GBE

Genome Biology and Evolution

SMBE

Domestication and the Mitochondrial Genome: Comparing Patterns and Rates of Molecular Evolution in Domesticated Mammals and Birds and Their Wild Relatives

Camile Moray[1,*], Robert Lanfear[1,2], and Lindell Bromham[1]

[1]Centre for Macroevolution and Macroecology, Division of Evolution Ecology and Genetics, Research School of Biology, Australian National University, Canberra, Australian Capital Territory, Australia

[2]National Evolutionary Synthesis Center, Durham, NC

*Corresponding author: E-mail: camile.moray@anu.edu.au.

Accepted: December 30, 2013

Data deposition: GenBank accessions used in the analysis are provided in supplementary table S3, Supplementary Material online. All alignment and data files used in this analysis are available on the Dryad Digital Repository (http://doi.org/10.5061/dryad.d85ng).

Abstract

Studies of domesticated animals have led to the suggestion that domestication could have significant effects on patterns of molecular evolution. In particular, analyses of mitochondrial genome sequences from domestic dogs and yaks have yielded higher ratios of nonsynonymous to synonymous substitutions in the domesticated lineages than in their wild relatives. These results are important because they imply that changes to selection or population size operating over a short timescale can cause significant changes to the patterns of mitochondrial molecular evolution. In this study, our aim is to test whether the impact on mitochondrial genome evolution is a general feature of domestication or whether it is specific to particular examples. We test whether domesticated mammals and birds have consistently different patterns of molecular evolution than their wild relatives for 16 phylogenetically independent comparisons of mitochondrial genome sequences. We find no consistent difference in branch lengths or d_N/d_S between domesticated and wild lineages. We also find no evidence that our failure to detect a consistent pattern is due to the short timescales involved or low genetic distance between domesticated lineages and their wild relatives. However, removing comparisons where the wild relative may also have undergone a bottleneck does reveal a pattern consistent with reduced effective population size in domesticated lineages. Our results suggest that, although some domesticated lineages may have undergone changes to selective regime or effective population size that could have affected mitochondrial evolution, it is not possible to generalize these patterns over all domesticated mammals and birds.

Key words: relaxed selection, artificial selection, mitochondria, d_N/d_S, effective population size, comparative analysis.

Introduction

Does domestication influence rates and patterns of molecular evolution? Analysis of single-nucleotide polymorphisms from the dog nuclear genome suggests a higher ratio of nonsynonymous to synonymous alleles relative to the wolf, which has been interpreted as the signature of relaxed selection and reduction in effective population size associated with domestication (Cruz et al. 2008). Similarly, studies have found that rice (Lu et al. 2006) and a laboratory strain of yeast (Gu et al. 2005) have higher ratios of nonsynonymous to synonymous changes (d_N/d_S) than their wild relatives. Comparison of dog, yak, pig, and silkworm mitochondrial genomes with their respective wild relatives have also shown that the domesticated lineages have higher d_N/d_S than their wild relatives (Björnerfeldt et al. 2006; Wang et al. 2011; Hughes 2013).

These studies raise the possibility that domestication has significant effects on molecular evolution. If true, this would demonstrate that rates and patterns of molecular evolution are labile on relatively short timescales. It is widely assumed that all domesticated lineages were established less than 15,000 years ago, so any detectable effects of domestication on molecular evolution must be due to recent changes having a significant and measurable impact on molecular evolution. Domesticated lineages might therefore provide an interesting case study for the influence of population changes or

© The Author(s) 2014. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from http://gbe.oxfordjournals.org/ by guest on January 25, 2014

Genome Biology and Evolution

SMBE

alteration of selective regime on patterns and rates of molecular evolution. On a practical level, observation of widespread impacts of domestication on molecular evolution would suggest that caution must be exercised when estimating the date of origin of domesticated lineages from molecular data or when including sequences from domesticated lineages in dating analyses.

Broadly speaking, there are three ways that domestication could affect patterns of molecular evolution: artificial selection, relaxed selective constraints, and reduced effective population size in domesticated lineages. Direct or indirect selection for traits during domestication may increase the rate of nonsynonymous substitutions at specific loci associated with selected traits (e.g., coat color in pigs) (Fang et al. 2009). Similar effects may be detected in loci that are linked to sites under artificial selection, as selective sweeps can drive fixation of neutral or nearly neutral linked alleles (Innan and Kim 2004; Kim and Nielsen 2004; Rubin et al. 2010). Artificial selection could also have genome-wide impacts on the rates and patterns of molecular evolution if selection for novelty promotes the evolution of mechanisms that increase the production of variation. For example, Burt and Bell (1987) found that domesticated mammals have higher chiasmata frequencies than other mammals with similar ages of maturity, which they suggested reflects "adaptation to an environment characterized by intense selection in small populations for novel combinations of traits." Otto and Barton (2001) also found several examples across different kingdoms that suggest a link between artificial selection regimes and increased recombination. Strong directional selection pressure and/or reduced effective population size could potentially increase the mutation rate (Sniegowski et al. 1997; Lynch 2010, 2011), though any increase in production of novel traits comes at the cost of a higher rate of deleterious mutations (King and Kashi 2007). While mitochondrial genomes of mammals and birds rarely if ever recombine, if domestication does indirectly select for generation of variation through recombination or mutation (Burt and Bell 1987; Denamur and Matic 2006; Dobney and Larson 2006; Bromham 2009), it could potentially influence rates of molecular evolution.

Relaxed selection could influence molecular evolution in domesticated lineages by permitting a greater proportion of nonsynonymous mutations to persist. Some of the traits that experience relaxed selection during domestication may be related to changes in environmental conditions and lifestyle (Clutton-Brock 1999; Björnerfeldt et al. 2006; Driscoll et al. 2009; Rubin et al. 2010). For example, the higher proportion of nonsynonymous changes in the mitochondrial genomes of dogs (Björnerfeldt et al. 2006) and domestic yaks (Wang et al. 2011) has been attributed to relaxed selection on metabolic efficiency in domesticated lineages, due to humans changing their habitat, selecting for tameness, and providing protection from predators.

Domesticated populations may often experience reductions in effective population sizes due to inbreeding and genetic bottlenecks (Vilà et al. 2005; Xia et al. 2009). Reduced effective population size increases the chance of fixing slightly deleterious mutations through drift, which should be reflected in increased d_N/d_S (Kimura and Ohta 1971; Ohta 1992). This effect is thought to account for patterns such as the correlation between body size and d_N/d_S in mammals (Nikolaev et al. 2007; Popadin et al. 2007; Nabholz et al. 2013). Domesticated lineages may undergo extreme bottlenecks on foundation. However, the domestication process has likely occurred over long periods of time and may have included few or many bottlenecks interspersed with introgression and population expansion (Allaby et al. 2008; Meyer and Purugganan 2013). This process could allow a lineage to recover from dramatic bottlenecks (Vilà et al. 2005). For example, although Taurine cattle may have originally descended from less than one hundred female founders (Bollongino et al. 2012), the high level of current genetic diversity has led to estimates of an ancestral wild population of 90,000 (MacEachern et al. 2009). Ongoing selective breeding and narrowing of the breeding pool may have also reduced effective population size in some domesticated lineages (Medugorac et al. 2009). For example, dogs are likely to have experienced a prehistoric bottleneck from wolves (Vilà et al. 1997), but it is likely that some dog populations have experienced more severe bottlenecks in recent history from breeding pressure (Wayne and Ostrander 2007).

Changes in population structure or conditions during domestication may be expected to have significant impacts on molecular evolution. However, the generality of the relationship between domestication and patterns of molecular evolution has not been established. Is it confined to a few well-studied examples, or is it a more general feature of all domesticated lineages? Not all studies support higher nonsynonymous rates in domestic lineages. For example, Rokas (2009) found a lower d_N/d_S in the proteome of a domesticated fungus compared with its wild relative. Here, we aim to ask whether increased d_N/d_S is a general feature of the mitochondrial genomes of domesticated lineages by comparing sequences from the maximum available number of phylogenetically independent comparisons of domesticated mammals and birds and their wild relatives.

We focus on the mitochondrial genome for several reasons. The animal mitochondrial genome has a higher rate of molecular evolution than the nuclear genome (Rand 1994; Ballard and Whitlock 2004), so is more likely to reflect any recent changes in rates and patterns of molecular evolution than the nuclear genome. The mitochondrial genome also has a smaller effective population size than the nuclear genome because it is haploid, rarely if ever recombines, and is maternally inherited (Harrison 1989; Moore 1995; Rokas et al. 2003; Ballard and Whitlock 2004), so it is expected to have a higher rate of fixation of nearly neutral substitutions (Ohta 1992),

Downloaded from http://gbe.oxfordjournals.org/ by guest on January 25, 2014

Genome Biology and Evolution SMBE

which are thought to dominate mitochondrial genome evolution (Rand and Kann 1996; Bazin et al. 2006). As our aim is to include as many independent domestic lineages as possible, there is a much wider availability of mitochondrial genomes than whole nuclear genome sequences.

To test whether domesticated animals have significantly different patterns of molecular evolution in mitochondrial genomes, we compared complete or nearly complete mitochondrial genome sequences between 16 phylogenetically independent comparisons of domesticated mammals and birds and their close wild relatives. We took two complementary approaches to analyze the data. We used a sister pairs approach to compare branch length, synonymous and nonsynonymous differences, and their ratios in wild and domesticated lineages. We also analyzed all taxa together in a single phylogenetic ("whole tree") analysis. We found no evidence of a consistent difference between rates and patterns of molecular evolution in the mitochondrial genomes of domesticated mammals and birds and their wild relatives.

Materials and Methods

Selection of Comparisons

We defined domesticated lineages as genetically distinct populations of organisms that have been purposely bred to suit the needs of the domesticator (Blumler et al. 1991; Diamond 2002). We identified the wild relatives of each domesticate from the literature and collected information on the age and history of each domestication event (see supplementary material, Supplementary Material online). We verified using published sources that the chosen wild relative and domestic populations could be identified as well-supported, independent lineages from genetic data and that the domesticated and wild taxa were considered distinct based on morphology, behavior, or geography.

To maintain phylogenetic independence among comparisons of domesticates and their wild relatives, we did not include multiple domesticated lineages that share the same wild relatives. For example, the llama and alpaca are suspected of sharing a wild relative (Kadwell et al. 2001; Cui et al. 2007), so we could only use one of these domesticates in our study. However, we were able to obtain whole mitochondrial genomes associated with two independently domesticated lineages for the dog (Björnerfeldt et al. 2006) and the pig (Wu et al. 2007). For both the dog and pig, the two domesticate–wild comparisons were analyzed as quartets, where one domesticate–wild relative pair acted as the outgroup for the other comparison.

DNA Sequences

We found 16 comparisons of domesticates and their wild relatives with complete or nearly complete mitochondrial genome sequences available on GenBank (www.ncbi.nlm.nih.gov/genbank/, last accessed May 2013). For each comparison, we collected a complete or nearly complete mitochondrial genome sequence for the domesticate, its wild relative, and a closely related outgroup (see supplementary table S3, Supplementary Material online, for accession numbers and alignment lengths). We preferentially collected sequences for the most closely related wild relative and outgroup for each domesticated lineage for which we could obtain a complete or nearly complete mitochondrial genome sequence. We preferentially selected sequences from published articles that explicitly stated whether the sequences came from wild or domesticated individuals. Sequences were not always available for the closest known wild relative, so in some cases we had to choose a more distant wild taxon. We conducted analyses with and without these more distant comparisons (for details see supplementary material, Supplementary Material online). Similarly, in some cases, there is evidence in the literature for population bottlenecks in the wild relatives, and this parallel change may make it harder to detect any effect of reduction in effective population size in the domesticated lineages. We repeated the sister pair and whole tree analyses excluding these comparisons to account for these potentially problematic comparisons.

We used a single mitochondrial genome to represent each taxon. This is because we wished to maximize the number of independent comparisons included in order to gauge general patterns of mitochondrial evolution in domesticated lineages. Multiple sequences are available for relatively few appropriate comparisons, and in many cases the lineages are not clearly monophyletic, which complicates the comparison of rates of substitution or levels of polymorphism (Hughes 2013). Use of a single sequence also avoids the problem of node density effect (Hugall and Lee 2007), especially because the level of polymorphism or number of substitutions may be overestimated in domesticated lineages if a greater number of sequences from domesticated lineages are included than sequences from the wild relatives. By using only a single sequence per lineage, we are unable to distinguish between substitutions (present in all members of a population) and polymorphisms (present in some but not all members of a population).

Sister Pairs Analysis

We aligned the mitochondrial sequences (including protein-coding genes, rRNA, tRNA, and control region sequences) for each domesticate–wild relative comparison and outgroup. We also constructed alignments of only protein-coding genes for estimating nonsynonymous (d_N) and synonymous (d_S) substitution rates. All alignments were performed by eye in Geneious (Drummond et al. 2011). We deleted any sites or codons that contained gaps in either the domesticate or wild relative sequence so that each base was comparable between sister species and thus informative for a sister pairs analysis.

Downloaded from http://gbe.oxfordjournals.org/ by guest on January 25, 2014

Genome Biology and Evolution

SMBE

For the whole genome alignments for each comparison, we estimated branch lengths in BASEML (Yang 2007) using the TN93 substitution model and unconstrained rates (clock = 0 in PAML). We estimated d_S, d_N, and d_N/d_S for the protein-coding sequences in CODEML in PAML (version 4.4b, Yang 2007), using the F3x4 codon frequency model (clock = 0). We tested for significant differences in branch length for each comparison using a likelihood ratio test (LRT).

We combined all 16 independent comparisons into a single analysis in order to ask whether the domesticated lineages have consistently different patterns of molecular evolution than their wild relatives. Each independent comparison contributed one data point to a nonparametric analysis of the differences in branch length, d_N, d_S, and d_N/d_S between domesticates and their wild relatives. We used both a sign test and the Wilcoxon signed-ranks test (Wilcoxon 1946).

As older divergences have had more time to accumulate substitutions, it may be that the power to detect a significant difference increases over time. If this were the case, we expect that if we compare age of domestication (years before present) or divergence of each sister pair (sum of domesticate and wild relative branch lengths) with the difference between domesticate and wild relative in d_N/d_S, d_N, d_S, and total substitution rate, we would find that the older or more divergent comparisons are more likely to show a positive association between domestication and molecular evolution. To test this prediction, we used Spearman's rank correlation to test for an association between mean age of domestication (measured in years before present, table 1) and differences in branch lengths, d_S, d_N, and d_N/d_S between domesticates and wild relatives. We also used Spearman's rank correlation to test for an association between the genetic distance between domesticated and wild lineages (measured as the sum of both the domestic and wild branches in each comparison) and differences in branch length, d_S, d_N, and d_N/d_S.

Whole Tree Analysis

In addition to the sister pairs approach, we performed a whole tree analysis where we combined the domesticated and wild taxa together into a single phylogeny. Because not all sequences could be confidently aligned between birds and mammals, we created three different alignments: 1) all sequences for all bird taxa; 2) all sequences for all mammal taxa; and 3) protein-coding sequences for all birds and mammals. The D-loop region was excluded from the whole tree analysis because it could not be confidently aligned across all taxa and was not available for several of the domesticate–wild relative comparisons.

For each of these three alignments, we estimated a phylogeny using the following procedure. First, we established data partitions for each alignments using a greedy search in PartitionFinder v1.0.1 (Lanfear et al. 2012), with linked branch lengths, constraining the models of evolution to those available in RAxML, and using AICc for model selection (a measure of AIC corrected for small sample sizes, Hurvich and Tsai 1989). In PartitionFinder, we defined initial data blocks that separated protein-coding genes by gene and codon position. For alignments 1 (all bird genes) and 2 (all mammal genes), we treated the 12S and 16S rRNA genes as separate data blocks and combined all tRNA sequences into one data block. Then, using the best partitions identified with PartionFinder, we analyzed the three alignments in RAxML version 7.0.4 (Stamatakis et al. 2008) to estimate a maximum likelihood phylogeny for each alignment, with 1,000 bootstrap replicates generated using the rapid bootstrapping algorithm. For the phylogenies based on alignments 1 and 2, we estimated branch lengths in BASEML (Yang 2007) using the REV model, unpartitioned data, and no molecular clock (clock = 0). For the phylogeny based on the protein-coding genes for birds and mammals, we used CODEML (Yang 2007) to estimate d_N/d_S in domesticated and wild lineages using the F3x4 codon frequency model, unpartitioned data, and no molecular clock (clock = 0).

For all phylogenies, we then tested for a significant difference in branch length between domesticated lineages and nondomesticated lineages using a LRT, comparing a one-rate model, where all taxa have the same rate, and a two-rate model, where one rate was estimated for all domesticates and a second rate for all wild relatives. A significant result from the LRT would allow us to reject the hypothesis of uniform rates over the phylogeny.

All alignment and data files used in this analysis are available on the Dryad Digital Repository http://doi.org/10.5061/dryad.d85ng) and can also be obtained from the corresponding author.

Results

Sister Pairs Analysis

We analyzed differences in branch length, synonymous (d_S) and nonsynonymous (d_N) differences, and d_N/d_S for 16 sister pairs between domesticated birds and mammals and their wild relatives using a sign test and Wilcoxon signed-rank test (table 1). We found no evidence for a consistent difference between domesticated and wild lineages in branch length (sign test $P = 0.80$, Wilcoxon signed-ranks $P = 0.32$), synonymous rates (d_S: sign test $P = 0.45$, Wilcoxon signed-ranks $P = 0.78$), nonsynonymous rates (d_N: sign test $P = 1.00$, Wilcoxon signed-ranks $P = 1.00$), nor d_N/d_S (sign test $P = 1.00$, Wilcoxon signed-ranks $P = 0.75$).

Six out of 16 comparisons showed a significant difference in branch length between the domesticated and wild lineages (presented in bold in table 1). In three of these comparisons (llama and both pig lineages), the domesticated lineages had a significantly longer branch length. In the remaining three

Downloaded from http://gbe.oxfordjournals.org/ by guest on January 25, 2014

Table 1

Comparison of Mitochondrial Genomes from 16 Domesticated Mammals and Birds and Their Wild Relatives

Domesticate (Common Name)	Domesticate (Scientific Name)	Wild Relative (Scientific Name)	Mean Domestication Age	d_N/d_S			d_S			d_N			Branch Length		
				D	W	Sign	D	W	Sign	D	W	Sign	D	W	Sign
Dog	*Canis lupus familiaris*[a]	*Canis lupus*[a]	13,500	0.1673	0.1006	+	0.0083	0.0082	+	0.0014	0.0008	+	0.0045	0.0047	–
Dog	*Canis lupus familiaris*[a]	*Canis lupus*[a]	13,500	0	0	–	0.0038	0.0019	+	0	0	–	0.0016	0.0010	+
Cat	*Felis catus*	*Lynx rufus*	9,750	0.0330	0.0278	+	0.2811	0.3106	–	0.0093	0.0086	+	0.0814	0.0905	–
Horse	*Equus caballus*	*Equus przewalskii*	6,000	0.0441	0.3065	–	0.0057	0.0030	+	0.0002	0.0009	–	0.0021	0.0013	+
Donkey	*Equus asinus*	*Equus asinus somalicus*	5,000	0.0465	0.0648	–	0.0325	0.0234	+	0.0015	0.0015	–	0.0098	0.0065	+
Pig	***Sus scrofa domesticus***[a]	***Sus scrofa***[a]	9,000	0.1728	0.1932	–	0.0136	0.0096	+	0.0024	0.0019	+	**0.0056**	**0.0037**	**+**
Pig	***Sus scrofa domesticus***[a]	***Sus scrofa***[a]	9,000	0.0575	0.0497	+	0.0129	0.0074	+	0.0007	0.0004	+	**0.0033**	**0.0018**	**+**
Goat	*Capra hircus*	*Capra falconeri*	11,000	0.0662	0.0601	+	0.0474	0.0476	–	0.0031	0.0029	+	0.0141	0.0176	–
Sheep	***Ovis aries***	***Ovis ammon***	8,500	0.0562	0.0391	+	0.0369	0.0596	–	0.0021	0.0023	–	**0.0135**	**0.0155**	–
Cow	***Bos taurus***	***Bos gaurus***	9,000	0.0360	0.0400	–	0.1566	0.1486	+	0.0056	0.0059	–	**0.0405**	**0.0426**	–
Yak	*Bos grunniens*	*Bos grunniens*	4,750	0.0870	0.0634	+	0.0090	0.0123	–	0.0008	0.0008	+	0.0034	0.0038	–
Water buffalo	*Bubalus bubalis*	*Bubalus depressicornis*	5,000	0.0476	0.1009	–	0.0447	0.0429	+	0.0021	0.0043	–	0.0099	0.0162	–
Llama	***Lama glama***	***Lama guanicoe***	6,000	0.0431	0.0171	+	0.0378	0.0313	+	0.0016	0.0005	+	**0.0113**	**0.0080**	**+**
Bactrian camel	*Camelus bactrianus*	*Camelus ferus*	5,000	0.0591	0.0730	–	0.0305	0.0329	–	0.0018	0.0024	–	0.0089	0.0102	–
Chicken	*Gallus gallus*	*Gallus gallus gallus*	8,250	0.1083	0.1098	–	0.0033[b]	0.0033[b]	+	0.0004[b]	0.0004[b]	–	0.0011	0.0010	+
Goose	***Anser anser***	***Anser albifrons***	3,500	0.0511	0.0302	+	0.0310	0.0443	–	0.0016	0.0013	+	**0.0071**	**0.0122**	–
		No. positive comparisons				8			10			8			7
		Total no. comparisons				16			16			16			16
		Sign test *P*				1			0.45			1			0.80
		Wilcoxon signed-rank test *P*				0.75			0.78			1			0.32

NOTE.—The age of domestication was calculated from the mean of published estimated ranges of timing of domestication events in years before present (see supplemental material, Supplemental Material online, for sources and details on the domesticate–wild comparisons chosen). Estimates of synonymous (d_S) and nonsynonymous (d_N) substitutions rates, d_N/d_S, and total substitution rate (substitutions per site) were estimated in PAML v 4.4b (Yang 2007). Comparisons with significantly different branch lengths between domesticated and wild lineages (see Materials and Methods) are presented in bold. Columns marked with D represent estimates for domesticates, and columns labeled W represent estimates for wild relatives. The sign columns represent the sign of the difference between domesticate and wild relative values. Positive symbols represent values that are larger in domesticates compared with their wild relatives, and negative symbols represent smaller values in domesticates.

[a]Domesticate–wild relative comparisons represented by two independent lineages (the dog and the pig).

[b]While these values are nearly equal, there is a small difference reflected in the direction of the sign of the d_S (chicken 0.003297, red junglefowl 0.003293) and d_N (chicken 0.000357, red junglefowl 0.000362). When the chicken comparison is removed from the data set, the sign and Wilcoxon signed-ranks tests are still not significant.

Downloaded from http://gbe.oxfordjournals.org/ by guest on January 25, 2014

Genome Biology and Evolution

SMBE

comparisons (sheep, cow, and goose), the wild relative had a significantly longer branch length.

A Spearman's rank correlation test revealed no evidence of a correlation between the age of the domestication event and direction of the difference between domesticated and wild lineages in branch length ($\rho = 0.01$, $P = 0.97$), d_S ($\rho = 0.08$, $P = 0.76$), d_N ($\rho = 0.40$, $P = 0.13$), nor in d_N/d_S ($\rho = 0.14$, $P = 0.62$). We also found no evidence of a correlation between domestication age and genetic distance between sister pairs ($\rho = 0.01$, $P = 0.96$), suggesting that, in the mitochondrial genome, the older comparisons included in this study do not always have the greatest genetic distance.

We found no significant relationship between genetic distance (sum of wild and domesticate branch lengths) and difference in d_S ($\rho = -0.26$, $P = 0.34$), d_N ($\rho = 0.07$, $P = 0.80$), or d_N/d_S ($\rho = 0.07$, $P = 0.79$), but we did find a significant negative relationship between genetic distance and difference in branch length ($\rho = -0.65$, $P = 0.01$). This suggests that in the most divergent comparisons, the wild relative is more likely to have the longer branch length. The relationship is robust to the removal of either the cat or the cow comparisons, which are the most divergent comparisons (supplementary fig. S1, Supplementary Material online); however, removing both of these comparisons makes the relationship nonsignificant. This result suggests that the net amount of molecular change between the sequences could influence the chance of detecting a difference in rate between the domesticated and wild relatives, but that this effect is unlikely to be responsible for our failure to detect more genetic change in domesticated lineages, as the relationship is in the opposite direction (greater genetic distance is associated with longer branches in the wild relative).

Whole Tree Analysis

For the whole tree analysis, we found no significant difference between the one- and two-rate models for any of the three alignments we tested: 1) no significant difference in d_N/d_S for the alignment of protein-coding genes for all birds and mammals ($P = 0.42$); 2) no significant difference in branch length for whole genome alignment for all birds ($P = 0.98$); 3) no significant difference in branch length for whole genome alignment for all mammals ($P = 0.95$).

We repeated the sister pair and whole tree analyses, removing comparisons for which we were unable to use the closest wild relatives (cat, goat, cow, water buffalo, and goose; either because of sequence availability or because the closest relative is extinct) or where we found evidence in the literature that the wild relatives have experienced genetic bottlenecks (camel, pigs, horse, donkey, and water buffalo, see supplementary material and supplementary tables S1 and S2, Supplementary Material online). Only a small part of the divergence in the distant comparisons (cat, goat, cow, water buffalo, and goose) may actually correspond to molecular changes influenced by domestication, which could make these comparisons less informative. Furthermore, if reduced effective population size influences molecular rates in domesticates, we may have had difficulty detecting that signal when comparing a domesticate with a wild relative that has also experienced reduced effective population size. We repeated the analyses removing comparisons with suspected bottlenecks in the wild relatives: camel, pigs, horse, donkey, and the water buffalo. We repeated this analysis with and without the water buffalo as the wild relative, the lowland anoa, has only recently experienced a genetic bottleneck (see supplementary material, Supplementary Material online). In addition to experiencing a recent genetic bottleneck, the lowland anoa is an island endemic, which could be associated with a reduced effective population size and, thus, increased molecular rates (Woolfit and Bromham 2005).

When repeating the sister pairs analysis without comparisons where the wild relative has experienced bottlenecks (camel, pigs, horse, donkey, and the water buffalo), we found that domesticates have a significantly higher d_N/d_S than their wild relatives (Wilcoxon signed-ranks $P = 0.02$). Therefore, it is possible that in the pairs with bottlenecks in the wild relatives, reduced effective population size has had parallel effects in both domesticated lineages and their wild relatives, reducing the chance of detecting differences between them. All other alternative sister pair and whole tree analyses were not significant (supplementary tables S1 and S2, Supplementary Material online).

Discussion

We find no evidence for a general and consistent difference in the tempo and mode of mitochondrial molecular evolution of domesticated birds and mammals when compared with their wild relatives. Given that higher d_N/d_S has been reported for a number of domestic lineages, why do we fail to find evidence for a general increase in d_N/d_S across all the domestic lineages included in this study?

It is possible that lack of statistical power has prevented us from identifying significant differences in some comparisons. Our power is unavoidably limited by the nature of the question. We are unable to include more comparisons because there are relatively few fully domesticated animal lineages, and we had to leave some lineages out of this study due to lack of sequence data from appropriate wild relatives (e.g., turkey: see supplementary material, Supplementary Material online). It may be informative to apply this comparative approach to domesticated plants, which are more diverse. Furthermore, all domestication events are young on an evolutionary scale, so there has been only a short period of time for differences in tempo and mode of molecular evolution to make a detectable impression on patterns of sequence differences.

Downloaded from http://gbe.oxfordjournals.org/ by guest on January 25, 2014

Genome Biology and Evolution SMBE

If the relatively small number of sequence differences between recently diverged genomes was obscuring a result, then we would expect the six sister pair comparisons with a significant difference in branch length to be more likely to show longer branch lengths, or higher d_N/d_S, in the domesticated lineage. But only half of the comparisons with a significant difference in branch length show more genetic change in the domesticated lineage, a pattern indistinguishable from chance. We also find that in the more divergent comparisons (those with a greater net genetic distance between the domestic and wild lineages), it is the wild relative that is more likely to have a longer branch length. Older domesticated lineages, which have had more time to accumulate evidence of distinct patterns of molecular evolution, do not show a greater tendency to have higher rates of change than their wild relatives. So we do not think that lack of power to detect differences in rate of change explains the lack of a consistent pattern in our comparisons. However, it may be possible that bottlenecks in wild relative populations may impact our power to detect a difference in d_N/d_S between domesticated lineages and their wild relatives.

One way to increase power to detect changes in the tempo and mode of molecular evolution in domesticated lineages is to take a population-level approach, with multiple individual samples for each domesticate and wild lineage. Recent population-level studies have found increased d_N/d_S or ratio of nonsynonymous to synonymous diversity (π_N/π_S) in a number of domesticated lineages compared with their wild relatives (Wang et al. 2011; Hughes 2013). However, these studies have included an uneven number of domesticated and wild samples (254 from the dog vs. 19 from the wolf; 59 from the domestic pig vs. 27 from wild boar; 41 from the domestic chicken vs. 17 from the red junglefowl in Hughes 2013, and 51 domestic yaks vs. 21 wild yaks in Wang et al. 2011). Many short, recently diverged branches can increase estimates of d_N/d_S (Rocha et al. 2006), so higher d_N/d_S is more likely to be reported if an analysis includes more branches in a domesticate population than a wild one.

To avoid the measurement bias due to the node density effect, we only sampled one individual per domesticated and wild lineage. Choosing only one sequence per lineage also helps us to avoid the problem of lack of monophyly in analyses of population-level data. Backcrossing and interbreeding with wild relatives can shape the molecular evolution of domesticated and wild lineages (Vilà et al. 2005), and these processes may have varied substantially between lineages. For example, Hughes (2013) reported that the phylogenies of domesticated and wild lineages of chickens, dogs, and pigs are not monophyletic but intermixed, which could be a signature of ancestral polymorphisms or interbreeding in these populations. We have attempted to minimize this effect on our results by choosing wild lineages that may not be the closest relative but have less chance of being influenced by recent introgression (see Materials and Methods and supplementary material, Supplementary Material online). However, by choosing only one sequence per lineage, we are unable to distinguish substitutions from polymorphisms. Our approach could mask higher rates of change in the domesticated lineage if wild lineages consistently retained comparatively more ancestral polymorphisms.

If the majority of substitutions in the mitochondrial genome are neutral or slightly deleterious, rather than under positive selection, then we would expect d_N/d_S estimates in the mitochondrial genome to be higher within species than between species (Hasegawa et al. 1998; Rand and Kann 1998; Weinreich and Rand 2000; Ho et al. 2005). Therefore, population-level estimates of mitochondrial d_N/d_S that do not account for the effect of ancestral polymorphism are expected to be higher than those estimated at the lineage level. As such, we would expect our d_N/d_S estimates to be lower than those from population-level studies. Concordant with these population-level studies, we found a higher d_N/d_S in one dog, one pig, and the yak comparison. Although we cannot compare our d_N/d_S estimates with the π_N/π_S reported in Hughes (2013), our d_N/d_S estimate for the domesticated and wild yaks are, as expected, lower than those reported by Wang et al. (2011) (our d_N/d_S for wild yaks: 0.06, their d_N/d_S for wild yaks: 0.07, our d_N/d_S for domesticated yaks: 0.09, their d_N/d_S for domesticated yaks 0.23).

It could be argued that the housekeeping genes of the mitochondria are unlikely to experience a dramatic change in selective regime, which could explain why we found no consistent pattern associated with domestication in the mitochondrial genome. Actually, many studies of domestication report changes in traits associated with metabolism (Xia et al. 2009; Gibbons et al. 2012). For example, selective sweeps in chickens raised for meat production are connected to genes associated with growth, appetite, and metabolic regulation (Rubin et al. 2010). It is therefore possible that artificially selected traits could be associated with growth and metabolism, which could potentially increase d_N in mitochondrial loci (MacEachern et al. 2009; Rubin et al. 2010; Akey et al. 2010; Amaral et al. 2011; Kijas et al. 2012). However, our study is designed to detect changes in genome-wide rates of change, rather than focusing on the effect of selection on particular genes.

Our results do not preclude an impact of domestication on patterns of mitochondrial evolution, but they do suggest that there is no consistent, detectable difference between all domesticated lineages and their wild relatives. It may be that domestication can influence mitochondrial molecular evolution, but that it does not do so consistently and uniformly across all domesticated lineages in comparison to their wild relatives. Each domestication history has involved different levels of human intervention, and the observed genetic and morphological changes in domesticated lineages are variable (Zeder 2006). For example, it has been suggested that domestic sheep and cats may have undergone less severe genetic

Downloaded from http://gbe.oxfordjournals.org/ by guest on January 25, 2014

Genome Biology and Evolution

SMBE

bottlenecks than other domesticated animals (Driscoll et al. 2007; Kijas et al. 2009), but as both of these lineages have higher d_N/d_S estimates (table 1), this does not seem to provide an explanation for the lack of a general pattern of higher d_N/d_S across domesticated lineages. Similarly, some domesticated lineages, like the horse, cat, and camel, may have experienced less artificial selection than others (Clutton-Brock 1999; Driscoll et al. 2009), yet the horse and cat have higher d_N/d_S than their wild relatives, and the camel has lower d_N/d_S.

In addition to considering the heterogeneity of processes affecting the domesticated lineages, population processes in the wild relatives may also impact on our ability to detect changes in the tempo and mode of molecular evolution in domesticated lineages. If similar changes have occurred in both the domesticated lineages and their wild relatives, then we may be unable to detect a significant difference between them. In particular, some wild relatives may have experienced significant genetic bottlenecks. For example, the wild relative of the water buffalo, the lowland anoa, is an island endemic, which could be associated with a reduced effective population size and, thus, increased d_N/d_S (Woolfit and Bromham 2005). Other examples of wild relatives that may have undergone population size reduction are the wild Bactrian camels (Hare 1997; Silbermayr et al. 2010), wild boar (Scandura et al. 2008), Przewalski's horses (Clutton-Brock 1999; Vilà 2001), and the Somali wild ass (Moehlman 2002). When we analyzed a reduced set of comparisons, removing comparisons where we found evidence that the wild relative had undergone a population bottleneck, we found that domesticated lineages had a higher d_N/d_S than their wild relatives. Although the sample size for this test is small ($N = 10$), this result is consistent with the hypothesis that domestication reduces a lineage's effective population size and thus may increase the accumulation of slightly deleterious, nonsynonymous changes in the mitochondrial genome.

In this analysis of 16 domesticated mammals and birds, we find no evidence of a general, consistent pattern in the rates or patterns of molecular evolution in the mitochondria. However, we do find that in a subset of comparisons, there is evidence of higher d_N/d_S in domesticated lineages, which may be a signature of changes in effective population size. We conclude that differences in d_N/d_S between particular domesticated lineages and their wild relatives in the mitochondrial genome (Björnerfeldt et al. 2006; Wang et al. 2011) are best explained by specific factors in the biology or domestication history of particular lineages and not a generally predictable result of domestication.

Supplementary Material

Supplementary material, figure S1, and tables S1 and S2 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgment

This work was funded by the Australian Research Council.

Literature Cited

Akey JM, et al. 2010. Tracking footprints of artificial selection in the dog genome. Proc Natl Acad Sci U S A. 107:1160–1165.

Allaby RG, Fuller DQ, Brown TA. 2008. The genetic expectations of a protracted model for the origins of domesticated crops. Proc Natl Acad Sci U S A. 105:13982–13986.

Amaral AJ, et al. 2011. Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA. PLoS One 6:e14782.

Ballard JWO, Whitlock MC. 2004. The incomplete natural history of mitochondria. Mol Ecol. 13:729–744.

Bazin E, Glémin S, Galtier N. 2006. Population size does not influence mitochondrial genetic diversity in animals. Science 312:570–572.

Björnerfeldt S, Webster MT, Vilà C. 2006. Relaxation of selective constraint on dog mitochondrial DNA following domestication. Genome Res. 16:990–994.

Blumler MA, et al. 1991. The ecological genetics of domestication and the origins of agriculture. Curr Anthropol. 32:23–54.

Bollongino R, et al. 2012. Modern Taurine cattle descended from small number of Near-Eastern founders. Mol Biol Evol. 29:2101–2104.

Bromham L. 2009. Why do species vary in their rate of molecular evolution? Biol Lett. 5:401–404.

Burt A, Bell G. 1987. Mammalian chiasma frequencies as a test of two theories of recombination. Nature 326:803–805.

Clutton-Brock J. 1999. A natural history of domesticated mammals, 2nd ed. Cambridge (UK): Cambridge University Press.

Cruz F, Vilà C, Webster MT. 2008. The legacy of domestication: accumulation of deleterious mutations in the dog genome. Mol Biol Evol. 25:2331–2336.

Cui P, et al. 2007. A complete mitochondrial genome sequence of the wild two-humped camel (*Camelus bactrianus ferus*): an evolutionary history of Camelidae. BMC Genomics 8:241.

Denamur E, Matic I. 2006. Evolution of mutation rates in bacteria. Mol Microbiol. 60:820–827.

Diamond J. 2002. Evolution, consequences and future of plant and animal domestication. Nature 418:700–707.

Dobney K, Larson G. 2006. Genetics and animal domestication: new windows on an elusive process. J Zool. 269:261–271.

Driscoll CA, Macdonald DW, O'Brien SJ. 2009. From wild animals to domestic pets, an evolutionary view of domestication. Proc Natl Acad Sci U S A. 106:9971–9978.

Driscoll CA, et al. 2007. The Near Eastern origin of cat domestication. Science 317:519–523.

Drummond A, et al. 2011. Geneious v5.4. [cited 2013 Jun 1]. Available from: http://www.geneious.com/.

Fang M, Larson G, Soares Ribeiro H, Li N, Andersson L. 2009. Contrasting mode of evolution at a coat color locus in wild and domestic pigs. PLoS Genet. 5:e1000341.

Gibbons JG, et al. 2012. The evolutionary imprint of domestication on genome variation and function of the filamentous fungus *Aspergillus oryzae*. Curr Biol. 22:1403–1409.

Gu Z, et al. 2005. Elevated evolutionary rates in the laboratory strain of *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A. 102:1092–1097.

Hare J. 1997. The wild Bactrian camel *Camelus bactrianus ferus* in China: the need for urgent action. Oryx 31:45–48.

Harrison RG. 1989. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends Ecol Evol. 4:6–11.

Hasegawa M, Cao Y, Yang Z. 1998. Preponderance of slightly deleterious polymorphism in mitochondrial DNA: nonsynonymous/synonymous

Downloaded from http://gbe.oxfordjournals.org/ by guest on January 25, 2014

rate ratio is much higher within species than between species. Mol Biol Evol. 15:1499–1505.
Ho SYW, Phillips MJ, Cooper A, Drummond AJ. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Mol Biol Evol. 22:1561–1568.
Hugall AF, Lee MSY. 2007. The likelihood node density effect and consequences for evolutionary studies of molecular rates. Evolution 61: 2293–2307.
Hughes AL. 2013. Accumulation of slightly deleterious mutations in the mitochondrial genome: a hallmark of animal domestication. Gene 515:28–33.
Hurvich CM, Tsai CL. 1989. Regression and time series model selection in small samples. Biometrika 76:297–307.
Innan H, Kim Y. 2004. Pattern of polymorphism after strong artificial selection in a domestication event. Proc Natl Acad Sci U S A. 101: 10667–10672.
Kadwell M, et al. 2001. Genetic analysis reveals the wild ancestors of the llama and the alpaca. Proc Biol Sci. 268:2575–2584.
Kijas JW, et al. 2009. A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. PLoS One 4:e4668.
Kijas JW, et al. 2012. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. PLoS Biol. 10:e1001258.
Kim Y, Nielsen R. 2004. Linkage disequilibrium as a signature of selective sweeps. Genetics 167:1513–1524.
Kimura M, Ohta T. 1971. Theoretical aspects of population genetics. Princeton (NJ): Princeton University Press.
King DG, Kashi Y. 2007. Indirect selection for mutability. Heredity 99: 123–124.
Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol. 29:1695–1701.
Lu J, et al. 2006. The accumulation of deleterious mutations in rice genomes: a hypothesis on the cost of domestication. Trends Genet. 22: 126–131.
Lynch M. 2010. Evolution of the mutation rate. Trends Genet. 26: 345–352.
Lynch M. 2011. The lower bound to the evolution of mutation rates. Genome Biol Evol. 3:1107–1118.
MacEachern S, et al. 2009. Molecular evolution of the Bovini tribe (Bovidae, Bovinae): is there evidence of rapid evolution or reduced selective constraint in domestic cattle? BMC Genomics 10:179.
Medugorac I, et al. 2009. Genetic diversity of European cattle breeds highlights the conservation value of traditional unselected breeds. Mol Ecol. 18:3394–3410.
Meyer RS, Purugganan MD. 2013. Evolution of crop species: genetics of domestication and diversification. Nat Rev Genet. 14:840–852.
Moehlman PD. 2002. Status and action plan for the African wild ass. In: Moehlman PD, editor. Equids: zebras, asses and horses, status survey and conservation action plan. Gland (Switzerland): IUCN. p. 2–9.
Moore WS. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution 49:718–726.
Nabholz B, Uwimana N, Lartillot N. 2013. Reconstructing the phylogenetic history of long-term effective population size and life-history traits using patterns of amino acid replacement in mitochondrial genomes of mammals and birds. Genome Biol Evol. 5:1273–1290.
Nikolaev SI, et al. 2007. Life-history traits drive the evolutionary rates of mammalian coding and noncoding genomic elements. Proc Natl Acad Sci U S A. 104:20443–20448.
Ohta T. 1992. The nearly neutral theory of molecular evolution. Annu Rev Ecol Syst. 23:263–286.
Otto SP, Barton NH. 2001. Selection for recombination in small populations. Evolution 55:1921–1931.
Popadin K, Polishchuk LV, Mamirova L, Knorre D, Gunbin K. 2007. Accumulation of slightly deleterious mutations in mitochondrial protein-coding genes of large versus small mammals. Proc Natl Acad Sci U S A. 104:13390–13395.
Rand DM. 1994. Thermal habit, metabolic rate and the evolution of mitochondrial DNA. Trends Ecol Evol. 9:125–131.
Rand DM, Kann LM. 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. Mol Biol Evol. 13:735–748.
Rand DM, Kann LM. 1998. Mutation and selection at silent and replacement sites in the evolution of animal mitochondrial DNA. Genetica 102–103:393–407.
Rocha EPC, et al. 2006. Comparisons of d_N/d_S are time dependent for closely related bacterial genomes. J Theor Biol. 239:226–235.
Rokas A. 2009. The effect of domestication on the fungal proteome. Trends Genet. 25:60–63.
Rokas A, Ladoukakis E, Zouros E. 2003. Animal mitochondrial DNA recombination revisited. Trends Ecol Evol. 18:411–417.
Rubin CJ, et al. 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. Nature 464:587–591.
Scandura M, et al. 2008. Ancient vs. recent processes as factors shaping the genetic variation of the European wild boar: are the effects of the last glaciation still detectable? Mol Ecol. 17: 1745–1762.
Silbermayr K, et al. 2010. High mitochondrial differentiation levels between wild and domestic Bactrian camels: a basis for rapid detection of maternal hybridization. Anim Genet. 41:315–318.
Sniegowski PD, Gerrish PJ, Lenski RE. 1997. Evolution of high mutation rates in experimental populations of *E. coli*. Nature 387: 703–705.
Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol. 57:758.
Vilà C. 2001. Widespread origins of domestic horse lineages. Science 291: 474–477.
Vilà C, Seddon J, Ellegren H. 2005. Genes of domestic mammals augmented by backcrossing with wild ancestors. Trends Genet. 21: 214–218.
Vilà C, et al. 1997. Multiple and ancient origins of the domestic dog. Science 276:1687–1689.
Wang Z, et al. 2011. Domestication relaxed selective constraints on the yak mitochondrial genome. Mol Biol Evol. 28:1553–1556.
Wayne RK, Ostrander EA. 2007. Lessons learned from the dog genome. Trends Genet. 23:557–567.
Weinreich DM, Rand DM. 2000. Contrasting patterns of nonneutral evolution in proteins encoded in nuclear and mitochondrial genomes. Genetics 156:385–399.
Wilcoxon F. 1946. Individual comparisons of grouped data by ranking methods. J Econ Entomol. 39:269.
Woolfit M, Bromham L. 2005. Population size and molecular evolution on islands. Proc R Soc Lond B Biol Sci. 272:2277–2282.
Wu GS, et al. 2007. Population phylogenomic analysis of mitochondrial DNA in wild boars and domestic pigs revealed multiple domestication events in East Asia. Genome Biol. 8:R245.
Xia Q, et al. 2009. Complete resequencing of 40 genomes reveals domestication events and genes in silkworm (*Bombyx*). Science 326: 433–436.
Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol Biol Evol. 24:1586–1591.
Zeder MA. 2006. Central questions in the domestication of plants and animals. Evol Anthropol. 15:105–117.

Associate editor: Michael Purugganan

Genome Biology and Evolution SMBE

Downloaded from http://gbe.oxfordjournals.org/ by guest on January 25, 2014