ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

The evolutionary history of grey wolf Y chromosomes

Linnéa Smeds[1] 🆔 | Ilpo Kojola[2] | Hans Ellegren[1] 🆔

[1]Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden

[2]Natural Resources Institute Finland (Luke), Rovaniemi, Finland

Correspondence
Hans Ellegren, Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden.
Email: Hans.Ellegren@ebc.uu.se

Abstract

Analyses of Y chromosome haplotypes uniquely provide a paternal picture of evolutionary histories and offer a very useful contrast to studies based on maternally inherited mitochondrial DNA (mtDNA). Here we used a bioinformatic approach based on comparison of male and female sequence coverage to identify 4.7 Mb from the grey wolf (*Canis lupis*) Y chromosome, probably representing most of the male-specific, nonampliconic sequence from the euchromatic part of the chromosome. We characterized this sequence and then identified ≈1,500 Y-linked single nucleotide polymorphisms in a sample of 145 resequenced male wolves, including 75 Finnish wolf genomes newly sequenced in this study, and in 24 dogs and eight other canids. We found 53 Y chromosome haplotypes, of which 26 were seen in grey wolves, that clustered in four major haplogroups. All four haplogroups were represented in samples of Finnish wolves, showing that haplogroup lineages were not partitioned on a continental scale. However, regional population structure was indicated because individual haplotypes were never shared between geographically distant areas, and genetically similar haplotypes were only found within the same geographical region. The deepest split between grey wolf haplogroups was estimated to have occurred 125,000 years ago, which is considerably older than recent estimates of the time of divergence of wolf populations. The distribution of dogs in a phylogenetic tree of Y chromosome haplotypes supports multiple domestication events, or wolf paternal introgression, starting 29,000 years ago. We also addressed the disputed origin of a recently founded population of Scandinavian wolves and observed that founding as well as most recent immigrant haplotypes were present in the neighbouring Finnish population, but not in sequenced wolves from elsewhere in the world, or in dogs.

KEYWORDS
bioinfomatics/phyloinfomatics, conservation genetics, haplotypes, population genomics, Y chromosome

1 | INTRODUCTION

The sex-limited chromosome—the sex chromosome unique to one sex (i.e., the Y chromosome in organisms with male heterogamety, such as mammals and *Drosophila*, and the W chromosome in organisms with female heterogamety, such as birds)—differs from other nuclear DNA in that it is uniparentally inherited and does not recombine. This has several consequences that make it unusual with respect to function, evolution and use in population genetic analyses (Bachtrog, 2013; Ellegren, 2011; Jobling & Tyler-Smith, 2003; Lahn,

Pearson, & Jegalian, 2001). The mammalian Y chromosome evolved from an ancestral pair of autosomes in which arrest of recombination leading to independent X and Y chromosome lineages was initiated 180 (Cortez et al., 2014) to 240–320 million years ago (Lahn & Page, 1999; Ross et al., 2005; Sandstedt & Tucker, 2004). Several subsequent steps of recombination cessation in different lineages further delimited the Y chromosome (Lahn & Page, 1999; Ross et al., 2005; Sandstedt & Tucker, 2004), a process potentially driven by sexually antagonistic alleles and mediated by Y chromosome inversions (Charlesworth, 2017; Charlesworth, Charlesworth, & Marais, 2005; Ponnikas, Sigeman, Abbott, & Hansson, 2018). In addition, a combination of different forms of rearrangements, such as transposition, retrotransposition, deletion and duplication, has meant that the Y chromosome today bears little resemblance to its ancestral homologue, the X chromosome.

In the absence of recombination, Y chromosomes become highly degenerate over time (Charlesworth, 1978; Charlesworth & Charlesworth, 2000; Nei, 1970; Rice, 1994) and accumulate a rich repertoire of repetitive sequences, including large heterochromatic structures. Degeneration and rearrangements imply that Y chromosome evolution is highly dynamic; indeed, the structure and gene content of the Y chromosome differ significantly among mammalian lineages (Chang, Yang, Retzel, & Liu, 2013; Cortez et al., 2014; Hughes et al., 2012,2010; Li et al., 2013; Skinner et al., 2016; Soh et al., 2014). In most lineages, only a few genes have survived Y chromosome degeneration (Lahn & Page, 1997; Skaletsky et al., 2003). Preservation of such genes could be driven by selection in males to maintain two copies of dosage-sensitive genes (Bellott et al., 2014; Cortez et al., 2014; Lahn & Page, 1997). Alternatively, the acquisition of male-specific function (e.g., in testis development and spermatogenesis) would set the stage for adaptive evolution and introduce strong selective constraint against degeneration of Y-linked genes (Lahn & Page, 1997; Skaletsky et al., 2003). An interesting feature of genes of the latter category noted in several mammals is that they often reside within campliconic structures harbouring multiple gene copies (Bhowmick, Satta, & Takahata, 2007; Rozen et al., 2003; Soh et al., 2014).

Because the Y chromosome does not recombine, except for the pseudoautosomal region, it is clonally inherited as a single haplotype from father to son. This means that male lineages can be traced back in time (Semino et al., 2000; Underhill et al., 2000), offering a very useful contrast to phylogenetic or population genetic patterns provided by analyses of maternally inherited mtDNA, and of autosomes (Poznik et al., 2013; Underhill & Kivisild, 2007). There are many examples of sex-specific demographic histories related to, for example, mating system, domestication, dispersal and introgression, and such differences between sexes can only be revealed by analyses of both paternally and maternally inherited genetic markers (Jones & Searle, 2015; Lippold et al., 2014).

Although potentially powerful, population genetic and phylogenetic inferences from Y chromosome haplotypes have only been made for a very limited number of organisms, notably humans (Karafet et al., 2008; Karmin et al., 2015; Poznik et al., 2016;

Underhill et al., 2000; for a recent review see Jobling & Tyler-Smith, 2017), but also in some domestic animals including dogs (Brown, Darwent, Wictum, & Sacks, 2015; Ding et al., 2012; Oetjens, Martin, Veeramah, & Kidd, 2018; Sacks et al., 2013). One reason for this is that many genome assemblies lack sequence data from the sex-limited chromosome because an individual of the homogametic sex was used for genome sequencing (to increase coverage of the X chromosome). When an individual of the heterogametic sex has been used, sequence data from the sex-limited chromosome may be limited due to low coverage (of a haploid chromosome), or to the very high repeat content that is characteristic of nonrecombining chromosomes. Moreover, because linkage maps cannot be produced for nonrecombining chromosomes, Y chromosome assembly cannot be assisted by ordering or orientation of scaffolds based on genetic data (Tomaszkiewicz, Medvedev, & Makova, 2017). Painstaking bacterial artificial clone (BAC) sequencing and single-haplotype iterative mapping were used for assembling the human (Skaletsky et al., 2003), chimpanzee (Hughes et al., 2010), rhesus monkey (Hughes et al., 2012), mouse (Soh et al., 2014), and dog and cat Y chromosome (Li et al., 2013), but is impractical for widespread use across organisms. However, recently developed bioinformatic approaches based on comparisons of male and/or female sequence coverage, or k-mers distribution, can identify scaffolds from the sex-limited chromosome in essentially any species (Carvalho & Clark, 2013; Chen, Bellott, Page, & Clark, 2012; Hall et al., 2013; Smeds et al., 2015; Tomaszkiewicz et al., 2016).

The grey wolf (*Canis lupis*) is an iconic carnivore species with a complex history of diverse types of relationships with humans. With the rise of agriculture, wolves started to pose a threat to human settlements when free -range livestock formed easy targets for predation (Leonard, Vilà, & Wayne, 2005). Fostered by symbolic evil roles in religion and, later, in stories and tales, strong antipathy toward wolves developed. Extermination of wolves escalated over time in both Europe and North America, and was in modern times supported by bounty programmes and facilitated by the use of poison and more efficient weapons. As a result of eradication campaigns, wolves have disappeared from many parts of the world where they once were common. In other parts they are considered endangered, with several small populations suffering from inbreeding and potentially inbreeding depression (Åkesson et al., 2016; Aspi, Roininen, Ruokonen, Kojola, & Vilà, 2006; Gómez-Sánchez et al., 2018; Kardos et al., 2018; Liberg et al., 2005; Pilot et al., 2013; Randi et al., 2000; Sastre et al., 2011; Vilà, Walker et al., 2003). In Scandinavia, wolves went extinct in the 1960s but, somewhat surprisingly, again started to breed in southern parts of the Scandinavian peninsula in the 1980s, >1,000 km from the nearest regular occurrence in Finland and Russia. A special relationship between wolves and humans owes to the fact that dogs were domesticated from wolves, a process that has been given considerable attention by geneticists (Freedman et al., 2014; Leonard et al., 2002; Pang et al., 2009; Savolainen, Zhang, Luo, Lundeberg, & Leitner, 2002; Skoglund, Ersmark, Palkopoulou, & Dalen, 2015; Thalmann et al., 2013; Vilà et al., 1997; vonHoldt et al., 2010).

Here we used a computational approach to identify and characterize 4.7 Mb of sequence from the wolf Y chromosome. Based on genomic resequencing data from 145 male wolves we identified Y chromosome haplotypes and studied their phylogenetic relationship. Finally, we use Y chromosome haplotypes to study the disputed origin of an endangered Scandinavian wolf population.

2 | MATERIAL AND METHODS

2.1 | Mapping of reads and the identification of Y-specific scaffolds based on coverage

Raw reads from 10 male and 10 female wolves (Kardos et al., 2018, for accession numbers, see Supporting Information Table S1) were mapped onto a male wolf genome assembly (Gopalakrishnan et al., 2017, downloaded from https://sid.erda.dk/wsgi-bin/ls.py?share_xml:id=f1ppDgUPQG), using BWA version 0.7.13 (Li & Durbin, 2009). The reads were sorted with SAMTOOLS version 1.5 (Li et al., 2009) and merged and deduplicated with PICARD version 2.10.3 (http://broadinstitute.github.io/picard/). Y chromosome-specific scaffolds were identified following the procedure of Smeds et al. (2015). In short, perfectly mapping reads (i.e., reads that mapped without any mismatches) were extracted, and the mean and median coverage per scaffold were calculated with BEDTOOLS version 2.25.0 (Quinlan & Hall, 2010) for males and females separately. Combining data for all 10 individuals of each sex gave a mean coverage of 257× for males (median 271×) and 238× (251×) for females; note that this will mostly reflect the coverage of autosomal sequence, which constitutes the majority of the genome. We then selected scaffolds with a median male coverage of >50× and a female-to-male median coverage ratio of 0 (meaning female median coverage = 0). These scaffolds were considered to represent sequences from the Y chromosome (see Supporting Information Figure S1 for examples). Relaxed settings did not add much sequence; for example, using >20× male coverage and a female-to-male coverage ratio of 0.05 only added 1% more data, <50 kb.

Five scaffolds that did not pass the strictly set thresholds described above were still considered to originate from the Y chromosome from the observation that they contained genes (or part of genes) known to be Y-linked in dogs. Importantly, three of these scaffolds (scaffold_5293, scaffold_5774 and scaffold_7290) had very high male coverage (>1,000×, suggesting they are part of repetitive structures in which several similar repeat copies have been collapsed in the assembly), but failed identification due to nonzero female coverage (1–2×). The latter could result from similar (repeat) sequences also present elsewhere in the genome. A fourth scaffold (scaffold_3306) had zero female coverage, but only 6× perfect male coverage. Many additional male reads mapped to this scaffold but with mismatches, indicating errors in the assembled scaffold. The fifth scaffold (scaffold_242), which also turned out to be the largest putatively Y-linked scaffold, consisted of an ~425-kb-long region that met the threshold for Y-linkage, followed by a 1.74-Mb-long segment with equal male and female coverage (Supporting Information Figure S1). The latter sequence aligned to the dog

pseudo-autosomal region (PAR) and this scaffold hence included the border between the PAR and the male-specific region on the Y chromosome (MSY) region. For further analysis, we only kept the 425-kb MSY region of this scaffold. The Y scaffolds were aligned to a previously published dog Y assembly (Li et al., 2013) using nucmer from the MUMMER package (version 3.9.4). The -maxmatch parameter was used to output all possible anchors between the two, not just unique anchors. Anchors longer than 500 bp with at least 99% identity were visualized with CIRCOS version 0.69-6 (Krzywinski et al., 2009). Scaffolds were manually ordered to match the dog Y assembly, which is based on BACs and hence is continuous.

The sequences were repeat masked using REPEATMASKER version 4.0.6 (Smit, Hubley, & Green, RepeatMasker Open-4.0. 2013–2015 http://www.repeatmasker.org) with "canidae" as species.

2.2 | Gene discovery and annotation

We downloaded paired-end transcriptome data from a male dog (SRA Accession numbers SRS072744–SRS072749; Hoeppner et al., 2014) and mapped them to the wolf assembly using HISAT2 (Kim, Langmead, & Salzberg, 2015). After extracting reads mapping to Y-linked scaffolds, we ran CUFFLINKS (Trapnell et al., 2010) separately for each tissue. We also used annotated, Y-linked genes from dog (GenBank accession KP081776.1) and blasted them onto the scaffolds. We then manually inspected all loci in IGV (Thorvaldsdottir, Robinson, & Mesirov, 2013) and corrected the annotation provided by CUFFLINKS if needed, such as when duplications were evident from genomic coverage. We also used BLAST results from dog Y-linked genes to combine parts of transcripts present on different scaffolds.

2.3 | Using read coverage to estimate copy number of multicopy gene families

Per-base coverage of all coding sequences was extracted for each individual using BEDTOOLS GENOMECOVERAGEBED. We calculated the mean and median coverage for each gene, and also noted the mode, namely the most abundant coverage for each gene. All these values were normalized relative to the coverage of all known single-copy genes. This should give mean and median numbers close to one for single-copy genes, close to two for genes with two copies, and so on.

2.4 | Sequencing

We sequenced 75 male wolves from Finland following the same protocol as described by Kardos et al. (2018). Briefly, DNA was prepared from frozen tissue samples and sequenced on an Illumina HiSeqX device, using a pair-end approach with 150-bp read length and 350-bp insert size, aiming for >30× mean autosomal coverage per individual.

2.5 | Variant detection

The following refers to male samples only. We downloaded whole-genome Illumina paired-end sequence data from 53 grey wolves

from Scandinavia (mostly Sweden plus some from Norway), 11 from China, one each from Korea, India and Italy, and three from the United States. We also downloaded data from two eastern wolves (*Canis lycaon*, also known as eastern timberland or Algonquin wolves), two red wolves (*Canis rufus*), four coyotes (*Canis latrans*) and 24 dogs (one dingo, 10 pure breeds and 13 marked as "indigenous") (see Supporting Information Table S2 for accession numbers). As this study focused on grey wolves, the inclusion of a set of dogs merely served as a reference to previously identified Y chromosome haplotypes in canids. All these sequences, together with the 75 newly sequenced Finnish wolves, were mapped to the wolf genome assembly with BWA, in the same manner as above. The use of the whole assembly at this step, rather than Y-linked scaffolds only, was motivated by the fact that reads from elsewhere in the genome (e.g., from related sequences on the X chromosome) could falsely map to the Y chromosome in the absence of their true target sequence. Reads mapping to Y-linked scaffolds were deduplicated as above and realigned using GATK version 3.8 (McKenna et al., 2010). Finnish samples were base calibrated with BQSR according to the "GATK Best Practices" (Van der Auwera et al., 2013).

Variants were first called in each sample separately with GATK's HAPLOTYPECALLER. Then joint genotyping was performed merging all samples using GATK's GENOTYPEGVCFS. Because we lack a proper Y chromosome reference single nucleotide polymorphism (SNP) set, variant quality score recalibration (VQSR) could not be performed. Instead, SNPs were hard filtered with VARIANTFILTRATION using the settings --filterExpression "QD < 2.0 || FS > 60.0 || MQ < 40.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0", following Alternative Protocol 2 in the "GATK Best Practices" (Van der Auwera et al., 2013). We used both diploid and haploid SNP calling, which at first glance may seem unnecessary for a haploid chromosome. However, given the repetitive nature of the Y chromosome with multiple copies of similar sequences potentially collapsing in the assembly, identification of heterozygous "SNPs" provided a means for detection of collapsed duplicates. We noted that the distribution of distances between adjacent heterozygous calls was heavily skewed towards zero (Supporting Information Figure S2), consistent with collapsed regions manifesting as a strong clustering of (seemingly) heterozygous sites. We used a 95% cut-off from this distribution and removed those regions located between pairs of adjacent heterozygous sites closer to each other than this cut-off (181 bp). We also removed 90 bp (half of 181) flanking sequence around each separate cluster. The few heterozygous sites that failed to be filtered with this method were excluded. Another option would have been to filter based on coverage, for example, removing all sites that had higher than twice the expected coverage. However, we noted that sequence coverage varied substantially over the Y chromosome, indicating that this method would have removed a high proportion of nonduplicated sites while keeping many low-coverage regions with heterozygous sites. We also removed all repeats, insertions and deletions called by GATK, as well as gaps in the reference and coding sequence. Single genotypes were additionally filtered based on genotype quality and callability, so that any call with GQ < 30 or not reported as callable in GATKs CALLABLELOCI for a specific individual was set to N.

Two of the North American wolves (Yellowstone) were father and son (Fan et al., 2016). In the initial variant calling they differed at a single site, which upon manual inspection was found to be probably due to a genotyping error. For the remaining analyses, one of these individuals was removed.

We also called variants in 10 female wolves from Kardos et al. (2018), following the above procedure and criteria (for accession numbers see Supporting Information Table S3). This was used as an extra control to avoid repetitive, ambiguous or misassembled regions. None of the variant sites detected in males was called in females, strengthening the inference that all selected scaffolds were on the Y chromosome.

2.6 | Imputation of missing data

We imputed missing calls using the method described by Barbieri et al. (2016), and adapted from Lippold et al. (2014). In short, a missing call was imputed if the three genetically most similar individuals, based on pairwise distances, all had the same genotype. If they disagreed, or if any of them also failed to be called, the site was discarded. In total, 1% of the individual genotypes were imputed.

2.7 | Haplotype analysis

A median-joining haplotype network based on 1,177 variable sites genotyped in all 176 individuals was constructed and plotted with POPART (Leigh, Bryant, & Nakagawa, 2015). To relate our sequences to existing dog Y chromosome haplotype nomenclature, we downloaded Y haplotype fragments from Ding et al. (2012) and Natanaelsson et al. (2006) (accession numbers: DQ973626.1–DQ973805.1 and HQ389365–HQ389435) and used BLAST to find their location in the wolf genome assembly. We found the positions for all variable sites (given in data set 2 of Ding et al., 2012) except one, which fell into an assembly gap. We could then match our individuals with the given haplotypes H1–H31. The haplogroups defined by Ding et al were later slightly reorganized by Shannon et al. (2015); we use the latter for assigning the haplotypes into haplogroups.

2.8 | Estimating the divergence time

We used BEAST version 1.8.4 (Drummond, Suchard, Xie, & Rambaut, 2012) for phylogenetic tree reconstruction and estimation of coalescent times. Only one individual per haplotype was used, reducing the number of individuals substantially due to the high level of haplotype sharing in the Scandinavian and Finnish samples. We removed the potentially hybridizing eastern and red wolves, and used only one coyote as an outgroup (an individual from California determined to be a pure coyote in vonHoldt et al., 2016). The GTR+I+G model was chosen as the best-fitting model based on JMODELTEST (Darriba, Taboada, Doallo, & Posada, 2012), and we used a strict clock and set the chain length to 100,000,000 steps, logging parameters every 10,000 steps. We ran five independent runs that were combined with LOGCOMBINER. A maximum

clade credibility (MCC) tree was generated using a 10% burn-in with BEAST'S TreeAnnotator and drawn with FIGTREE (http://tree.bio.ed.ac.uk/software/figtree/).

In the absence of calibration points based on, for example, fossil records, mutation rate estimates are necessary for estimating divergence times. The mutation rate in wolves or dogs has not been directly estimated. Several authors have used 1×10^{-8} per site and generation referring to Lindblad-Toh et al. (2005). However, that study merely used this rate as a predefined parameter in coalescent simulations without justification. Three recent studies have estimated the genomic mutation rate using ancient samples (one wolf and two dogs) as calibration points. The resulting estimates were 0.4×10^{-8} (Skoglund et al., 2015), $0.3–0.45 \times 10^{-8}$ (Frantz et al., 2016) and 0.56×10^{-8} (Botigué et al., 2017). These studies based their calculations on a generation time set to 3 years. However, it has been suggested that the generation time of wolves is 4.2–5 years (e.g. Mech, Barber-Meyer, & Erb, 2016; vonHoldt et al., 2008). Using 4.5 years as generation time increases the cited mutation rate estimates to an interval of $0.45–0.84 \times 10^{-8}$. If we set the rate to 0.6×10^{-8} and assume a male-to-female mutation rate ratio of 2.0 (Wilson Sayres, Venditti, Pagel, & Makova, 2011), a male (Y chromosome) mutation rate of 0.8×10^{-8} is obtained, which we use in this study. We acknowledge that many assumptions are needed to arrive at this estimate, and that it should be seen as an approximation.

Additionally, we used the rho statistic (Jobling, Hollox, Hurles, Kivisild, & Tyler-Smith, 2014) to estimate divergence times based on averaging the counts from each tip to the root. The counts were taken from the network in Figure 3a.

2.9 | Inferring haplogroup in an ancient wolf sample

We downloaded whole genome data from an ancient wolf excavated in Taimyr, Russia (Skoglund et al., 2015), accession number ERR868147 and mapped the reads in the same way as above. However, because coverage was less than 1× it was not possible to genotype it together with the other samples. Instead, we only considered sites already identified as variable in our set of wolves and dogs, and called the Taimyr wolf at those sites covered by at least one read. We are aware of the inexactness of this method, especially considering the high amount of errors in ancient DNA sequences, but the focus on sites and alleles known to segregate in wolf populations should reduce the risk of erroneous calls. Furthermore, we were only interested in placing the ancient sample in relation to the modern samples based on diagnostic SNPs from the different haplogroups. The variant sites covered by the Taimyr wolf were used for drawing a maximum parsimony tree in SEAVIEW v4.5 (Gouy, Guindon, & Gascuel, 2010).

3 | RESULTS

3.1 | Identification of wolf Y chromosome scaffolds

By comparing genomic coverage from resequencing data of multiple males and females we identified 120 Y-linked scaffolds in a wolf short-read genome assembly (Gopalakrishnan et al., 2017), summing to 4.68 Mb (Supporting Information Table S4). One of the scaffolds additionally contained 1.75 Mb of the known PAR. Synteny between the identified wolf Y chromosome scaffolds and a BAC-derived dog Y chromosome assembly (Li et al., 2013) is shown in Figure 1. Our approach identified more Y-linked sequence in total than the dog assembly, but an ampliconic region assembled in the latter was completely collapsed in the short-read assembly.

The Y chromosome scaffolds were repeat-rich with 55.9% of the nongap bases masked, compared to 39.8% in the full assembly. The most abundant repeat types were long interspersed nuclear elements (36.6%) and short interspersed nuclear elements (11.1%). There was a marked difference between the repeat landscapes of wolf Y chromosome and autosomal sequences (Supporting Information Figure S3). Most notably, the Y chromosome showed a pronounced relatively recent (corresponding to ≈10% sequence divergence) activity of LINE insertions, not seen in autosomes.

3.2 | Gene annotation

We found 22 genes on the wolf Y chromosome (Table 1), including all 18 genes previously reported as Y-linked in dogs (Li et al., 2013). Due to the fragmented nature of assembled Y chromosome sequences, some exons or parts of exons were not covered in the assembly. The multicopy genes TSPY, SRY and CUL4BY (see below) were found to be completely collapsed, with each of them assembled into a single copy. Similarly, the duplicated BCORY gene (BCORY1 and BCORY2) was partly collapsed and we found three new exons, not reported before, comprising the 5′ untranslated region and the start of the coding sequence. Genomic coverage suggested that this part was single copy and hence only present in either BCORY1 or BCORY2.

New Y chromosome genes not present in the dog Y chromosome assembly included TMSB4Y, which is Y-linked in primates but previously thought to be lost in Laurasiatheria (Cortez et al., 2014). It also included a canine orthologue of EIF2S3Y, in dogs previously only detected in RNA sequencing (Cortez et al., 2014; Li et al., 2013). We found two genes that to our knowledge have not been reported as Y-linked in other mammals. One was AP1S2 (AP-1 complex subunit sigma-2) present on the X chromosome in many mammals, including dog; we refer to the wolf Y chromosome copy as AP1S2Y. Transcriptome data show that it is broadly expressed across tissues in dogs, although the level of expression in different tissues has not been quantified (Supporting Information Table S5). A BLAST search with AP1S2Y identified a predicted gene (NCBI accession number XM_008683968) in polar bear (Ursus maritimus), with Y-linkage supported by male/female coverage differences in this species (Bidon, Schreck, Hailer, Nilsson, & Janke, 2015). This suggests that AP1S2Y has been retained on the Y chromosome of several species within Carnivora. The other new gene was WWC3, which is X-linked in dog and other mammals. The Y chromosome copy (here referred to as WWC3Y) was split over three different scaffolds, with several exons missing in between.
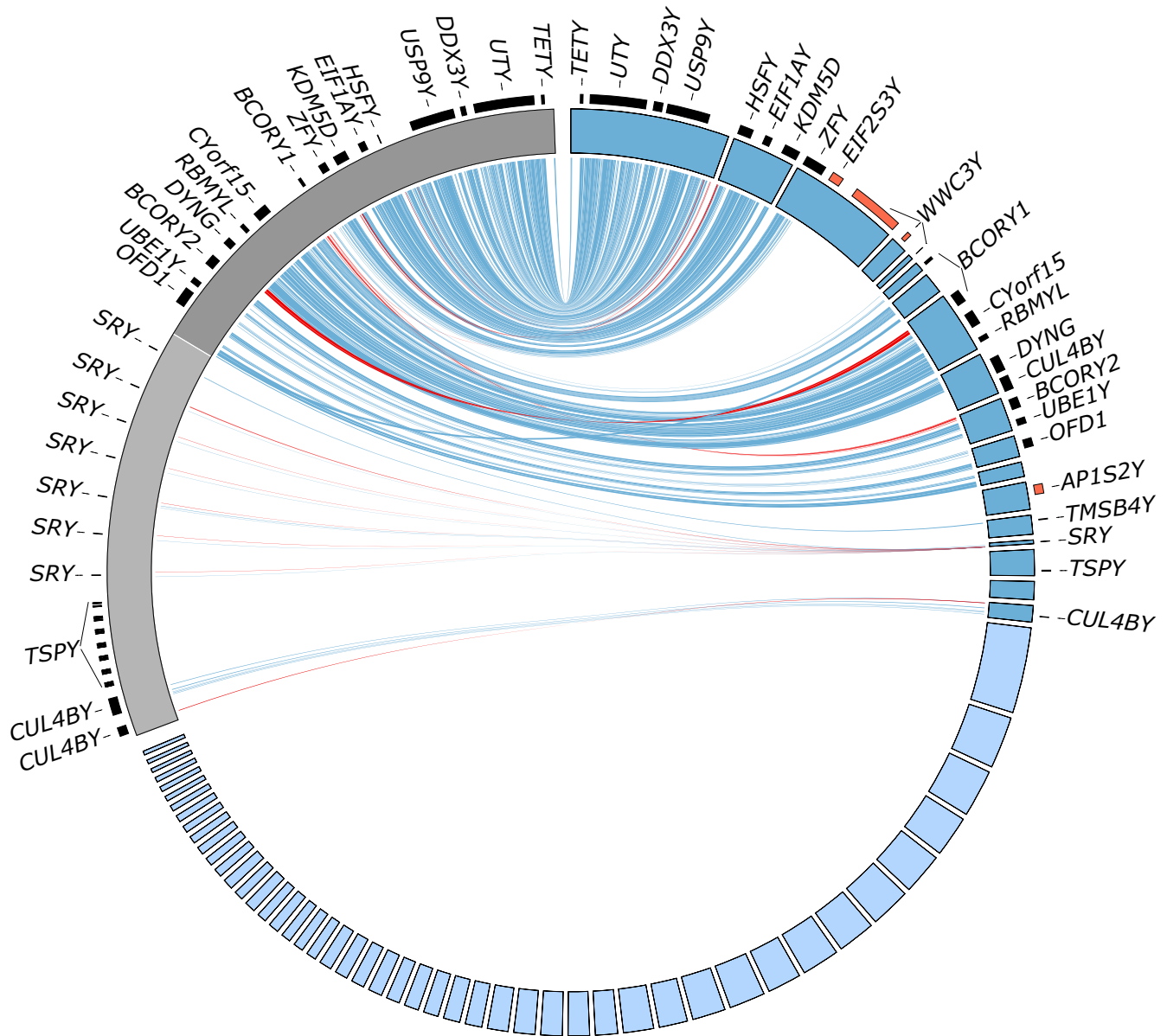
FIGURE 1 All-to-all nucmer alignments >500 bp between wolf Y chromosome scaffolds from Gopalakrishnan et al. (2017) identified in this study (right, in blue) and a dog Y chromosome bacterial artificial clone (BAC) assembly (left, in grey, from Li et al., 2013). Forward alignments are drawn in blue, reverse alignments in red. Wolf scaffolds with alignment anchors and/or genes are ordered according to the dog assembly. These are followed by scaffolds with unknown position (lighter blue) in descending order. New genes found in wolf are marked in red. The ampliconic region in the dog assembly is shown in lighter grey. Note that some wolf genes span more than one scaffold and that *CUL4BY* spans two nonadjacent scaffolds according to the alignment [Colour figure can be viewed at wileyonlinelibrary.com]

3.3 | Copy number variation of Y-linked genes

We estimated copy number based on normalized coverage for each Y-linked gene (Figure 2a). *TSPY* was present in ~100 copies, *SRY* in three and *UBE1Y* in two. Coverage in *OFD1* suggested that it was duplicated, but manual inspection revealed female coverage, and blasting of the X-linked copy of *OFD1* from dog suggested that the X- and Y-linked copies were collapsed in the assembly. *BCORY2* and *CUL4BY* were both ambiguous, with some exons apparently presenting as single-copy whereas others were amplified. This divided both genes into one single-copy and one

multicopy part, with the amplified region of *CUL4BY* represented by ≈10 copies.

Variation in coverage among individuals for some of the multicopy genes (Figure 2a) could indicate that there is copy-number variation in wolf populations, but may also simply reflect technical limitations of copy-number estimation related to stochastic variation in sequence coverage. To reduce the noise in estimation from single individuals, we grouped wolves by the four main Y chromosome haplogroups to be described below, and treated dogs as a separate group. This suggested that there was indeed copy-number variation for *TSPY*, with dogs having the fewest and wolves from haplogroup

HG0 the largest number of copies (Figure 2b). It also indicated that the number of *CUL4BY* gene copies varied among haplogroups.

3.4 | Wolf Y chromosome haplotypes

We called single nucleotide variants in the identified Y chromosome sequences from a sample of 144 resequenced male wolves. These included 75 Finnish wolves newly sequenced for the purpose of this study, and publicly available data from 53 wolves from Scandinavia and 16 wolves from elsewhere in the world. We also included available sequences from two red wolves, two eastern wolves, 24 dogs including one dingo, and four coyotes. Because of the high fraction of repetitive sequence and many collapsed duplicate regions in the Y chromosome scaffolds, we performed extensive filtering to decrease the risk of calling false variants (see Section 22). After filtering, ~600 kb of single-copy, high-quality Y chromosome sequence remained, in which we were able to call 1,524 SNPs. We consider this to represent a conservative assessment of the amount of Y chromosome polymorphism present in the sample. After imputation, we

TABLE 1 Genes identified on wolf Y chromosome scaffolds. The position of genes within is each scaffold is indicated

Gene	Location in assembly (scaffold ID and position in scaffold)
AP1S2Y	scaffold_2775: 9–36 kb
BCORY1[a]	scaffold_2091: 134–169 kb; scaffold_3549: 8–16 kb
BCORY2[a]	scaffold_2578: 8–37 kb
CUL4BY[a]	scaffold_7290: 0.5–0.7 kb; scaffold_5293: 1.1–1.3 kb; scaffold_3306: 25–26 kb; scaffold_5774: 1.1–1.2 kb; scaffold_2411: 120–82 kb
CYorf15	scaffold_2091: 62–102 kb
DDX3Y	scaffold_242: 200–224 kb; scaffold_6535: 0–1,3 kb
EIF1AY	scaffold_2073: 84–104 kb
EIF2S3Y	scaffold_1620: 196–229 kb
HSFY[a]	scaffold_2073: 140–178 kb
KDM5D	scaffold_2073: 1–45 kb
OFD1	scaffold_2992: 38–14 kb
DYNG[a]	scaffold_2411: 21–65 kb
RBMYL[a]	scaffold_2091: 16–29 kb
SRY	scaffold_4057: 2–3 kb
TETY[a]	scaffold_242: 426–419 kb
TMSB4Y	scaffold_3047: 10–13 kb
TSPY	scaffold_2802: 55–58 kb
UBE1Y	scaffold_2578: 66–84 kb
USP9Y	scaffold_242: 69–187 kb
UTY	scaffold_242: 245–400 kb; scaffold_8081: 0.4–0.5 kb
WWC3Y	scaffold_1620: 150–4 kb; scaffold_3205: 22–34 kb; scaffold_3892: 10–6 kb
ZFY	scaffold_1620: 253–312 kb

[a]Specific to testis in transcriptome sequencing of blood, brain, heart, liver, lung, muscle and testis (Hoeppner et al., 2014).

then limited the data set to 1,177 variable sites genotyped in all 176 individuals.

We found 53 distinct Y chromosome haplotypes, 26 of which were seen in grey wolves (Supporting Information Table S2). A median-joining network revealed that grey wolves spread out over four larger clusters of haplotypes (Figure 3a). The network was well resolved without reticulations. Three of the four clusters corresponded to the previously defined dog haplogroups HG1–3/HG6, HG9 and HG23, whereas the fourth cluster was unique to wolves and was previously referred to as H27 by Ding et al. (2012), or simply "Asian wolves" by Oetjens et al. (2018); as shown below, it is also seen in European wolves. As other studies have used H27 to denote a different, dog-specific haplotype (Oetjens et al., 2018; Shannon et al., 2015), we renamed this haplogroup HG0 to avoid confusion (cf. de Groot et al., 2015). The number of diagnostic mutations for each haplogroup was 84 for HG0, 62 for HG23, 59 for HG9 and 46 for HG1–3/HG6 (see Supporting Information Table S6 for list of all diagnostic mutations). HG1–3/HG6 has previously been defined as two or even three separate haplogroups, but because these were so similar to each other compared to the three other haplogroups (HG6 had only four and HG1–3 two private mutations, respectively), we found it appropriate to refer to them as a single haplogroup. In this context we note that the definition of haplogroups among Y chromosome lineages is somewhat arbitrary and is always sensitive to the degree of resolution given by the number of individuals analysed, and their origin.

There was no clear geographical structure among the analysed grey wolf samples at a global scale (Figure 3b); European wolves were found in all four haplogroups and Asian wolves in three. That extensive Y chromosome diversity is present even on small geographical scales was evident from the finding of haplotypes from all four major haplogroups in the sample of Finnish wolves. At the same time, there were some instances of regional signatures of population structure in the form of phylogenetically very similar haplotypes detected in the same geographical area (Supporting Information Table S7). Examples of this included Finnish wolves within H23, Chinese wolves within HG0 and North American wolves within HG1–3/HG6. Because genome sequence data from only three male North American wolves were available, two of which were father–son (Fan et al., 2016), we cannot make strong phylogeographical conclusions for the New World. A recent study has suggested that all extant grey wolf lineages in the New World derive from a single colonization event of North America when a land bridge connected Eurasia and North America >23,000 years ago (Koblmüller et al., 2016; but see, e.g., Wayne, Lehman, Allard, & Honeycutt, 1992).

The four coyotes split into two deep branches, with the two eastern wolves clustering with coyotes from one of the branches and the two red wolves with coyotes from the other (Figure 3a). The relatively limited set of dogs included in the study mapped to two haplogroups, in two different clades within each group.

We used BEAST for phylogenetic reconstruction and dating of lineage splitting, including estimation of the time to the most recent common ancestor (TMRCA) of all grey wolf Y chromosome
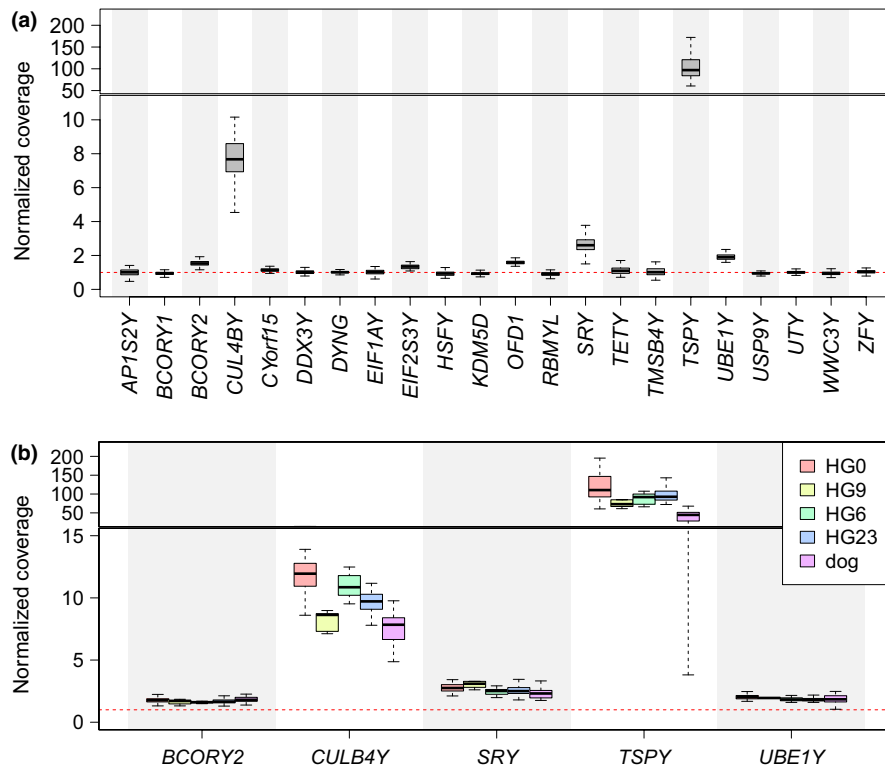
FIGURE 2 Normalized male read coverage in wolf Y chromosome coding sequences (as a proxy for gene copy number). Boxes with whiskers (outside the upper and lower quartiles) show the range of individual coverage within each gene. Note that the y-axes have two different scales. The red dashed line corresponds to one copy. (a) All genes. (b) Multicopy genes with wolves grouped into haplogroups. All dogs are combined into one group irrespective of their haplogroup [Colour figure can be viewed at wileyonlinelibrary.com]

haplotypes and using a coyote as outgroup. Critical to dating, we used a male mutation rate of 0.8×10^{-8} per site and generation, and assumed a generation time of 4.5 years (see Section 22). All wolf Y chromosome sequences coalesce 125,000 years ago (95% highest posterior density interval [HPDI] 103,000–126,000 years ago), representing the split between haplogroup HG0 and other lineages (Figure 4). HG23 diverged from HG1–3/HG6 and HG9 87,000 years ago (HPDI 74,000–92,000), and the two latter groups diverged 71,000 years ago (HPDI 62,000–78,000). Adding more outgroups (coyotes, red wolf, eastern wolf) did not significantly affect estimated coalescence times or changed the phylogeny (Supporting Information Figure S4). Estimates of divergence times using rho statistics gave very similar datings, with a TMRCA of all wolf haplotypes of 111,000 years ago (HPDI 94,000–123,000). The split between grey wolf and coyote lineages was estimated to have occurred 200,000 years ago (HPDI 188,000–223,000).

The age of dog patrilines is dependent on the number of domestication events. Assuming three such events in the haplogroups/subhaplogroups in which dog haplotypes were found (each group potentially representing a domestication event), the most recent splits between dog and wolf haplotypes in those groups were 29,000 (HPDI 23,000–32,000; haplogroup HG1–3), 26,000 (HPDI 17,000–26,000; HG6) and 24,000 (HPDI 13,000-–21,000; HG23) years ago.

We conservatively assumed one domestication event within HG23 because the single ingroup Italian wolf lineage could potentially be the result of an old wolf–dog hybridization event.

We also considered a 35,000-year-old wolf sample from Taimyr, Russia (Skoglund et al., 2015). Due to very low sequencing coverage we could only obtain variant calls from 267 of the variant sites. Of these, the Taimyr wolf matched 20 out of 25 (80%) variants unique to HG0, while it matched none of the variants unique to either HG1–3/HG6, HG9 or HG23. The Taimyr wolf thus appears to belong to haplogroup HG0. A maximum parsimony tree based on the 267 sites and including the Taimyr wolf supports this inference (Supporting Information Figure S5).

3.5 | Degree of Y chromosome diversity in Finnish grey wolves

Because our sampling scheme was strongly biased towards wolves from northern Europe, with the Scandinavian population recently passing through a sharp bottleneck followed by intensive inbreeding (see further below), we cannot obtain a representative estimate of global nucleotide diversity of wolf Y chromosomes. However, given that the Finnish population harboured haplotypes from all four major haplogroups, the degree of diversity in this population would provide some indication of wolf Y chromosome variability. Using one

FIGURE 3 Median joining network of Y chromosome haplotypes in dogs, wolves and coyotes. (a) All 176 samples in which the size of circles corresponds to the number of identical samples. Grey wolves are coloured according to country of origin. Haplotype labels are given for grey wolf haplotypes. (b) Only one individual per grey wolf haplotype, coloured according to their geographical origin. No haplotype was shared between grey wolves from different geographical regions. Bars indicate the number of substitutions per branch. Note that the networks were not rooted [Colour figure can be viewed at wileyonlinelibrary.com]
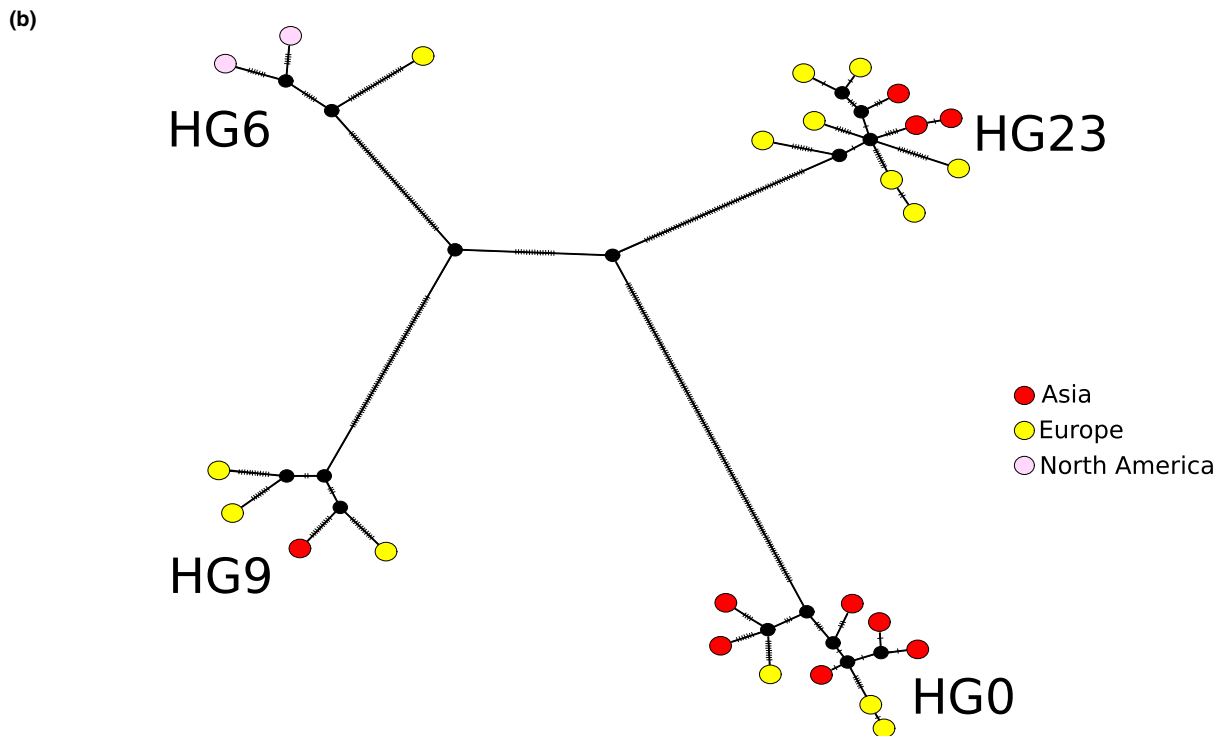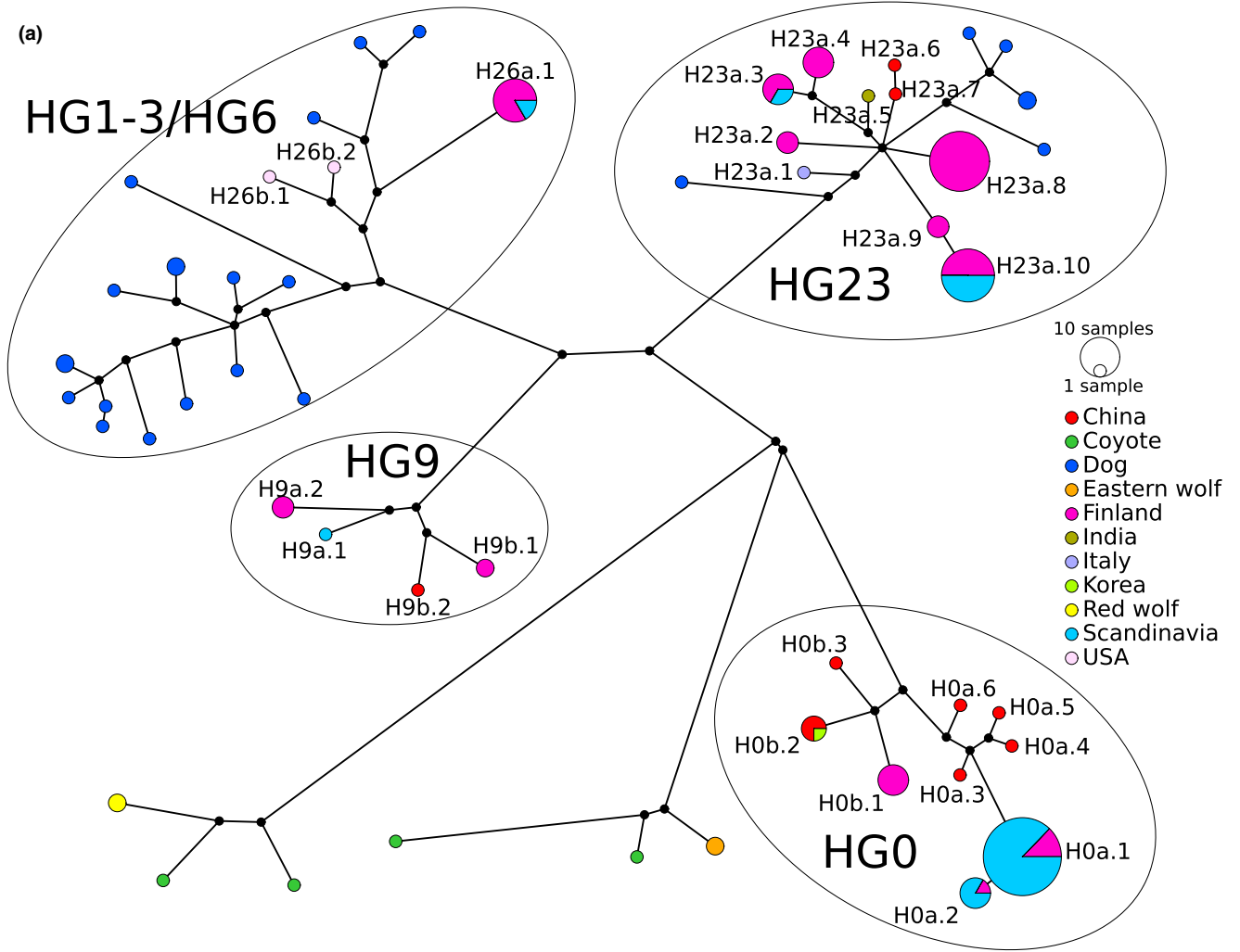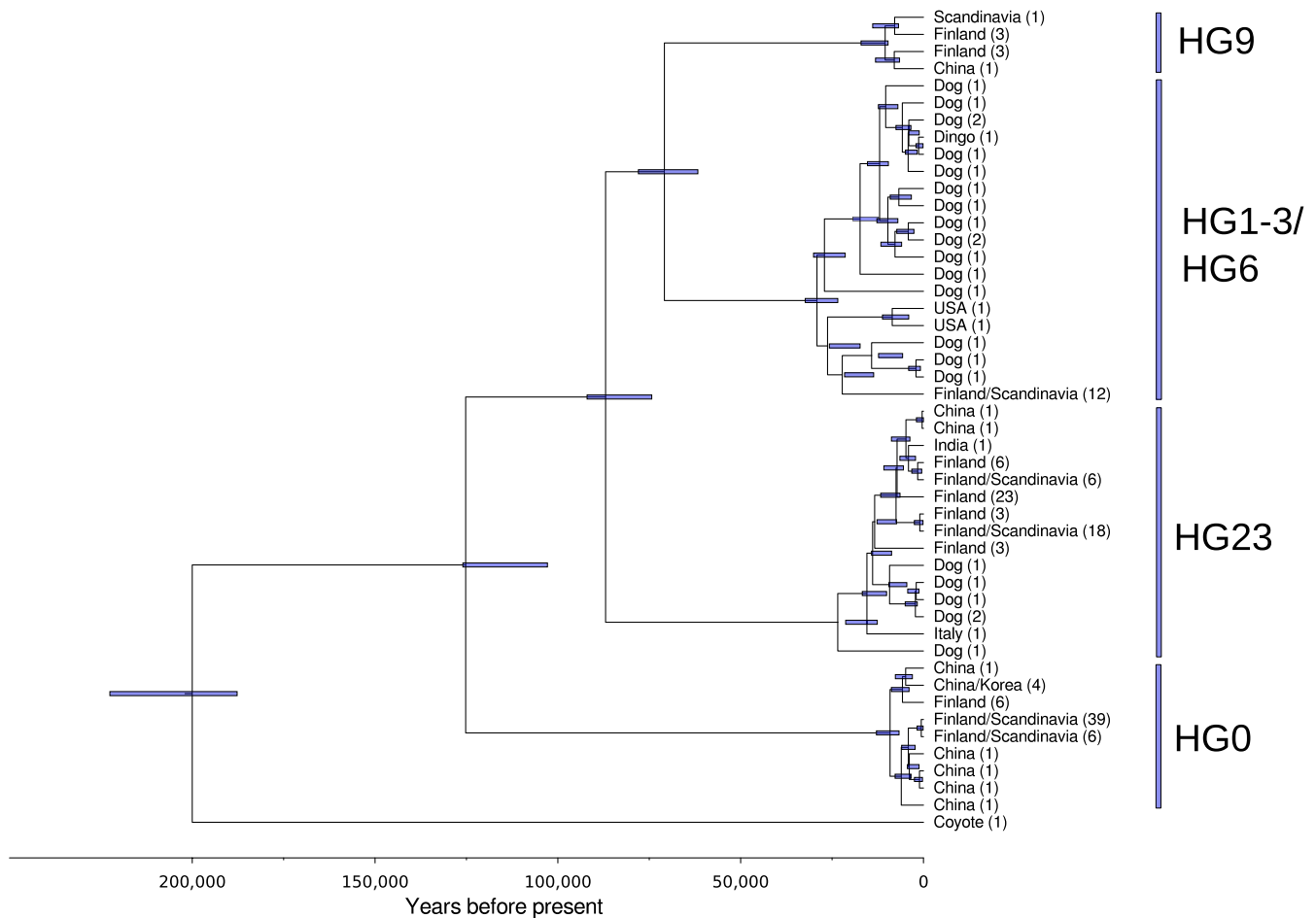
FIGURE 4 Phylogenetic tree of canid Y chromosome haplotypes reconstructed in BEAST. The time scale is based on a male mutation rate of 0.8×10^{-8} and a generation time of 4.5 years. Wolves are presented with their country of origin. Numbers in parentheses denote the number of samples with identical haplotypes. The 95% highest density posterior intervals are shown as blue horizontal bars. For dating of domestication events, one event was assumed in each of the two basal lineages within haplogroup HG1–3/HG6, and one within HG23 [Colour figure can be viewed at wileyonlinelibrary.com]

individual from each haplotype found in this population, the mean pairwise Y chromosome sequence difference between Finnish wolves was 0.00028, that is, on average one variable site per 3,569 bp.

3.6 | Origin of the Scandinavian wolf population

It has been assumed that the contemporary Scandinavian wolf population was founded by a single male and female in the early 1980s, followed by another male arriving in 1991 (Wabakken, Sand, Liberg, & Bjärvall, 2001; see Section 44). More recently, immigrants have been detected (Seddon, Sundqvist, Björnerfeldt, & Ellegren, 2006), some of which have reproduced (Åkesson et al., 2016). The temporal occurrence of different Y chromosome haplotypes in Scandinavia was consistent with this scenario. Haplotype H0a.1 was present in the Scandinavian population from 1984 and onwards, and haplotype H0a.2 from 1993 and onwards (Figure 5). These were the only haplotypes detected in the breeding population until 2008 when new immigrants started to reproduce. Seven immigrants from 2002–2013 showed four different haplotypes, none of them being haplotype H0a.1 or H0a.2. Immigrants probably originate from Finland and/or Russia.

Importantly, both haplotype H0a.1 and haplotype H0a.2 were also detected in the Finnish wolf population. The same was true for three out of the four haplotypes displayed by recent immigrants. None of the wolf Y chromosome haplotypes seen in Scandinavia, neither those of the founders nor those of recent immigrants, were detected in the Italian wolves or among Asian and American wolves. They were also not detected among the analysed dogs. These data are consistent with the hypothesis that the Scandinavian wolf population originates from immigrants from a geographically close population. However, in the absence of a large number of genome sequences from, for example, other parts of Europe, we cannot exclude other potential scenarios.

4 | DISCUSSION

We used a strategy for identification of wolf Y chromosome sequences based on comparison of male and female coverage in whole-genome resequencing reads mapped to a wolf genome assembly. This resulted in the assignment of 4.68 Mb from 120
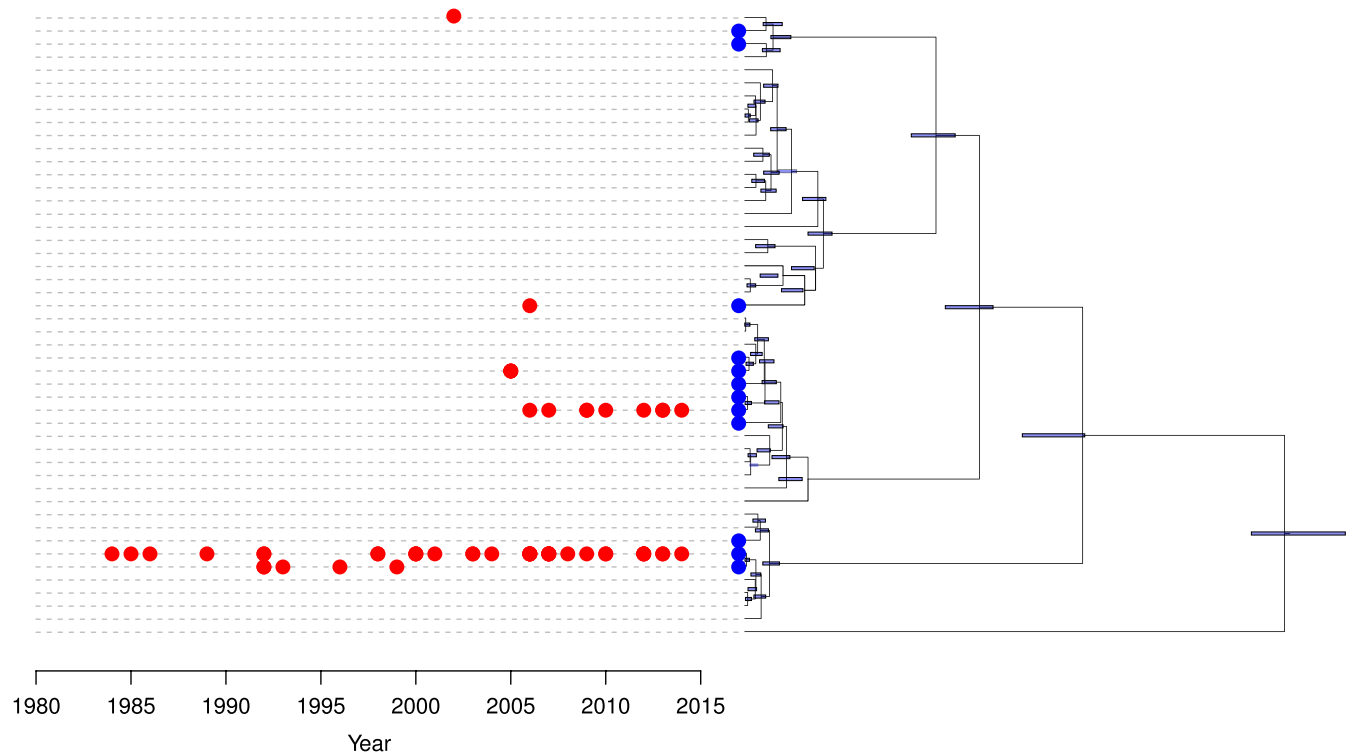
FIGURE 5 Detection of the different Y chromosome haplotypes in Scandinavia (red). Haplotypes present in Finland are marked in blue in the phylogenetic tree from Figure 4 shown to the right [Colour figure can be viewed at wileyonlinelibrary.com]

scaffolds originating from the Y chromosome and almost doubled the amount of available canine Y chromosome sequence; previously, Li et al. (2013) assembled ≈2.5 Mb of Y-linked sequence using dog BAC sequencing.

The canine Y chromosome has an estimated size of ≈20 Mb (Li et al., 2013). The p arm, which represents about half the chromosome, constitutes the heterochromatic nucleolus organizer region (NOR). On the q arm, the PAR is 6.6 Mb while the rest is divided into a single-copy and an ampliconic segment respectively (Li et al., 2013). With nearly 5 Mb of male-specific sequence now available, most of the euchromatic part of the Y chromosome should have been identified. The main differences between the sequence contained within the Y-linked scaffolds identified herein and the dog Y chromosome assembly of Li et al. (2013) are that we significantly extended the portion of unique sequence and that the dog assembly revealed the structure of the ampliconic region, which was largely collapsed in the short-read-derived wolf scaffolds.

4.1 | Geographical structure of wolf populations

Sequencing of population samples rather than genotyping of SNPs ascertained in a small panel of individuals was a major advantage of this work. Specifically, whole-genome resequencing of 75 male Finnish wolves augmented with available male sequence data from 69 grey wolves sampled elsewhere, 24 dogs and eight other canids allowed us to reconstruct the phylogenetic relationships among wolf Y chromosome lineages. The evolutionary history of grey wolves has been considered complex and previous work has revealed

conflicting evidence concerning population differentiation and geographical structure. One confounding factor is that hybridization between canid species is common (Gottelli et al., 1994; Hailer & Leonard, 2008; Wayne, 1993), including hybridization between grey wolves and dogs (Fan et al., 2016; Godinho et al., 2011; Pilot et al., 2018; Pires et al., 2017; Randi et al., 2000). In particular, hybridization events between coyotes, grey wolves and potentially other ancestral canid species/lineages in North America have muddled the taxonomic status of several extant wolf ecotypes in the New World (Hohenlohe et al., 2017; vonHoldt et al., 2016, 2011; Wayne & Jenks, 1991; Wheeldon, Rutledge, Patterson, White, & Wilson, 2013). Although not a focus of this study, our Y chromosome data confirmed admixture among North American canids (vonHoldt et al., 2016; but see Hohenlohe et al., 2017) and were consistent with paternal coyote (or other ancestral canid) introgression into red wolf and eastern wolf populations (Figure 3a). The analysed coyotes were represented by two very divergent haplotype lineages, which were as divergent from each other as they were from grey wolf lineages.

Initial mtDNA analyses indicated limited worldwide phylogeographical structure in grey wolves (Ersmark et al., 2016; Vilà et al., 1999; Wayne et al., 1992), a somewhat unexpected observation for a mammalian species even when considering the significant dispersal capacity of wolves (Kojola et al., 2006; Linnell, Brøseth, Solberg, & Brainerd, 2005). More extensive sampling and, in particular, use of nuclear markers have subsequently provided evidence for subpopulation structure on geographical and/or environmental scales (Aspi et al., 2006; Carmichael, Nagy, Larter, & Strobeck, 2002; Ersmark et al., 2016; Musiani et al., 2007; Pilot et al., 2010,2006; Schweizer,

Robinson et al., 2016a; Schweizer, vonHoldt et al., 2016b; vonHoldt et al., 2011; Zhang et al., 2014). An analysis of genome sequences from some 20 wolves sampled across the world recently demonstrated that the divergence between New World and Old World wolves constituted the earliest branching event among wolf populations (Fan et al., 2016). It also separated European from Middle Eastern wolves, and lowland from highland Asian wolves.

Interestingly, our Y chromosome data provided a mixed pattern of genetic structure. The main haplogroups were not partitioned on a continental scale, similar to the situation for mtDNA lineages (Vilà et al., 1999). In fact, all four Y chromosome haplogroups were represented among wolves from a relatively small geographical area in Finland. Moreover, three of the four haplogroups were present in the relatively limited sample (12 individuals) of Chinese wolves. In contrast, the lack of sharing of individual haplotypes between geographically distant areas indicated population structure on a regional scale. This was best illustrated by the extensive sampling of Finnish wolves. None of the 12 haplotypes detected in Finland were found in samples representing other parts of the world, except in neighbouring Sweden. However, they were frequently resampled in Finland. This resembles the situation seen in a comparison of microsatellite-based Y chromosome haplotypes in Italian and Russian wolves (Sastre et al., 2011). Population structure was also indicated because genetically similar haplotypes were only found within the same geographical region. We will discuss these results further below.

4.2 | Divergence times

There has been considerable discrepancy among the results from previous attempts to estimate the divergence times of wolves and wolf populations. Vilà et al. (1999) estimated that wolf mtDNA lineages coalesced 290,000 years ago. Later studies using mtDNA from both extant and ancient samples have shortened this estimate (Koblmüller et al., 2016; Matsumura, Inoshima, & Ishiguro, 2014; Skoglund et al., 2015; Thalmann et al., 2013), but the precise datings have varied depending on, among other things, the samples included in the analyses. For example, Skoglund et al. (2015) dated the most basal split of wolf mtDNA lineages to close to 80,000 years ago but excluded outgroup Chinese wolf samples from the analysis. Moreover, most of these results derive from a commonly used calibration point of 1–2 million years for the split between grey wolves and coyotes, a divergence time that recently has been questioned in favour of a much more recent split between the two species (Fan et al., 2016; vonHoldt et al., 2016).

Recent work has used genome sequence data and the generalized phylogenetic coalescent sampler method (G-PhoCS, a Bayesian demography inference method) to estimate the divergence time of wolf populations (Fan et al., 2016; Freedman et al., 2014; vonHoldt et al., 2016). This has suggested much more recent splits than indicated by mtDNA analyses, with wolf populations diverging 11,000–13,000 years ago in Fan et al. (2016) and 11,700–15,100 years ago in Freedman et al. (2014). We estimated the TMRCA of wolf Y chromosome lineages to 125,000 years ago

(95% HPDI, 103,000–126,000), considerably older than these recent estimates of the divergence of wolf populations. Estimates of coalescent and divergence times are critically dependent on the mutation rate applied and there are several sources of uncertainty behind the Y chromosome mutation rate of 0.8×10^{-8} that we used—see Section 22, and Jobling and Tyler-Smith (2017) for a general discussion. These include uncertainty about the overall genomic mutation rate in wolves and the extent of the male mutation bias. Mutation rate heterogeneity related to base composition (notably CpG mutability) and sequence neutrality are other factors that can bias divergence time estimation. Moreover, time estimates are clearly also dependent on the generation time used. If we apply a generation time of 3 years and a mutation rate of 1×10^{-8} (as used by Fan et al., 2016), the basal split between HG0 and other wolf Y chromosome lineages is estimated to 50,000 years ago, still considerably longer ago than the previous estimates.

The relatively old age of wolf Y chromosome lineages in relation to the timing of population divergence is not necessarily surprising because estimating coalescence times of individual loci is not the same as estimating the time of population divergence (Rosenberg & Feldman, 2002). Segregation of ancestral variation at the time of population divergence implies that coalescence times of extant lineages is the sum of the time to population divergence and the TMRCA of lineages segregating at the time of population divergence. There is one population divergence time but there is a unique coalescence time for any locus and any given set of individuals analysed (see Section 44 of G-PhoCS in this context by Gronau, Hubisz, Gulko, Danko, & Siepel, 2011). Patrilines represent just one realization of the evolutionary process, similar to the situation at a particular autosomal locus. As a single segregating unit, the Y chromosome is sensitive to stochasticity in lineage sorting.

Demographic analyses have suggested a relatively recent bottleneck of wolves approximately 10,000–20,000 years ago (Fan et al., 2016; Freedman et al., 2014; but see Gopalakrishnan et al., 2017, for the timing), with at least a three-fold reduction in the effective population size. This should have promoted genetic differentiation in connection with wolf populations diverging at about the same time, thereby contributing to signatures of global and regional population structure seen in genomic data. However, it does not exclude late Pleistocene/early Holocene survival of ancient Y chromosome or mtDNA lineages. This is related to the finding of a 35,000-year-old grey wolf from Taimyr assigned to HG0, showing that the lineage leading to present-day HG0 haplotypes existed long before the Last Glacial Maximum. The observation that the Taimyr wolf Y chromosome haplotype represented the deepest split within HG0 was not unexpected given the age of the Taimyr sample and the estimation of the TMRCA of present-day HG0 haplotypes to about 10,000 years ago.

4.3 | Relationship to other canine Y chromosome studies

A few previous studies have characterized canid Y chromosome haplotypes by sequence analysis or genotyping of SNP arrays (Ding et

al., 2012; Natanaelsson et al., 2006; Sacks et al., 2013; Shannon et al., 2015). These studies were mostly designed to study dog domestication with a limited number of wolf samples included. With 144 analysed wolf genomes, our sample vastly exceeds what have been used in previous genomic work of wolves in general and of Y chromosomes in particular. On the other hand, with genome sequence data from a large population sample (Finland) augmented with available data from mostly a limited number of wolves from different parts of the world, we do not have a fully representative data set from the whole distribution range of wolves.

The study that is most relevant to the present work is that of Oetjens et al. (2018), who mapped reads from 13 wolf genomes, and many dog genomes, to the 2.5-Mb Y chromosome sequence identified by Li et al. (2013). At the haplogroup level, the topology of the Y chromosome phylogenies in the two studies is identical. However, Oetjens et al. (2018) found that a Great Lake wolf sampled in Minnesota was a distant outgroup to all four major haplogroups of grey wolves. This sample was also included in our study, named "Eastern wolf," and had a haplotype identical to another eastern wolf from Algonquin Provincial Park, Canada. This haplotype clearly assigned our "Eastern wolf" to one of the two coyote lineages that was not sampled by Oetjens et al. (2018).

Oetjens et al. (2018) obtained much older time estimates for the split of Y chromosome lineages, with the basal divergence between HG0 and other haplogroups placed at 767,000 years ago (HPDI 303,000–1,392,000), compared to 125,000 years ago in our study. This difference seems primarily to be due to the calibration of their estimates by assuming a TMRCA of wolves and coyotes of 1.5 million years. From this they obtained a mutation rate estimate of $2.5–3.1 \times 10^{-10}$ per year, which corresponds to $1.1–1.4 \times 10^{-9}$ per generation assuming a generation time of 4.5 years. The latter an order of magnitude lower than the rate we applied (0.8×10^{-8}) and also much lower than those used in other recent genomic studies (e.g., Fan et al., 2016; Freedman et al., 2014; but see Skoglund et al., 2015).

Several studies have also used a limited set of microsatellites, mainly four markers that we previously described (Sundqvist, Ellegren, Olivier, & Vilà, 2001), to define Y chromosome haplotypes in wolves and, mostly, dogs (Benson, Patterson, & Wheeldon, 2012; Brown et al., 2015; Fabbri et al., 2014; Randi et al., 2014; Sacks et al., 2013; Vilà, Walker et al., 2003; Wheeldon et al., 2013; Wilson, Rutledge, Wheeldon, Patterson, & White, 2012). These microsatellite-derived haplotypes have not been anchored to sequence-based canine Y chromosome haplotypes, and it is therefore difficult to compare studies using the two different approaches. An important topic for future research will be to establish the relationship between canid Y chromosome haplotypes defined by microsatellites and SNPs (cf. de Groot et al., 2015).

4.4 | The origin of the Scandinavian grey wolf population

Following a long period of population decline due to human persecution, grey wolves became functionally extinct on the Scandinavian

peninsula in the late 1960s (e.g., Linnell et al., 2005). In the early 1980s, grey wolves again started to reproduce in Scandinavia, initiating a steady settlement with an increasing population size reaching several hundred individuals in the last few decades (Åkesson et al., 2016; Wabakken et al., 2001). Genetic data have confirmed that the population was founded by a single pair, followed by the arrival of another male in 1991 (Vilà, Sundqvist et al., 2003). In the absence of further successful immigration until 2005, the population has become highly inbred with inbreeding coefficients of some individuals estimated from pedigree records of up to 0.5 (Åkesson et al., 2016; Bensch et al., 2006; Liberg et al., 2005). There is extensive linkage disequilibrium in the population (Hagenblad, Olsson, Parker, Ostrander, & Ellegren, 2009) and huge tracts of runs of homozygosity forming large genomic regions of identity-by-descent are seen in many individuals (Kardos et al., 2018).

The fact that the recolonization event somewhat unexpectedly took place in southern Sweden, >1,000 km from the nearest regular breeding grounds in northern Finland, fuelled a public debate on the origin of the contemporary population (Linnell et al., 2005). This has remained an issue over the years despite genetic evidence in favour of an origin by immigrants from the east (Flagstad et al., 2003). The controversy has included alternative scenarios such as illegal reintroduction of wild or captive wolves, and dog ancestry (Linnell et al., 2005). As late as 2015, the Norwegian parliament called for new investigations concerning the origin of the Scandinavian wolf population, including potential hybridization with dogs (https://www.stortinget.no/no/Saker-og-publikasjoner/Vedtak/Vedtak/Sak/?p=65090). In addition, there is strong controversy regarding management of the Scandinavian wolf population, in particular concerning culling (Immonen & Husby, 2016). Following approval of licensed hunting by the Swedish Environmental Protection Agency, objections have been made by the European Commission to the Swedish government that such hunting would be an infringement of The Habitats Directive (Darpö, 2016; Epstein, 2016).

Our results are important for the question of the origin of the Scandinavian wolf population in at least two respects. First, the haplotypes of the two founding males in the early 1980s and 1991 (H0a.1 and H0a.2) were both present in the contemporary Finnish population, but were not detected in wolves from elsewhere in the world or in dogs. Similarly, most haplotypes of recent immigrants to Scandinavia were also present in the Finnish population, but not seen in other wolves or in dogs. This provides strong support that the patrilines of the Scandinavian wolf population originated from a geographically close wolf population, and that the same applies to recent immigrants.

Second, and in relation to potential hybridization with dogs, we note that both founder haplotypes were from haplogroup HG0. None of the 24 dogs that we included belonged to this haplogroup. More importantly, haplogroup HG0 appears to be wolf-specific as previous work based on other Y chromosome marker sets has failed to detect dogs with haplotypes from this group, despite >1,000 dogs from numerous breeds being genotyped (Shannon et al., 2015). This provides strong evidence against the recent paternal contribution of

dogs to the Scandinavian wolf population. It does not exclude the possibility that hybridization with dogs occurred in the past, but in that case there appear to be no surviving dog paternal lineages left. Wolf–dog hybridization is typically asymmetrical in the direction of male dog × female wolf, although rare instances of the opposite have been reported (Hindrikson, Mannil, Ozolins, Krzywinski, & Saarma, 2012; Muñoz-Fuentes, Darimont, Paquet, & Leonard, 2010).

4.5 | Dog domestication

This study was not designed to address dog domestication, but the Y chromosome haplotype tree provides some information pertinent to this question. First, dogs were found in two of the major haplogroups (HG1–3/HG6 and HG23) and in three clusters if dividing the large HG1–3/HG6 haplogroup into HG1–3 and HG6. This is consistent with several independent domestication events (Frantz et al., 2016), or at least several instances of wolf paternal introgression into dog populations. Second, datings of the most recent split between wolf and dog haplotypes within each haplogroup/cluster (29,000, 26,000 and 24,000 years ago, respectively) support an Late Palaeolithic origin of domestic dogs. This is in line with genomic data suggesting that dogs were domesticated before the rise of agriculture (Botigué et al., 2017; Freedman et al., 2014; Skoglund et al., 2015). However, it is important to emphasize that only a limited number of wolf genome sequences were available from some parts of the world. Moreover, the sample of dogs analysed may not represent the full range of genetic diversity present among dogs.

4.6 | Y chromosome genes and copy number variation

All four new genes (*AP1S2Y*, *EIF2S3Y*, *TMSB4Y* and *WWC3Y*) that we identified on the wolf Y chromosome are gametologous copies of X-linked genes. They are expressed in testis in dogs but none of them shows testis-specific expression. As such, they may be considered to represent the group of dosage-sensitive genes that have survived Y chromosome degeneration by selection for maintenance of gene dose of functionally equivalent gametologues in males and females (Bellott et al., 2014). However, *AP1S2Y* is not Y-linked in most other mammals and thus does not belong to a core set of genes broadly retained on the Y chromosome in different mammalian lineages during sex chromosome evolution.

Sequence coverage confirmed the presence of several multi-copy genes on the wolf Y chromosome (cf. Li et al., 2013). Our data showed that *TSPY* is present in numerous copies, in the order of 100, with indications of copy number variation within wolf populations. Resolution was not sufficient for precise estimation of copy number at the individual level but when we partitioned wolves based on Y chromosome haplogroups and treated dogs as a separate group, the latter clearly had fewer copies (median = 45) than all wolves (haplogroup medians of 110, 73, 92 and 93). Using quantitative PCR (qPCR), Li et al. (2013) estimated *TSPY* copy number in three dogs to 25–35. Partitioned analysis of *CUL4BY* also suggested there being

copy-number variation, in this case visible between wolf haplogroups. Frequent structural variation including copy-number variation appears to be a hallmark of the ampliconic region of the mammalian Y chromosome (Oetjens, Shen, Emery, Zou, & Kidd, 2016; Repping et al., 2006), and may potentially have functional implications.

The male-determining *SRY* gene is present in multiple copies in some mammalian lineages, notably in many rodent species (Nagamine et al., 1994) but also in Carnivora (Pearks Wilkerson et al., 2008), Perissodactyla (Han et al., 2017) and Lagomorpha (Geraldes & Ferrand, 2006). Our data suggested that there are three *SRY* copies in wolves and dogs, and clearly not more than this. In contrast, qPCR data from three dogs suggested them as having seven copies (Li et al., 2013). Moreover, our coverage data suggested that *UBE1Y* is present in two copies while Li et al. (2013) identified it as a single-copy gene with qPCR. These discrepancies call for further work and may illustrate the technical challenges associated with copy-number determination based on sequence coverage or with qPCR. Alternatively, the differences might be due copy-number variation within these genes.

AUTHOR CONTRIBUTION

L.S. performed all analyses. I.K. provided samples. L.S. and H.E. wrote the paper. H.E. conceived of the study and supervised the project.

ORCID

Linnéa Smeds https://orcid.org/0000-0002-8415-9259

Hans Ellegren https://orcid.org/0000-0002-5035-1736

REFERENCES

Åkesson, M., Liberg, O., Sand, H., Wabakken, P., Bensch, S., & Flagstad, Ø. (2016). Genetic rescue in a severely inbred wolf population. *Molecular Ecology*, *25*(19), 4745–4756. https://doi.org/10.1111/mec.13797

Aspi, J., Roininen, E., Ruokonen, M., Kojola, I., & Vilà, C. (2006). Genetic diversity, population structure, effective population size and demographic history of the Finnish wolf population. *Molecular Ecology*, *15*(6), 1561–1576. https://doi.org/10.1111/j.1365-294X.2006.02877.x

Bachtrog, D. (2013). Y-chromosome evolution: Emerging insights into processes of Y-chromosome degeneration. *Nature Reviews Genetics*, *14*(2), 113–124. https://doi.org/10.1038/nrg3366

Barbieri, C., Hübner, A., Macholdt, E., Ni, S., Lippold, S., Schröder, R., … Pakendorf, B. (2016). Refining the Y chromosome phylogeny with southern African sequences. *Human Genetics*, 135(5), 541–553. https://doi.org/10.1007/s00439-016-1651-0

Bellott, D. W., Hughes, J. F., Skaletsky, H., Brown, L. G., Pyntikova, T., Cho, T.-J., … Page, D. C. (2014). Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature*, 508(7497), 494–499. https://doi.org/10.1038/nature13206

Bensch, S., Andrén, H., Hansson, B., Pedersen, H. C., Sand, H., Sejberg, D., … Liberg, O. (2006). Selection for heterozygosity gives hope to a wild population of inbred wolves. *PLoS One*, 1(1), e72. https://doi.org/10.1371/journal.pone.0000072

Benson, J. F., Patterson, B. R., & Wheeldon, T. J. (2012). Spatial genetic and morphologic structure of wolves and coyotes in relation to environmental heterogeneity in a *Canis* hybrid zone. *Molecular Ecology*, 21(24), 5934–5954. https://doi.org/10.1111/mec.12045

Bhowmick, B. K., Satta, Y., & Takahata, N. (2007). The origin and evolution of human ampliconic gene families and ampliconic structure. *Genome Research*, 17(4), 441–450. https://doi.org/10.1101/gr.5734907

Bidon, T., Schreck, N., Hailer, F., Nilsson, M. A., & Janke, A. (2015). Genome-wide search identifies 1.9 Mb from the polar bear Y chromosome for evolutionary analyses. *Genome Biology and Evolution*, 7(7), 2010–2022. https://doi.org/10.1093/gbe/evv103

Botigué, L. R., Song, S., Scheu, A., Gopalan, S., Pendleton, A. L., Oetjens, M., … Veeramah, K. R. (2017). Ancient European dog genomes reveal continuity since the Early Neolithic. *Nature Communications*, 8, 16082. https://doi.org/10.1038/ncomms16082

Brown, S. K., Darwent, C. M., Wictum, E. J., & Sacks, B. N. (2015). Using multiple markers to elucidate the ancient, historical and modern relationships among North American Arctic dog breeds. *Heredity*, 115(6), 488–495. https://doi.org/10.1038/hdy.2015.49

Carmichael, L. E., Nagy, J. A., Larter, N. C., & Strobeck, C. (2002). Prey specialization may influence patterns of gene flow in wolves of the Canadian Northwest. *Molecular Ecology*, 10(12), 2787–2798. https://doi.org/10.1046/j.0962-1083.2001.01408.x

Carvalho, A. B., & Clark, A. G. (2013). Efficient identification of Y chromosome sequences in the human and Drosophila genomes. *Genome Research*, 23(11), 1894–1907. https://doi.org/10.1101/gr.156034.113

Chang, T.-C., Yang, Y., Retzel, E. F., & Liu, W.-S. (2013). Male-specific region of the bovine Y chromosome is gene rich with a high transcriptomic activity in testis development. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 12373–12378. https://doi.org/10.1073/pnas.1221104110

Charlesworth, B. (1978). Model for evolution of Y-chromosomes and dosage compensation. *Proceedings of the National Academy of Sciences of the United States of America*, 75(11), 5618–5622. https://doi.org/10.1073/pnas.75.11.5618

Charlesworth, B., & Charlesworth, D. (2000). The degeneration of Y chromosomes. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 355(1403), 1563–1572. https://doi.org/10.1098/rstb.2000.0717

Charlesworth, D. (2017). Evolution of recombination rates between sex chromosomes. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 372(1736), 20160456. https://doi.org/10.1098/rstb.2016.0456

Charlesworth, D., Charlesworth, B., & Marais, G. (2005). Steps in the evolution of heteromorphic sex chromosomes. *Heredity*, 95(2), 118–128. https://doi.org/10.1038/sj.hdy.6800697

Chen, N., Bellott, D. W., Page, D. C., & Clark, A. G. (2012). Identification of avian W-linked contigs by short-read sequencing. *BMC Genomics*, 13, 183. https://doi.org/10.1186/1471-2164-13-183

Cortez, D., Marin, R., Toledo-Flores, D., Froidevaux, L., Liechti, A., Waters, P. D., … Kaessmann, H. (2014). Origins and functional evolution of Y chromosomes across mammals. *Nature*, 508(7497), 488–493. https://doi.org/10.1038/nature13151

Darpö, J. (2016). The commission: A sheep in wolf's clothing?: On infringement proceedings as a legal device for the enforcement of EU law on the environment, using Swedish wolf management as an example. *Journal for European Environmental and Planning Law*, 13(3), 270–293. https://doi.org/10.1163/18760104-01303004

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, 9, 772. https://doi.org/10.1038/nmeth.2109

de Groot, G. A., Nowak, C., Skrbinšek, T., Andersen, L. W., Aspi, J., Fumagalli, L., … Muñoz-Fuentes, V. (2015). Decades of population genetic research reveal the need for harmonization of molecular markers: The grey wolf *Canis lupus* as a case study. *Mammal Review*, 46(1), 44–59. https://doi.org/10.1111/mam.12052

Ding, Z.-l., Oskarsson, M., Ardalan, A., Angleby, H., Dahlgren, L.-g., Tepeli, C., … Zhang, Y.-p. (2012). Origins of domestic dog in Southern East Asia is supported by analysis of Y-chromosome DNA. *Heredity*, 108(5), 507–514. https://doi.org/10.1038/hdy.2011.114

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29(8), 1969–1973. https://doi.org/10.1093/molbev/mss075

Ellegren, H. (2011). Sex-chromosome evolution: Recent progress and the influence of male and female heterogamety. *Nature Reviews Genetics*, 12(3), 157–166. https://doi.org/10.1038/nrg2948

Epstein, Y. (2016). Favourable conservation status for species: Examining the Habitats Directive's key concept through a case study of the Swedish wolf. *Journal of Environmental Law*, 28(2), 221–244. https://doi.org/10.1093/jel/eqw006

Ersmark, E., Klütsch, C. F. C., Chan, Y. L., Sinding, M.-H., Fain, S. R., Illarionova, N. A., … Savolainen, P. (2016). From the past to the present: Wolf phylogeography and demographic history based on the mitochondrial control region. *Frontiers in Ecology and Evolution*, 4, 134. https://doi.org/10.3389/fevo.2016.00134

Fabbri, E., Caniglia, R., Kusak, J., Galov, A., Gomerčić, T., Arbanasić, H., … Randi, E. (2014). Genetic structure of expanding wolf (*Canis lupus*) populations in Italy and Croatia, and the early steps of the recolonization of the Eastern Alps. *Mammalian Biology*, 79(2), 138–148. https://doi.org/10.1016/j.mambio.2013.10.002

Fan, Z., Silva, P., Gronau, I., Wang, S., Armero, A. S., Schweizer, R. M., … Wayne, R. K. (2016). Worldwide patterns of genomic variation and admixture in gray wolves. *Genome Research*, 26(2), 163–173. https://doi.org/10.1101/gr.197517.115

Flagstad, O., Walker, C. W., Vila, C., Sundqvist, A.-K., Fernholm, B., Hufthammer, A. K., … Ellegren, H. (2003). Two centuries of the Scandinavian wolf population: Patterns of genetic variability and migration during an era of dramatic decline. *Molecular Ecology*, 12(4), 869–880. https://doi.org/10.1046/j.1365-294X.2003.01784.x

Frantz, L. A. F., Mullin, V. E., Pionnier-Capitan, M., Lebrasseur, O., Ollivier, M., Perri, A., … Larson, G. (2016). Genomic and archaeological evidence suggest a dual origin of domestic dogs. *Science*, 352(6290), 1228–1231. https://doi.org/10.1126/science.aaf3161

Freedman, A. H., Gronau, I., Schweizer, R. M., Ortega-Del Vecchyo, D., Han, E., Silva, P. M., … Novembre, J. (2014). Genome sequencing highlights the dynamic early history of dogs. *PLoS Genetics*, 10(1), e1004016. https://doi.org/10.1371/journal.pgen.1004016

Geraldes, A., & Ferrand, N. (2006). A 7-bp insertion in the 3′ untranslated region suggests the duplication and concerted evolution of the rabbit SRY gene. *Genetics Selection Evolution*, 38(3), 313. https://doi.org/10.1186/1297-9686-38-3-313

Godinho, R., Llaneza, L., Blanco, J. c., Lopes, S., Álvares, F., García, E. J., … Ferrand, N. (2011). Genetic evidence for multiple events of hybridization between wolves and domestic dogs in the Iberian Peninsula. *Molecular Ecology*, 20(24), 5154–5166. https://doi.org/10.1111/j.1365-294X.2011.05345.x

Gómez-Sánchez, D., Olalde, I., Sastre, N., Enseñat, C., Carrasco, R., Marques-Bonet, T., … Ramírez, O. (2018). On the path to

extinction: Inbreeding and admixture in a declining gray wolf population. *Molecular Ecology*, 27(18), 3599–3612. https://doi.org/10.1111/mec.14824

Gopalakrishnan, S., Samaniego Castruita, J. A., Sinding, M.-H., Kuderna, L. F. K., Räikkönen, J., Petersen, B., … Gilbert, M. T. P. (2017). The wolf reference genome sequence (*Canis lupus lupus*) and its implications for *Canis* spp. population genomics. *BMC Genomics*, 18, 495. https://doi.org/10.1186/s12864-017-3883-3

Gottelli, D., Sillero-zubiri, C., Applebaum, G. D., Roy, M. S., Girman, D. J., Garcia-moreno, J., … Wayne, R. K. (1994). Molecular genetics of the most endangered canid: The Ethiopian wolf *Canis simensis*. *Molecular Ecology*, 3(4), 301–312. https://doi.org/10.1111/j.1365-294X.1994.tb00070.x

Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, 27(2), 221–224. https://doi.org/10.1093/molbev/msp259

Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G., & Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, 43, 1031–1034. https://doi.org/10.1038/ng.937

Hagenblad, J., Olsson, M., Parker, H. G., Ostrander, E. A., & Ellegren, H. (2009). Population genomics of the inbred Scandinavian wolf. *Molecular Ecology*, 18(7), 1341–1351. https://doi.org/10.1111/j.1365-294X.2009.04120.x

Hailer, F., & Leonard, J. A. (2008). Hybridization among three native North American *Canis* species in a region of natural sympatry. *PLoS One*, 3(10), e3333. https://doi.org/10.1371/journal.pone.0003333

Hall, A. B., Qi, Y., Timoshevskiy, V., Sharakhova, M. V., Sharakhov, I. V., & Tu, Z. (2013). Six novel Y chromosome genes in *Anopheles* mosquitoes discovered by independently sequencing males and females. *BMC Genomics*, 14, 273. https://doi.org/10.1186/1471-2164-14-273

Han, H., Zhao, X., Xia, X., Chen, H., Lei, C., & Dang, R. (2017). Copy number variations of five Y chromosome genes in donkeys. *Archives Animal Breeding*, 60, 391–397. https://doi.org/10.5194/aab-60-391-2017

Hindrikson, M., Mannil, P., Ozolins, J., Krzywinski, A., & Saarma, U. (2012). Bucking the trend in wolf–dog hybridization: First evidence from Europe of hybridization between female dogs and male wolves. *PLoS One*, 7(10), e46465. https://doi.org/10.1371/journal.pone.0046465

Hoeppner, M. P., Lundquist, A., Pirun, M., Meadows, J. R. S., Zamani, N., Johnson, J., … Grabherr, M. G. (2014). An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. *PLoS One*, 9(3), e91172. https://doi.org/10.1371/journal.pone.0091172

Hohenlohe, P. A., Rutledge, L. Y., Waits, L. P., Andrews, K. R., Adams, J. R., Hinton, J. W., … White, B. N. (2017). Comment on "Whole-genome sequence analysis shows two endemic species of North American wolf are admixtures of the coyote and gray wolf". *Science Advances*, 3(6), e1602250. https://doi.org/10.1126/sciadv.1602250

Hughes, J. F., Skaletsky, H., Brown, L. G., Pyntikova, T., Graves, T., Fulton, R. S., … Page, D. C. (2012). Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. *Nature*, 483(7387), 82–86. https://doi.org/10.1038/nature10843

Hughes, J. F., Skaletsky, H., Pyntikova, T., Graves, T. A., van Daalen, S. K. M., Minx, P. J., … Page, D. C. (2010). Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content. *Nature*, 463(7280), 536–539. https://doi.org/10.1038/nature08700

Immonen, E., & Husby, A. (2016). Norway wolf cull will hit genetic diversity. *Nature*, 539, 31. https://doi.org/10.1038/539031a

Jobling, M. A., Hollox, E., Hurles, M., Kivisild, T., & Tyler-Smith, C. (2014). *Human evolutionary genetics* (2nd ed.). New York, NY: Garland Science.

Jobling, M. A., & Tyler-Smith, C. (2003). The human Y chromosome: An evolutionary marker comes of age. *Nature Reviews Genetics*, 4(8), 598–612. https://doi.org/10.1038/nrg1124

Jobling, M. A., & Tyler-Smith, C. (2017). Human Y-chromosome variation in the genome-sequencing era. *Nature Reviews Genetics*, 18(8), 485–497. https://doi.org/10.1038/nrg.2017.36

Jones, E. P., & Searle, J. B. (2015). Differing Y chromosome versus mitochondrial DNA ancestry, phylogeography, and introgression in the house mouse. *Biological Journal of the Linnean Society*, 115(2), 348–361. https://doi.org/10.1111/bij.12522

Karafet, T. M., Mendez, F. L., Meilerman, M. B., Underhill, P. A., Zegura, S. L., & Hammer, M. F. (2008). New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Research*, 18(5), 830–838. https://doi.org/10.1101/gr.7172008

Kardos, M., Åkesson, M., Fountain, T., Flagstad, Ø., Liberg, O., Olason, P., … Ellegren, H. (2018). Genomic consequences of intensive inbreeding in an isolated wolf population. *Nature Ecology & Evolution*, 2(1), 124–131. https://doi.org/10.1038/s41559-017-0375-4

Karmin, M., Saag, L., Vicente, M., Sayres, M. A. W., Järve, M., Talas, U. G., … Kivisild, T. (2015). A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Research*, 25(4), 459–466. https://doi.org/10.1101/gr.186684.114

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, 12(4), 357–360. https://doi.org/10.1038/nmeth.3317

Koblmüller, S., Vilà, C., Lorente-Galdos, B., Dabad, M., Ramirez, O., Marques-Bonet, T., … Leonard, J. A. (2016). Whole mitochondrial genomes illuminate ancient intercontinental dispersals of grey wolves (*Canis lupus*). *Journal of Biogeography*, 43(9), 1728–1738. https://doi.org/10.1111/jbi.12765

Kojola, I., Aspi, J., Hakala, A., Heikkinen, S., Ilmoni, C., & Ronkainen, S. (2006). Dispersal in an expanding wolf population in Finland. *Journal of Mammalogy*, 87(2), 281–286. https://doi.org/10.1644/05-MAMM-A-061R2.1

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., … Marra, M. a. (2009). Circos: An information aesthetic for comparative genomics. *Genome Research*, 19(9), 1639–1645. https://doi.org/10.1101/gr.092759.109

Lahn, B. T., & Page, D. C. (1997). Functional coherence of the human Y chromosome. *Science*, 278(5338), 675–680. https://doi.org/10.1126/science.278.5338.675

Lahn, B. T., & Page, D. C. (1999). Four evolutionary strata on the human X chromosome. *Science*, 286(5441), 964–967. https://doi.org/10.1126/science.286.5441.964

Lahn, B. T., Pearson, N. M., & Jegalian, K. (2001). The human Y chromosome, in the light of evolution. *Nature Reviews Genetics*, 2(3), 207–216. https://doi.org/10.1038/35056058

Leigh, J. W., Bryant, D., & Nakagawa, S. (2015). POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6(9), 1110–1116. https://doi.org/10.1111/2041-210x.12410

Leonard, J. A., Vilà, C., & Wayne, R. K. (2005). Legacy lost: Genetic variability and population size of extirpated US grey wolves (*Canis lupus*). *Molecular Ecology*, 14(1), 9–17. https://doi.org/10.1111/j.1365-294X.2004.02389.x

Leonard, J. A., Wayne, R. K., Wheeler, J., Valadez, R., Guillen, S., & Vila, C. (2002). Ancient DNA evidence for Old World origin of New World dogs. *Science*, 298(5598), 1613–1616. https://doi.org/10.1126/science.1076980

Li, G., Davis, B. W., Raudsepp, T., Pearks Wilkerson, A. J., Mason, V. C., Ferguson-Smith, M., … Murphy, W. J. (2013). Comparative analysis of mammalian Y chromosomes illuminates ancestral structure and lineage-specific evolution. *Genome Research*, 23(9), 1486–1495. https://doi.org/10.1101/gr.154286.112

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Genome Project Data Processing, S (2009). The Sequence

Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liberg, O., Andren, H., Pedersen, H. C., Sand, H., Sejberg, D., Wabakken, P., … Bensch, S. (2005). Severe inbreeding depression in a wild wolf (*Canis lupus*) population. *Biology Letters*, *1*(1), 17–20. https://doi.org/10.1098/rsbl.2004.0266

Lindblad-Toh, K., Wade, C. M., Mikkelsen, T. S., Karlsson, E. K., Jaffe, D. B., Kamal, M., … Lander, E. S. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*, *438*(7069), 803–819. https://doi.org/10.1038/nature04338

Linnell, J. D. C., Brøseth, H., Solberg, E. J., & Brainerd, S. M. (2005). The origins of the southern Scandinavian wolf *Canis lupus* population: Potential for natural immigration in relation to dispersal distances, geography and Baltic ice. *Wildlife Biology*, *11*(4), 383–391. https://doi.org/10.2981/0909-6396(2005)11[383:TOOTSS]2.0.CO;2

Lippold, S., Xu, H., Ko, A., Li, M., Renaud, G., Butthof, A., … Stoneking, M. (2014). Human paternal and maternal demographic histories: Insights from high-resolution Y chromosome and mtDNA sequences. *Investigative Genetics*, *5*(1), 13. https://doi.org/10.1186/2041-2223-5-13

Matsumura, S., Inoshima, Y., & Ishiguro, N. (2014). Reconstructing the colonization history of lost wolf lineages by the analysis of the mitochondrial genome. *Molecular Phylogenetics and Evolution*, *80*, 105–112. https://doi.org/10.1016/j.ympev.2014.08.004

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., … DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Mech, L. D., Barber-Meyer, S. M., & Erb, J. (2016). Wolf (*Canis lupus*) generation time and proportion of current breeding females by age. *PLoS One*, *11*(6), e0156682. https://doi.org/10.1371/journal.pone.0156682

Muñoz-Fuentes, V., Darimont, C. T., Paquet, P. C., & Leonard, J. A. (2010). The genetic legacy of extirpation and re-colonization in Vancouver Island wolves. *Conservation Genetics*, *11*(2), 547–556. https://doi.org/10.1007/s10592-009-9974-1

Musiani, M., Leonard, J. a., Cluff, H. D., Gates, C. C., Mariani, S., Paquet, P. c., … Wayne, R. k. (2007). Differentiation of tundra/taiga and boreal coniferous forest wolves: Genetics, coat colour and association with migratory caribou. *Molecular Ecology*, *16*(19), 4149–4170. https://doi.org/10.1111/j.1365-294X.2007.03458.x

Nagamine, C. M., Shiroishi, T., Miyashita, N., Tsuchiya, K., Ikeda, H., Takao, N., … Kryukov, A. P. (1994). Distribution of the molossinus allele of Sry, the testis-determining gene, in wild mice. *Molecular Biology and Evolution*, *11*(6), 864–874.

Natanaelsson, C., Oskarsson, M. C., Angleby, H., Lundeberg, J., Kirkness, E., & Savolainen, P. (2006). Dog Y chromosomal DNA sequence: Identification, sequencing and SNP discovery. *BMC Genetics*, *7*, 45. https://doi.org/10.1186/1471-2156-7-45

Nei, M. (1970). Accumulation of nonfunctional genes on sheltered chromosomes. *American Naturalist*, *104*(938), 311–322. https://doi.org/10.1086/282665

Oetjens, M. T., Martin, A., Veeramah, K. R., & Kidd, J. M. (2018). Analysis of the canid Y-chromosome phylogeny using short-read sequencing data reveals the presence of distinct haplogroups among Neolithic European dogs. *BMC Genomics*, *19*, 350. https://doi.org/10.1186/s12864-018-4749-z

Oetjens, M. T., Shen, F. C., Emery, S. B., Zou, Z. T., & Kidd, J. M. (2016). Y-Chromosome structural diversity in the bonobo and chimpanzee lineages. *Genome Biology and Evolution*, *8*(7), 2231–2240. https://doi.org/10.1093/gbe/evw150

Pang, J.-F., Kluetsch, C., Zou, X.-J., Zhang, A.-B., Luo, L.-Y., Angleby, H., … Savolainen, P. (2009). mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. *Molecular Biology and Evolution*, *26*(12), 2849–2864. https://doi.org/10.1093/molbev/msp195

Pearks Wilkerson, A. J., Raudsepp, T., Graves, T., Albracht, D., Warren, W., Chowdhary, B. P., … Murphy, W. J. (2008). Gene discovery and comparative analysis of X-degenerate genes from the domestic cat Y chromosome. *Genomics*, *92*(5), 329–338. https://doi.org/10.1016/j.ygeno.2008.06.012

Pilot, M., Branicki, W., Jędrzejewski, W., Goszczyński, J., Jędrzejewska, B., Dykyy, I., … Tsingarska, E. (2010). Phylogeographic history of grey wolves in Europe. *BMC Evolutionary Biology*, *10*, 104. https://doi.org/10.1186/1471-2148-10-104

Pilot, M., Greco, C., vonHoldt, B. M., Jędrzejewska, B., Randi, E., Jędrzejewski, W., … Wayne, R. K. (2013). Genome-wide signatures of population bottlenecks and diversifying selection in European wolves. *Heredity*, *112*, 428–442. https://doi.org/10.1038/hdy.2013.122

Pilot, M., Greco, C., vonHoldt, B. M., Randi, E., Jędrzejewski, W., Sidorovich, V. E., … Wayne, R. K. (2018). Widespread, long-term admixture between grey wolves and domestic dogs across Eurasia and its implications for the conservation status of hybrids. *Evolutionary Applications*, *11*(5), 662–680. https://doi.org/10.1111/eva.12595

Pilot, M., Jedrzejewski, W., Branicki, W., Sidorovich, V. E., Jedrzejewska, B., Stachura, K., & Funk, S. M. (2006). Ecological factors influence population genetic structure of European grey wolves. *Molecular Ecology*, *15*(14), 4533–4553. https://doi.org/10.1111/j.1365-294X.2006.03110.x

Pires, A. E., Amorim, I. R., Borges, C., Simões, F., Teixeira, T., Quaresma, A., … Matos, J. (2017). New insights into the genetic composition and phylogenetic relationship of wolves and dogs in the Iberian Peninsula. *Ecology and Evolution*, *7*(12), 4404–4418. https://doi.org/10.1002/ece3.2949

Ponnikas, S., Sigeman, H., Abbott, J. K., & Hansson, B. (2018). Why do sex chromosomes stop recombining? *Trends in Genetics*, *34*(7), 492–503. https://doi.org/10.1016/j.tig.2018.04.001

Poznik, G. D., Henn, B. M., Yee, M.-C., Sliwerska, E., Euskirchen, G. M., Lin, A. A., … Bustamante, C. D. (2013). Sequencing Y chromosomes resolves discrepancy in time to common ancestor of males versus females. *Science*, *341*(6145), 562–565. https://doi.org/10.1126/science.1237619

Poznik, G. D., Xue, Y. L., Mendez, F. L., Willems, T. F., Massaia, A., & Sayres, M. A. W., … Genomes Project, C (2016). Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. *Nature Genetics*, *48*(6), 593–599. https://doi.org/10.1038/ng.3559

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Randi, E., Hulva, P., Fabbri, E., Galaverni, M., Galov, A., Kusak, J., … Caniglia, R. (2014). Multilocus detection of wolf × dog hybridization in Italy, and guidelines for marker selection. *PLoS One*, *9*(1), e86409. https://doi.org/10.1371/journal.pone.0086409

Randi, E., Lucchini, V., Christensen, M. F., Mucci, N., Funk, S. M., Dolf, G., & Loeschcke, V. (2000). Mitochondrial DNA variability in Italian and East European wolves: Detecting the consequences of small population size and hybridization. *Conservation Biology*, *14*(2), 464–473. https://doi.org/10.1046/j.1523-1739.2000.98280.x

Repping, S., van Daalen, S. K. M., Brown, L. G., Korver, C. M., Lange, J., Marszalek, J. D., … Rozen, S. (2006). High mutation rates have driven extensive structural polymorphism among human Y chromosomes. *Nature Genetics*, *38*, 463. https://doi.org/10.1038/ng1754

Rice, W. R. (1994). Degeneration of a nonrecombining chromosome. *Science*, *263*(5144), 230–232. https://doi.org/10.1126/science.8284674

Rosenberg, N. A., & Feldman, M. W. (2002). The relationship between coalescence times and population divergence times. In M. Slatkin &

M. Veuille (Eds.), *Modern developments in theoretical population genetics: The legacy of Gustave Malecot* (pp. 130–164). Oxford, UK: Oxford University Press.

Ross, M. T., Grafham, D. V., Coffey, A. J., Scherer, S., McLay, K., Muzny, D., ... Bentley, D. R. (2005). The DNA sequence of the human X chromosome. *Nature*, *434*(7031), 325–337. https://doi.org/10.1038/nature03440

Rozen, S., Skaletsky, H., Marszalek, J. D., Minx, P. J., Cordum, H. S., Waterston, R. H., ... Page, D. C. (2003). Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature*, *423*(6942), 873–876. https://doi.org/10.1038/nature01723

Sacks, B. N., Brown, S. K., Stephens, D., Pedersen, N. C., Wu, J. T., & Berry, O. (2013). Y chromosome analysis of dingoes and Southeast Asian village dogs suggests a Neolithic continental expansion from Southeast Asia followed by multiple Austronesian dispersals. *Molecular Biology and Evolution*, *30*(5), 1103–1118. https://doi.org/10.1093/molbev/mst027

Sandstedt, S. A., & Tucker, P. K. (2004). Evolutionary strata on the mouse X chromosome correspond to strata on the human X chromosome. *Genome Research*, *14*(2), 267–272. https://doi.org/10.1101/gr.1796204

Sastre, N., Vilà, C., Salinas, M., Bologov, V. V., Urios, V., Sánchez, A., ... Ramírez, O. (2011). Signatures of demographic bottlenecks in European wolf populations. *Conservation Genetics*, *12*(3), 701–712. https://doi.org/10.1007/s10592-010-0177-6

Savolainen, P., Zhang, Y. P., Luo, J., Lundeberg, J., & Leitner, T. (2002). Genetic evidence for an East Asian origin of domestic dogs. *Science*, *298*(5598), 1610–1613. https://doi.org/10.1126/science.1073906

Schweizer, R. M., Robinson, J., Harrigan, R., Silva, P., Galaverni, M., Musiani, M., ... Wayne, R. K. (2016a). Targeted capture and resequencing of 1040 genes reveal environmentally driven functional variation in gray wolves. *Molecular Ecology*, *25*, 357–379. https://doi.org/10.1111/mec.13467

Schweizer, R. M., vonHoldt, B., Harrigan, R., Knowles, J., Musiani, M., Coltman, D., ... Wayne, R. K. (2016b). Genetic subdivision and candidate genes under selection in North American gray wolves. *Molecular Ecology*, *25*, 380–402. https://doi.org/10.1111/mec.13364

Seddon, J. M., Sundqvist, A. K., Björnerfeldt, S., & Ellegren, H. (2006). Genetic identification of immigrants to the Scandinavian wolf population. *Conservation Genetics*, *7*(2), 225–230. https://doi.org/10.1007/s10592-005-9001-0

Semino, O., Passarino, G., Oefner, P. J., Lin, A. A., Arbuzova, S., Beckman, L. E., ... Underhill, P. A. (2000). The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: A Y chromosome perspective. *Science*, *290*(5494), 1155–1159. https://doi.org/10.1126/science.290.5494.1155

Shannon, L. M., Boyko, R. H., Castelhano, M., Corey, E., Hayward, J. J., McLean, C., ... Boyko, A. R. (2015). Genetic structure in village dogs reveals a Central Asian domestication origin. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(44), 13639–13644. https://doi.org/10.1073/pnas.1516215112

Skaletsky, H., Kuroda-Kawaguchi, T., Minx, P. J., Cordum, H. S., Hillier, L. D., Brown, L. G., ... Page, D. C. (2003). The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature*, *423*(6942), 825–837. https://doi.org/10.1038/nature01722

Skinner, B. M., Sargent, C. A., Churcher, C., Hunt, T., Herrero, J., Loveland, J. E., ... Tyler-Smith, C. (2016). The pig X and Y chromosomes: Structure, sequence, and evolution. *Genome Research*, *26*(1), 130–139. https://doi.org/10.1101/gr.188839.114

Skoglund, P., Ersmark, E., Palkopoulou, E., & Dalen, L. (2015). Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. *Current Biology*, *25*(11), 1515–1519. https://doi.org/10.1016/j.cub.2015.04.019

Smeds, L., Warmuth, V., Bolivar, P., Uebbing, S., Burri, R., Suh, A., ... Ellegren, H. (2015). Evolutionary analysis of the female-specific avian W chromosome. *Nature Communications*, *6*, 7330. https://doi.org/10.1038/ncomms8330

Soh, Y. Q. S., Alföldi, J., Pyntikova, T., Brown, L. G., Graves, T., Minx, P. J., ... Page, D. C. (2014). Sequencing the mouse Y chromosome reveals convergent gene acquisition and amplification on both sex chromosomes. *Cell*, *159*(4), 800–813. https://doi.org/10.1016/j.cell.2014.09.052

Sundqvist, A. K., Ellegren, H., Olivier, M., & Vilà, C. (2001). Y chromosome haplotyping in Scandinavian wolves (*Canis lupus*) based on microsatellite markers. *Molecular Ecology*, *10*(8), 1959–1966. https://doi.org/10.1046/j.1365-294X.2001.01326.x

Thalmann, O., Shapiro, B., Cui, P., Schuenemann, V. J., Sawyer, S. K., Greenfield, D. L., ... Wayne, R. K. (2013). Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. *Science*, *342*(6160), 871–874. https://doi.org/10.1126/science.1243650

Thorvaldsdottir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, *14*(2), 178–192. https://doi.org/10.1093/bib/bbs017

Tomaszkiewicz, M., Medvedev, P., & Makova, K. D. (2017). Y and W chromosome assemblies: Approaches and discoveries. *Trends in Genetics*, *33*(4), 266–282. https://doi.org/10.1016/j.tig.2017.01.008

Tomaszkiewicz, M., Rangavittal, S., Cechova, M., Sanchez, R. C., Fescemyer, H. W., Harris, R., ... Makova, K. D. (2016). A time- and cost-effective strategy to sequence mammalian Y chromosomes: An application to the de novo assembly of gorilla Y. *Genome Research*, *26*(4), 530–540. https://doi.org/10.1101/gr.199448.115

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., ... Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, *28*(5), 511–515. https://doi.org/10.1038/nbt.1621

Underhill, P. A., & Kivisild, T. (2007). Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annual Review of Genetics*, *41*, 539–564. https://doi.org/10.1146/annurev.genet.41.110306.130407

Underhill, P. A., Shen, P., Lin, A. A., Jin, L. i., Passarino, G., Yang, W. H., ... Oefner, P. J. (2000). Y chromosome sequence variation and the history of human populations. *Nature Genetics*, *26*(3), 358–361. https://doi.org/10.1038/81685

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., ... DePristo, M. A. (2013). From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*, *43*, 11.10.1–33. https://doi.org/10.1002/0471250953.bi1110s43

Vila, C., Amorim, I. R., Leonard, J. A., Posada, D., Castroviejo, J., Petrucci-Fonseca, F., ... Wayne, R. K. (1999). Mitochondrial DNA phylogeography and population history of the grey wolf *Canis lupus*. *Molecular Ecology*, *8*(12), 2089–2103. https://doi.org/10.1046/j.1365-294x.1999.00825.x

Vilà, C., Savolainen, P., Maldonado, J. E., Amorim, I. R., Rice, J. E., Honeycutt, R. L., ... Wayne, R. K. (1997). Multiple and ancient origins of the domestic dog. *Science*, *276*(5319), 1687–1689. https://doi.org/10.1126/science.276.5319.1687

Vilà, C., Sundqvist, A. K., Flagstad, Ø., Seddon, J., Bjornerfeldt, S., Kojola, I., ... Ellegren, H. (2003). Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. *Proceedings of the Royal Society B-Biological Sciences*, *270*(1510), 91–97. https://doi.org/10.1098/rspb.2002.2184

Vilà, C., Walker, C., Sundqvist, A.-k., Flagstad, Ø., Andersone, Z., Casulli, A., ... Ellegren, H. (2003). Combined use of maternal, paternal and bi-parental genetic markers for the identification of wolf-dog hybrids. *Heredity*, *90*(1), 17–24. https://doi.org/10.1038/sj.hdy.6800175

vonHoldt, B. M., Cahill, J. A., Fan, Z., Gronau, I., Robinson, J., Pollinger, J. P., … Wayne, R. K. (2016). Whole-genome sequence analysis shows that two endemic species of North American wolf are admixtures of the coyote and gray wolf. *Science Advances*, *2*(7), e1501714. https://doi.org/10.1126/sciadv.1501714

vonHoldt, B. M., Pollinger, J. P., Earl, D. A., Knowles, J. C., Boyko, A. R., Parker, H., … Wayne, R. K. (2011). A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. *Genome Research*, *21*(8), 1294–1305. https://doi.org/10.1101/gr.116301.110

vonHoldt, B. M., Pollinger, J. P., Lohmueller, K. E., Han, E., Parker, H. G., Quignon, P., … Wayne, R. K. (2010). Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature*, *464*, 898–902. https://doi.org/10.1038/nature08837

vonHoldt, B. M., Stahler, D. R., Smith, D. W., Earl, D. A., Pollinger, J. P., & Wayne, R. K. (2008). The genealogy and genetic viability of reintroduced Yellowstone grey wolves. *Molecular Ecology*, *17*(1), 252–274. https://doi.org/10.1111/j.1365-294X.2007.03468.x

Wabakken, P., Sand, H., Liberg, O., & Bjärvall, A. (2001). The recovery, distribution, and population dynamics of wolves on the Scandinavian peninsula, 1978–1998. *Canadian Journal of Zoology*, *79*(4), 710–725. https://doi.org/10.1139/z01-029

Wayne, R. K. (1993). Molecular evolution of the dog family. *Trends in Genetics*, *9*(6), 218–224. https://doi.org/10.1016/0168-9525(93)90122-X

Wayne, R. K., & Jenks, S. M. (1991). Mitochondrial DNA analysis implying extensive hybridization of the endangered red wolf *Canis rufus*. *Nature*, *351*, 565–568. https://doi.org/10.1038/351565a0

Wayne, R. K., Lehman, N., Allard, M. W., & Honeycutt, R. L. (1992). Mitochondrial DNA variability of the gray wolf: Genetic consequences of population decline and habitat fragmentation. *Conservation Biology*, *6*(4), 559–569. https://doi.org/10.1046/j.1523-1739.1992.06040559.x

Wheeldon, T. J., Rutledge, L. Y., Patterson, B. R., White, B. N., & Wilson, P. J. (2013). Y-chromosome evidence supports asymmetric dog introgression into eastern coyotes. *Ecology and Evolution*, *3*(9), 3005–3020. https://doi.org/10.1002/ece3.693

Wilson, P. J., Rutledge, L. Y., Wheeldon, T. J., Patterson, B. R., & White, B. N. (2012). Y-chromosome evidence supports widespread signatures of three-species *Canis* hybridization in eastern North America. *Ecology and Evolution*, *2*(9), 2325–2332. https://doi.org/10.1002/ece3.301

Wilson Sayres, M. A., Venditti, C., Pagel, M., & Makova, K. D. (2011). Do variations in substitution rates and male mutation bias correlate with life-history traits? A study of 32 mammalian genomes. *Evolution*, *65*(10), 2800–2815. https://doi.org/10.1111/j.1558-5646.2011.01337.x

Zhang, W., Fan, Z., Han, E., Hou, R., Zhang, L., Galaverni, M., … Zhang, Z. (2014). Hypoxia adaptations in the grey wolf (*Canis lupus chanco*) from Qinghai-Tibet plateau. *PLoS Genetics*, *10*(7), e1004466. https://doi.org/10.1371/journal.pgen.1004466

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.