divergent mammalian species pairs have been catalogued. In some cases, these hybrids are only able to produce offspring through backcrosses with a parent species due to unisexual sterility (Haldane's Rule). In other instances, F1 hybrids are able to produce F2s through matings with F1s. Here, we explicitly tested whether genetic distance can be used as a proxy to predict the relative fertility of the hybrid offspring resulting from numerous pairs of terrestrial mammals. We assessed the proxy's predictive power using a well characterised felid hybrid system, and applied it to modern and ancient hominins. Our results revealed a small overlap in mitochondrial genetic divergence values that distinguish species pairs whose divergence values fall within two categories: those whose hybrid offspring follow Haldane's Rule, and those whose hybrid offspring can produce F2s. The strong correlation between genetic distance and hybrid fertility demonstrated here suggests that this proxy can be employed to predict whether the hybrid offspring of two mammalian species will follow Haldane's Rule.

## Introduction

Though hybrids between mammalian species have been catalogued for decades [1], the extent and frequency of gene flow between evolutionarily divergent taxa has only been recognised since the availability of high-coverage nuclear genomes. Recent studies have revealed rampant gene flow between multiple species of bears[2], canids[3], felids[4–6], and cetaceans[7,8]. Genome analyses of other vertebrate and invertebrate lineages beyond mammals have also revealed similarly extensive patterns of ancient and contemporary introgression [9–14].

This demonstrated frequency of genomic introgression is perhaps surprising given the significant barriers that maintain reproductive isolation in species pairs that diverged up to millions of years ago. In mammals, genomic barriers manifest in accordance with Haldane's Rule [15] as the unisexual sterility of the heterogametic sex (XY males) in F1 hybrid offspring. In cases where matings between F1s fail to produce F2s, fertile offspring can often be produced through backcrosses between the fertile F1 females and males from one of the parent species. Occasionally,

however, F1s produced from interbreeding between closely related mammalian species pairs can result in viable and fertile F2 offspring.

If genetic divergence values between two species correlated with the ability of their hybrid offspring to produce F2s, these values could serve as a proxy to predict this occurrence. Though at least one study [16] reported that genetic divergence values do generally correlate with species boundaries, others [17,18] have questioned whether this correlation exists, and have instead stated that measures of species divergence are not reliable predictors of hybrid sterility. A recent empirical study of damselflies, however, demonstrated a strong correlation between the genetic distances between species pairs and their relative reproductive isolation [19].

Establishing whether genetic distance and reproductive isolation are correlated is also critical for our understanding of the genetic architecture of reproductive isolation. Doing so firstly requires knowing whether any two species are capable of producing viable or fertile offspring, but there is a general paucity of captive breeding experiments or field data that have unequivocally established this. An alternative approach is to develop a metric that can accurately predict the relative fertility of the F1 hybrids of any two species that makes use of interspecific crosses whose offspring have been reproductively characterized. Here we developed a robust, quantitative framework based on the correlation between mitochondrial genetic distance between mammalian species known to produce F1spairs to obtain a quantitative measure of whether F1 hybrids of both sexes are likely to be capable of breeding, or if they instead manifest Haldane's Rule. We tested the accuracy of the proxy in a well characterised felid hybrid system, and then applied it to a hominin case study to assess the relative potential sterility of hybrids between humans and their closest extinct relatives.

## Results and Discussion

### *Categorizing Hybrid Incompatibility*

We first explicitly defined two dichotomous categories along the spectrum of hybrid compatibility. Category 1 is defined by mammalian species pairs capable of producing fertile F1 offspring of both sexes that can reproduce without backcrossing with a parent species (even if there are observed asymmetries in gene flow and

variation in male fertility amongst the hybrids) (Table S1). Category 2 is defined by pairs of species that can produce viable F1 offspring, but follow Haldane's Rule, and thus only female F1s can reproduce by backcrossing with a parent species. Category 2 also includes species pairs whose hybrids are infertile (Table S1). We determined the categorical assignment of each species pair (Table S1, Fig. S1) by following a flowchart (Fig. S2) based upon empirical evidence derived from experimental studies of F1 hybrid fertility. We confidently placed seven species pairs into Category 1, and six others into Category 2.

Many additional live hybrid offspring have been reported in the literature than are included in Fig. 1 or in Fig. S1. We identified 17 species pairs known to produce viable offspring, but for which there was insufficient evidence to confidently assign them into either category (Table S2, Fig. S3). The framework and threshold values depicted in Fig. 1 allow us to predict the fertility of these offspring given the definitions described above and their placement into Categories 1 or 2. These pairs are listed in Table S2 and their relative positions are depicted in Fig. S3.

We then obtained published sequence data across all species (Table S3) from both the cytochrome b gene (*CYTB*) (n=1795) and complete mitochondrial genomes (n=30) (excluding the control region) from multiple individuals per species. By matching the phylogenies derived from the alignments to available nuclear species trees, and by only including sequences that fell into reciprocally monophyletic clades, we ensured that the selected mitochondrial sequences for each species were neither mislabelled, nor nuclear copies of mitochondrial genes (NUMTs), nor derived from hybrid populations. Using the sequence alignments, we then calculated average pairwise genetic distances between each species pair using both the number of raw differences, and differences scaled by several different nucleotide substitution models. In addition, we calculated genetic distances for four nuclear loci (*CHRNA1*, *GHR*, *ZFX*, *ZFY*) available between ten of the primate species pairs to test the correlation between the mitochondrial proxy and the nuclear genome estimates [20] (Fig. S4).

Plotting the calculated divergence estimates using *CYTB* revealed a small overlap in values associated with the two categories (Fig 1, Fig. S1). More specifically, the

Category 1 pair with the greatest divergence value was a pair of guinea pig species that were 8.0% divergent, and the Category 2 pair with the lowest divergence was a pair of vole species that were 7.2% divergent. Within this 0.8% overlapping region were several pairs of species that fall into both categories suggesting that this level of CYTB divergence is the region where F1 offspring begin to require a backcross to generate an F2. The existence of a genetic distance threshold separating the two categories also held true for the complete mitogenomes (Fig. S4).

In addition, both the male and female hybrid offspring of one of the two most divergent pairs, Domestic Pig (*Sus domesticus*) x Babirusa (*Babyrousa celebensis*), were shown to be infertile, and an exhaustive study that attempted to produce hybrids between mountain hares and European rabbits failed completely [21]. The significant divergence values between these species pairs corroborates previous studies showing that along the continuum of speciation, infertility in both sexes evolves prior to inviability [22–24]. In addition, a Mann-Whitney U test showed significantly lower genetic divergence values of species pairs in Category 1 relative to those in Category 2 ($p<6.3e-4$).

Importantly, the two categories of fertility defined here are not strictly linked with gene flow. For instance, though both male and female Category 1 hybrid offspring can reproduce without requiring a backcross with a parent species, gene flow asymmetries have been demonstrated in virtually all of these species pairs including house mice (*Mus musculus musculus* x *Mus musculus domesticus*) [25], and between brown bears (*Ursus arctos*) and polar bears (*Ursus maritimus*) [2] (Table S1). Gene flow has also been demonstrated between Category 2 species (including *Mus musculus* and *M. spretus [26]*). Since both fertility and the potential for gene flow vary along a continuum, it is striking that the divergence values associated with the two fertility categories defined here do not overlap more substantially.

The lack of available nuclear sequences (relative to mitochondria) limited our ability to test whether nuclear genes generally produced the same pattern as the mitochondria across all our species pairs. Despite this limitation, we were able to identify four nuclear loci: Zinc finger Y-chromosomal protein (*ZFY*), Zinc finger X-chromosomal protein (*ZFX*), Growth Hormone Receptor (*GHR*), Cholinergic

Receptor Nicotinic Alpha (*CHRNA1*), that have been sequenced in 10 primate pairs (Fig. S4) known to produce viable hybrid offspring [20]. We generated pairwise distances for each of these genes using the same method employed in the mitochondrial distance calculation. We then assigned each species pair to Category 1 or Category 2 based upon their *CYTB* divergence values within the original framework. In each case, though the order of the taxa based upon pairwise divergence values varied relative to the pattern generated using *CYTB* (owing to significantly fewer variable sites and thus smaller divergence values in nuclear loci), for the two most variable nuclear loci, *ZFX* and *GHR*, there was no overlap in divergence values between the two categories, consistent with the mitochondrial assessment (Fig. S4). The other two loci, *ZFY* and *CHRNA1*, possessed very limited interspecific nucleotide variability, but generally followed the same overall pattern.

***Testing the proxy using known mammalian hybrids***

In order to further substantiate both this correlation and the robustness of *CYTB* divergence as a proxy for hybrid sterility, we tested the utility of this system for predicting fertility in a well-known hybrid system. Cat breeders have crossed domestic cats (*Felis catus*) with several wild felids, including the Jungle cat (*F. chaus),* Leopard cat (*Prionailurus bengalensis*), and Serval (*Leptailurus serval*) [27], to create three exotic cat breeds: Chausies, Bengals, and Savannahs, respectively. In all cases, the F1 male hybrids are sterile. To regain fertility while maintaining some wild felid characteristics, breeders must backcross the F1 female offspring with male domestic cats to establish a breeding population of pets [27]. Given that multiple generations of unidirectional backcrossing was required for all three crosses to generate a fertile population, our proxy would firstly predict that the *CYTB* distances between all three pairs should be close to or greater than ~7.2%, and that they should all fall into the range encompassed by Category 2. Secondly, the pairs with larger genetic distance values should require a greater number of backcrosses with domestic cats (halving the wild cat ancestry with each subsequent generation) before fertility is restored in hybrid males and a breeding pet population is established.

Both of these predictions are borne out by the data (Fig 1, Table S3). All three pairs show *CYTB* distances greater than or equal to 7.5% and the increasing molecular

distances between the pairs correlate with an increase in the number of required backcross generations to regain fertility. Specifically, distances between domestic cats and Jungle cats, Leopard cats and Servals (7.54%, 10.94%, and 11.28% respectively) are consistent with both the observed minimum (2, 3, and 4 respectively) and average (3, 4, and 5, respectively) number of backcrosses with domestic cats required for hybrid males to acquire fertility [27]. These results are also consistent with an early hybrid experiment using guinea pigs in which hybrids between *Cavia fulgida* and *C. porcellus* (8.0% *CYTB* divergence) were able to regain male fertility after three generations of backcrossing [28](Table S1).

Accidental hybrids in zoos also confirm the predictive power of this proxy. In 2006, the Copenhagen Zoo placed a domestic sow (*Sus domesticus*) in a pen with a male babirusa (*Babyrousa celebensis*) with the expectation that the two species were sufficiently evolutionarily divergent that they would be incapable of producing offspring. Months later, however, five piglets were born and though two died from maternally induced trauma, the other three (two males, and one female), all survived and were shown to be infertile [29] (Table S1, Fig. S1). Historically, hybrid offspring between distantly related species have accidentally been produced in zoos (Table S2), though the relative fertility of the F1s was rarely established. In this case, the *CYTB* divergence between the two species (12.9%) is not much greater than the value between rhesus macaques and hamadryas baboons (12.5%) which were able to produce a live offspring, thus suggesting that live offspring between these suids was possible.

***Assessing hominin hybrid incompatibility***

The initial discovery of Neanderthals led some anthropologists, as early as the turn of the 20th century, to speculate that anatomically modern humans (AMH) and their closest extinct relatives were capable of producing hybrid offspring [30]. The absence of Neanderthal mitochondrial genomes in the extant human population, however, led some to suggest that AMH and Neanderthals did not hybridise [31–33]. More recent analyses of whole ancient genome sequences have demonstrated that, in fact, archaic hominins, including Neanderthals and Denisovans, did produce hybrid offspring, not only with AMH [34–36], but also with each other [37]. The generation of these ancient genomes has also allowed for an assessment of the role

that incompatibility may have played in the selection for and against hybrid introgression in modern humans [38]. The genomic confirmation of the existence of hominin hybrids supported the conclusions of two studies [39,40] that used a qualitative correlation between the divergence times between species pairs and the fertility of their hybrid offspring to suggest that, given their relatively recent temporal divergence, AMH and Neanderthals could have retained the ability to produce fertile offspring of both sexes.

We quantitatively assessed the relative fertility of hybrids between pairs of modern and ancient hominin lineages using the proxy established in this study. To do so, we calculated the average pairwise divergence in *CYTB* sequences between AMH and three extinct hominin lineages: Neanderthals, Denisovans, and the ancient population from the Sima de los Huesos cave in Spain [29,41]. To avoid overestimating the genetic divergence resulting from the comparison of modern and extinct populations, we generated distance values using the *CYTB* sequences derived solely from ancient AMH found in archaeological contexts (Supp Table 3).

The divergence values for each pairing of three *Homo* groups (Sima de los Huesos, Neanderthals, and AMH) occupy the bottom of the Category 1 range and are less than the divergence values between Polar Bears and Brown Bears, and between subspecific crosses of *Mus musculus* (Fig 1, Fig. S1, Table S1). When placed within this context, our data predict that ancient hominin lineages were likely not sufficiently divergent from each other to expect a significant biological impediment to the generation of fertile offspring. This is consistent with the ancient genomic evidence, which has shown not only that archaic populations interbred with AMH on at least four occasions [42], but also that introgression took place in both directions [43]. In addition, the divergence values of Denisovan-Neanderthal and Denisovan-AMH are the largest of the *Homo* pairings, and are consistent with the suggestion that Denisovans possessed a mitochondrial lineage that may have introgressed from another source population [44].

We also assessed hybrid sterility between more distantly diverged hominin lineages. Specifically, we calculated divergence values between humans and our two closest living relatives: chimpanzees (*Pan troglodytes*) and bonobos (*P. paniscus*). Female

chimpanzees inseminated with human sperm during a Soviet experiment in the 1920s failed to produce any offspring, and the reverse experiment did not progress beyond the planning stage [45]. Recent molecular clock assessments have suggested that AMH and chimps diverged ~5-6 Mya [46], well beyond both the two-million-year threshold suggested by other studies as the upper limit to hybrid fertility [39,40], and the average time to speciation [47]. Our analysis places the divergence values between AMH and chimps, and AMH and bonobos within Category 2 suggesting that even if hybrids could be produced, they would likely follow Haldane's Rule (Fig 1, Fig S1).

## Conclusions

The correlation demonstrated here between *CYTB* divergence (but also genetic divergence in general) and relative hybrid sterility suggests that distance values can be used as a proxy to accurately and rapidly predict the relative sterility of hybrids resulting from matings between pairs of mammalian species. Our emphasis here on mitochondrial DNA should not be misinterpreted as a claim that *CYTB* plays a causative role in hybrid sterility, though recent studies have proposed that speciation may be mediated by mitonuclear interactions [48,49]. Nor can the use of genetic divergence values as a proxy be perfectly predictive. For example, under the Dobzhansky-Müller model incompatibility can arise from as few as two mutations in isolated populations irrespective of time since divergence, meaning that it would be possible for closely related populations to be incapable of generating fertile hybrids [48,50], though no such examples have been described.

The value of any proxy is determined by both its predictive power and the ease of generating the proxy data. Publicly available mtDNA sequences from thousands of mammalian taxa already exist and calculating pairwise divergence values is inexpensive, simple and fast. As a result, mitogenomic distances have substantial value as a means to predict the potential for any two mammalian species to produce fertile offspring, and the relative degree of sterility in one or both sexes. As whole genomes become available from the same set of species, this analysis can be extended to determine which regions of the nuclear genome may also be more or less predictive.

The discovery of additional extinct hominin populations that survived into the last 250,000 years, including *H. floresiensis* [51] and *H. naledi* [52], has raised interest in understanding the limits to fertility and hybridization between extinct and extant *Homo* spp [53]. If and when mitochondrial genomes from these samples can be obtained, the approach described here may provide an answer, even if nuclear genomic data are not obtainable. Lastly, establishing which species pairs violate the predictions of the framework will identify unique systems that may lead to a better understanding of the process of reproductive isolation and the biological mechanisms responsible for hybrid sterility.

## Materials and Methods

### *Assessment of Hybrid Fertility and Rationale of Assignment into Categories*

In order to ascertain if there was a correlation between genetic divergence and the fertility of hybrid offspring between species, we first collected published examples of species pairs that were capable of producing live offspring. We then split the hybrid pairings into two categories. Category one consisted of seven species pairs that are capable of producing fertile F1 offspring of both sexes, and for which we were able to obtain evidence of captive breeding experiments showing that the F1s could mate to produce F2s. The evidence and rationale for placing each of these pairs into Category 1 is listed in Table S1 and the flow chart we used to determine the categorisation is shown in Table S2.

The hybrid offspring of all of six pairs of species in Category 2 are either completely infertile, or require one or more generations of female hybrid backcrosses with the male of a parent species to produce fertile offspring. For these pairs, we obtained evidence demonstrating no successful F2s from F1 hybrid pairings, an inability to produce offspring other than by backcrossing to a parent species, or other biological measurements (including histological assessments of the testes from the hybrid males) that demonstrated complete infertility (Table S1, Supp Fig. 5).

### *Genetic Distance Calculation*

Both *CYTB* sequences and full mitogenomes (excluding the control region) of multiple individuals of each species were collected from Genbank (Table S4) and aligned using Clustal Omega version 1.2.4 [54]. In order to ensure that none of the

sequences were either mislabelled, or were nuclear copies of mitochondrial genes, we constructed Neighbour-Joining trees using Geneious version 6.1.8 [55] and removed all individuals that did not fall into monophyletic clades consisting of individuals from each species. We first used *pModelTest* version 1.04 [56] to determine the best model for the alignment of each set of sequences for both species. We then calculated pairwise distances between each species pair using RAXML version 8 [57] and FASTTree version 2.1 [58]. We also generated raw distance values using the Hamming distance method which sums the number of base pair differences (ignoring transition or transversion status) and divides that number by the sequence length.

The distances were generated from the *CYTB* and nuclear gene alignments for each set of species pairings in fasta file format using a Python 2.7 wrapper to automate the terminal based programs RAXML, FastTree, and pModelTest. A custom Python 3.6 program was written to calculate Hamming distances of sequences making use of the distance v0.1.3 [59] and Biopython v1.66 [60] modules. Gaps in the aligned sequences were treated as missing data.

The mean distance and standard errors for each pairwise comparison were calculated using the bootstrapping method on the assumption that the sets of pairwise distances between related species would not be normally distributed. Each pairwise comparison group containing Hamming distances was randomly resampled into sets of equal sample size and processed using a helper function in the custom software which made use of the bootstrapped v0.0.2, NumPy v1.10.1 [61] and SciPy v0.16.0 [62] Python modules. The source code is available at https://github.com/BeebBenjamin/MrHamming. The *CYTB* distances were validated through comparison with data generated using the `compute between group mean distance` method in MEGA X for GNU/Linux [63], using the following settings:

a) Variance Estimation Method: Bootstrap Method
b) No of Bootstrap Replications: 500
c) Substitutions Type: Nucleotide
d) Model/Method: p-distance
e) Substitutions to Include:  d: Transitions + Transversions
f) Rates among Sites: Uniform Rates
g) Gaps/Missing/Data Treatment: Complete deletion

h) Select Codon Positions: 1st, 2nd, 3rd, Noncoding Site

Using a Student's t-test (two tailed), the differences between the results were found to be statistically non-significant (t=-0.11222, p= 0.912504).

***Mann-Whitney test of statistical difference between CYTB distance in hybrid categories***

The statistical significance of observed differences in CYTB divergence between Categories 1 and 2 was tested using the Mann-Whitney U test (p=6.216e-4) implemented in the R-software package [64].

## Figure Legends

### Figure 1

A depiction of the correlation between *CYTB* divergence between mammalian species pairs and the relative fertility of their hybrid offspring. In column A, the green circles represent species capable of producing fully fertile F1 offspring which can reproduce independently of their parent species (Category 1). Brown circles represent species pairs that follow Haldane's Rule and require backcrossing of a female F1 with a parent species, or both sexes are sterile (Category 2). The green and brown shaded regions represent the range of divergence values of the two categories. Column B depicts the divergence between three wild felid species and domestic cats, as well as the minimum number of generations of backcrosses with domestic cats before full fertility of the hybrid is restored. The white circles in Column C depict the divergence between three ancient hominins and AMH, as well as the distances between AMH and chimpanzees and bonobos (in Category 2). The asterisks represent those pairs that include modern samples of AMH. The lack of an asterisk signifies that only sequences derived from archaeological AMH were used to compute the divergence values. Details regarding the specific species pairs are listed in Fig. S1 and Table S1.

## Acknowledgments

We thank Simon Ho, Linda Maxson, Julie Wilson, Andrew Millard, John Hawks, Tom Higham, Kelly Harris, Christian Capelli, Shyam Gopalakirshnan, Montgomery Slatkin, Joshua Schraiber, Sam Turvey, Janet Kelso, and Murray Cox for advice and

discussion. We also thank the Zoological Society of London and the Bartlett Society for their assistance. GL was supported by the European Research Council (ERC-2013-StG 337574-UNDEAD) and the Natural Environment Research Council (NE/H005269/1 & NE/K005243/1).

## References

1. Gray AP. 1971 *Mammalian hybrids: A check-list with bibliography*. 2nd, revised, ed. Farnham Royal: Commonwealth Agricultural Bureaux.

2. Cahill JA, Stirling I, Kistler L, Salamzade R, Ersmark E, Fulton TL, Stiller M, Green RE, Shapiro B. 2015 Genomic evidence of geographically widespread effect of gene flow from polar bears into brown bears. *Mol. Ecol.* **24**, 1205–1217.

3. Gopalakrishnan S *et al.* 2019 Interspecific Gene Flow Shaped the Evolution of the Genus Canis. *Curr. Biol.* **29**, 4152.

4. Li G, Davis BW, Eizirik E, Murphy WJ. 2016 Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). *Genome Res.* **26**, 1–11.

5. Figueiró HV *et al.* 2017 Genome-wide signatures of complex introgression and adaptive evolution in the big cats. *Sci Adv* **3**, e1700299.

6. 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*. (doi:10.1093/molbev/msz139)

7. Árnason Ú, Lammers F, Kumar V, Nilsson MA, Janke A. 2018 Whole-genome sequencing of the blue whale and other rorquals finds signatures for introgressive gene flow. *Sci Adv* **4**, eaap9873.

8. Skovrind M, Castruita JAS, Haile J, Treadaway EC, Gopalakrishnan S, Westbury MV, Heide-Jørgensen MP, Szpak P, Lorenzen ED. 2019 Hybridization between two high Arctic cetaceans confirmed by genomic analysis. *Sci. Rep.* **9**, 7729.

9. Martin SH *et al.* 2013 Genome-wide evidence for speciation with gene flow in Heliconius butterflies. *Genome Res.* **23**, 1817–1828.

10. Martin SH, Davey JW, Salazar C, Jiggins CD. 2019 Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS Biol.* **17**, e2006288.

11. Edelman NB *et al.* 2019 Genomic architecture and introgression shape a butterfly radiation. *Science* **366**, 594–599.

12. Fontaine MC *et al.* 2015 Mosquito genomics. Extensive introgression in a malaria vector species complex revealed by phylogenomics. *Science* **347**, 1258524.

13. Lamichhaney S *et al.* 2015 Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* **518**, 371–375.

14. Runemark A, Trier CN, Eroukhmanoff F, Hermansen JS, Matschiner M, Ravinet M, Elgvin TO, Sætre G-P. 2018 Variation and constraints in hybrid genome formation. *Nat Ecol Evol* **2**, 549–556.

15. Haldane JBS. 1922 Sex ratio and unisexual sterility in hybrid animals. *J. Genet.* **12**, 101–109.

16. Bradley RD, Baker RJ. 2001 A Test of the Genetic Species Concept: Cytochrome-b Sequences and Mammals. *J. Mammal.* **82**, 960–973.

17. Jančúchová-Lásková J, Landová E, Frynta D. 2015 Are genetically distinct lizard species able to hybridize? A review. *Curr. Zool.* **61**, 155–180.

18. Edmands S. 2002 Does parental divergence predict reproductive compatibility? *Trends Ecol. Evol.* **17**, 520–527.

19. Sánchez-Guillén RA, Córdoba-Aguilar A, Cordero-Rivera A, Wellenreuther M. 2014 Genetic divergence predicts reproductive isolation in damselflies. *J. Evol. Biol.* **27**, 76–87.

20. Perelman P *et al.* 2011 A molecular phylogeny of living primates. *PLoS Genet.* **7**, e1001342.

21. Castle WE. 1925 The Hare-Rabbit, A Study in Evolution by Hybridization. *Am. Nat.* **59**, 280–283.

22. Presgraves DC. 2002 Patterns of postzygotic isolation in Lepidoptera. *Evolution* **56**, 1168.

23. Price TD, Bouvier MM. 2002 The evolution of F1 postzygotic incompatibilities in birds. *Evolution* **56**, 2083–2089.

24. Coyne JA, Orr HA. 1997 'Patterns of Speciation in Drosophila' Revisited. *Evolution* **51**, 295–303.

25. Teeter KC *et al.* 2008 Genome-wide patterns of gene flow across a house mouse hybrid zone. *Genome Res.* **18**, 67–76.

26. Song Y, Endepols S, Klemann N, Richter D, Matuschka F-R, Shih C-H, Nachman MW, Kohn MH. 2011 Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. *Curr. Biol.* **21**, 1296–1301.

27. Davis BW, Seabury CM, Brashear WA, Li G, Roelke-Parker M, Murphy WJ. 2015 Mechanisms Underlying Mammalian Hybrid Sterility in Two Feline Interspecies Models. *Mol. Biol. Evol.* **32**, 2534–2546.

28. Detlefsen JA. 1914 *Genetic studies on a cavy species cross*. Washington, D. C.: Carnegie Institution of Washington.

29. Thomsen PD, Schauser K, Bertelsen MF, Vejlsted M, Grøndahl C, Christensen K. 2011 Meiotic studies in infertile domestic pig-babirusa hybrids. *Cytogenet. Genome Res.* **132**, 124–128.

30. Arthur K. 1911 *Ancient Types of Man*. London: Harper & Brothers.

31. Currat M, Excoffier L. 2004 Modern Humans Did Not Admix with Neanderthals during Their Range Expansion into Europe. *PLoS Biol.* **2**, e421.

32. Serre D, Langaney A, Chech M, Teschler-Nicola M, Paunovic M, Mennecier P, Hofreiter M, Possnert G, Pääbo S. 2004 No evidence of Neandertal mtDNA contribution to early modern humans. *PLoS Biol.* **2**, E57.

488 33. Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S. 1997 Neandertal
489 DNA sequences and the origin of modern humans. *Cell* **90**, 19–30.

490 34. Green RE *et al.* 2010 A draft sequence of the Neandertal genome. *Science* **328**, 710–
491 722.

492 35. Meyer M *et al.* 2012 A high-coverage genome sequence from an archaic Denisovan
493 individual. *Science* **338**, 222–226.

494 36. Fu Q *et al.* 2015 An early modern human from Romania with a recent Neanderthal
495 ancestor. *Nature* **524**, 216–219.

496 37. Slon V *et al.* 2018 The genome of the offspring of a Neanderthal mother and a
497 Denisovan father. *Nature* **561**, 113–116.

498 38. Petr M, Pääbo S, Kelso J, Vernot B. 2019 Limits of long-term selection against
499 Neandertal introgression. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 1639–1644.

500 39. Holliday TW. 2006 Neanderthals and modern humans: an example of a mammalian
501 syngameon? In *Neanderthals Revisited: New Approaches and Perspectives* (eds J-J
502 Hublin, K Harvati, T Harrison), pp. 281–297. Dordrecht: Springer Netherlands.

503 40. Holliday TW, Gautney JR, Friedl L. 2014 Right for the Wrong Reasons. *Curr. Anthropol.*
504 **55**, 696–724.

505 41. Meyer M *et al.* 2013 A mitochondrial genome sequence of a hominin from Sima de los
506 Huesos. *Nature* **505**, 403–406.

507 42. Vernot B *et al.* 2016 Excavating Neandertal and Denisovan DNA from the genomes of
508 Melanesian individuals. *Science* **352**, 235–239.

509 43. Kuhlwilm M *et al.* 2016 Ancient gene flow from early modern humans into Eastern
510 Neanderthals. *Nature* **530**, 429–433.

511 44. Prüfer K *et al.* 2014 The complete genome sequence of a Neanderthal from the Altai
512 Mountains. *Nature* **505**, 43–49.

513 45. Etkind A. 2008 Beyond eugenics: the forgotten scandal of hybridizing humans and apes.
514 *Stud. Hist. Philos. Biol. Biomed. Sci.* **39**, 205–210.

515 46. Scally A *et al.* 2012 Insights into hominid evolution from the gorilla genome sequence.
516 *Nature* **483**, 169–175.

517 47. Hedges SB, Marin J, Suleski M, Paymer M, Kumar S. 2015 Tree of life reveals clock-like
518 speciation and diversification. *Mol. Biol. Evol.* **32**, 835–845.

519 48. Hill GE. 2016 Mitonuclear coevolution as the genesis of speciation and the
520 mitochondrial DNA barcode gap. *Ecol. Evol.* **6**, 5831–5842.

521 49. Ma H *et al.* 2016 Incompatibility between Nuclear and Mitochondrial Genomes
522 Contributes to an Interspecies Reproductive Barrier. *Cell Metab.* **24**, 283–294.

523 50. Orr HA, Turelli M. 2001 The evolution of postzygotic isolation: accumulating
524 Dobzhansky-Muller incompatibilities. *Evolution* **55**, 1085–1094.

525 51. Aiello LC. 2010 Five years of Homo floresiensis. *Am. J. Phys. Anthropol.* **142**, 167–179.

52. Berger LR, Hawks J, Dirks PH, Elliott M, Roberts EM. 2017 Homo naledi and Pleistocene hominin evolution in subequatorial Africa. *eLife*. **6**. (doi:10.7554/elife.24234)

53. Ovchinnikov IV. 2013 Hominin evolution and gene flow in the Pleistocene Africa. *Anthropologischer Anzeiger*. **70**, 221–227. (doi:10.1127/0003-5548/2013/0313)

54. Sievers F *et al.* 2011 Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539.

55. Kearse M *et al.* 2012 Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649.

56. Serra F. 2011 *pModelTest*. See https://github.com/etetoolkit/pmodeltest.

57. Stamatakis A. 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313.

58. Price MN, Dehal PS, Arkin AP. 2010 FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490.

59. Meyer M. 2013 *distance*. See https://github.com/doukremt/distance.

60. Cock PJA *et al.* 2009 Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423.

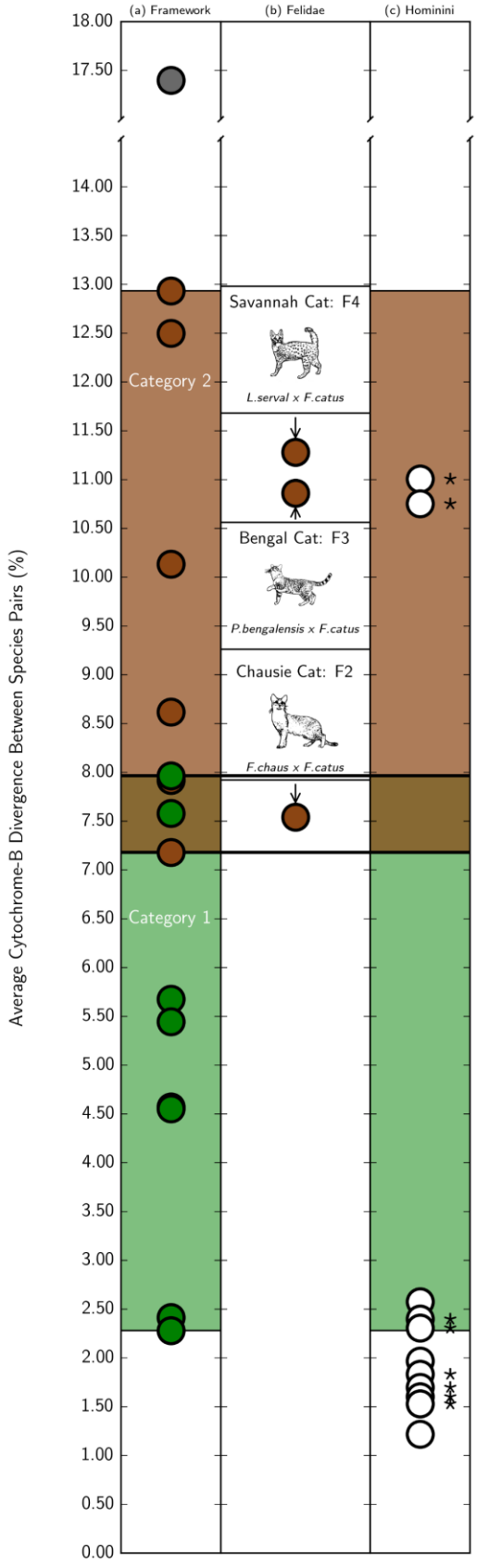
61. Oliphant TE. 2015 *Guide to NumPy*. 2nd edn. USA: CreateSpace Independent Publishing Platform.

62. Virtanen P *et al.* 2019 SciPy 1.0--Fundamental Algorithms for Scientific Computing in Python. *arXiv [cs.MS]*.

63. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018 MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**, 1547–1549.

64. R Core Team. 2017 *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See https://www.R-project.org/.

**Figure 1**


(a) Framework
(b) Felidae
(c) Hominini
Average Cytochrome-B Divergence Between Species Pairs (%)
18.00
17.50
14.00
13.50
13.00
12.50
12.00
11.50
11.00
10.50
10.00
9.50
9.00
8.50
8.00
7.50
7.00
6.50
6.00
5.50
5.00
4.50
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
Category 2
Category 1
Savannah Cat: F4
L.serval x F.catus
Bengal Cat: F3
P.bengalensis x F.catus
Chausie Cat: F2
F.chaus x F.catus