Genomic consequences of intensive inbreeding in an isolated wolf population

Marty Kardos[1,2], Mikael Åkesson[3], Toby Fountain[1], Øystein Flagstad[4], Olof Liberg[3], Pall Olason[5], Håkan Sand[3], Petter Wabakken[6], Camilla Wikenros[3], and Hans Ellegren[1,*]

1. Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden

2. Flathead Lake Biological Station, University of Montana, Polson, Montana 59860

3. Department of Ecology, Grimsö Wildlife Research Station, Swedish University of Agricultural Sciences, SE-730 91 Riddarhyttan, Sweden.

4. Norwegian Institute for Nature Research, P.O. Box 5685 Sluppen, Trondheim, NO-7485, Norway.

5. Wallenberg Advanced Bioinformatics Infrastructure (WABI), Science for Life Laboratory, Uppsala University, 75123 Uppsala, Sweden

6. Faculty of Applied Ecology and Agricultural Sciences, Campus Evenstad, Inland Norway University of Applied Sciences, NO-2480, Norway

*Corresponding author: Hans.Ellegren@ebc.uu.se

Kardos, Marty; Åkesson, Mikael; Fountain, Toby; Flagstad, Øystein; Liberg, Olof; Olason, Pall; Sand, Håkan; Wabakken, Petter; Wikenros, Camilla; Ellegren, Hans.
Genomic consequences of intensive inbreeding in an isolated wolf population. Nature Ecology & Evolution 2017
DOI:10.1038/s41559-017-0375-4

Abstract

Inbreeding (mating between relatives) is a major concern for conservation as it decreases the fitness of offspring and can increase the extinction risk of populations. While pedigrees have traditionally been used to measure individual inbreeding, molecular markers have opened up new avenues to characterize inbreeding. However, a limitation has been that small numbers of markers can only roughly measure the proportion of an individual's genome that is identical-by-descent (IBD) due to inbreeding. We used whole-genome resequencing of 97 grey wolves (*Canis lupus*) from the highly inbred Scandinavian wolf population, originally founded by only two individuals, to identify IBD chromosome segments as runs of homozygosity (ROH). This gave the very high resolution required to precisely measure the realized IBD fraction of the genome as F_{ROH}. We found a striking pattern of complete or near-complete homozygosity of entire chromosomes in many individuals. The majority of individual IBD was contributed by long IBD segments (>5cM) originating from common ancestors of parents within the last ~10 generations. However, although most IBD segments were very short (<0.02 cM) and originate from ancestors in deep history, they contributed little to the total amount of individual IBD. Individual inbreeding estimated with an extensive pedigree (F_P) was strongly correlated with realized inbreeding measured with the entire genome (r^2 = 0.86). However, inbreeding measured with the whole genome was more strongly correlated with multi-locus heterozygosity estimated with as few as 500 SNPs, and with F_{ROH} estimated with as few as 10,000 SNPs, than with F_P. Some immigrants were inbred, and two substantially so and also related,, which is counter to the assumptions of unrelated and non-inbred founders and immigrants in pedigree analysis of inbreeding. These results document in unique detail the genomic consequences of intensive inbreeding in a population of conservation concern.

Small populations are particularly vulnerable to extinction due to demographic stochasticity, reduced genetic variation, and inbreeding depression[1-4]. Inbreeding (mating between relatives) in small populations can lead to decreased individual fitness and population growth rate, owing to the expression of deleterious recessive alleles and increased homozygosity at loci with heterozygous advantage[3,5]. While inbreeding depression has long interested biologists, its strength and genetic basis in the wild are still not well understood[6,7]. A major challenge has been accurately measuring individual inbreeding in natural populations.

Individual inbreeding has classically been estimated with the pedigree inbreeding coefficient (F_P) for an individual using path analysis on a known pedigree[3,8,9]. F_P predicts F, the fraction of an individual's genome that is identical-by-descent (IBD), assuming that the pedigree founders, and any subsequent immigrants are non-inbred and unrelated. However, not only are the necessary multi-generation pedigrees difficult to obtain for most natural populations[10,11], but F_P often imprecisely measures F because of the stochastic effects of Mendelian segregation and linkage[7,12-17].

An alternative approach is to measure individual inbreeding indirectly by using genetic markers to estimate multi-locus heterozygosity (*MLH*)[18-21], as the major effect of inbreeding is to reduce the genome-wide heterozygosity of the offspring[5]. This reduction occurs because related parents pass on IBD chromosome segments that arise from a single chromosome copy in a shared ancestor, with these segments characterized by long stretches of homozygous genotypes (i.e., runs of homozygosity, ROH)[7]. *MLH* and similar statistics have the advantage of not requiring a pedigree, but suffer from low precision when using few loci[21-24].

High-throughput sequencing technologies can make it possible to measure genome-wide heterozygosity using thousands of genetic markers[25-28]. Importantly, whole-genome resequencing in species with high quality genome assemblies should facilitate the identification of IBD chromosome segments as

ROH, allowing the measurement of F as the fraction of the genome in long ROH (F_{ROH}) with very little error[29]. Additionally, whole-genome resequencing of many individuals from natural populations where high quality pedigrees are available would allow rigorous empirical evaluation of how well F_P, MLH and F_{ROH} based on smaller number of loci perform as estimators of F.

Here, we resequenced 97 genomes sampled from a semi-isolated and bottlenecked wolf population in Scandinavia, This population is of high conservation concern and has been subject to long-term studies of inbreeding, inbreeding depression and genetic rescue[30-34]. Importantly, the population represents a rare example of having a nearly complete pedigree available[30]. First, we sought to identify IBD chromosome segments and to quantify F among individuals in the population. Second, we evaluated the statistical performance of F_P, MLH, and F_{ROH} as measures of F. Finally, we searched for regions of the genome that may harbor alleles with large phenotypic effects contributing to inbreeding depression by scanning for chromosome segments where ROH were exceptionally rare or absent. To our knowledge, this is the first study combining whole-genome resequencing and pedigree information to study individual inbreeding in the wild.

Results

Study population, pedigree, and whole-genome resequencing

After a long period of population decline, wolves became functionally extinct from the Scandinavian Peninsula in 1960-1970s[35]. The contemporary Scandinavian wolf population was founded by two individuals in the early 1980s[33,36], and is characterized by prolonged periods of isolation with only rare reproductively successful immigrants[30,37]. We sampled 97 wolves from Scandinavia between 1977 and 2015, including 12 immigrants of which five were founders of the population. These individuals were chosen to represent the range of observed F_P values in the population, which were derived from a

pedigree extending back to the first breeding event in 1983[34]. F_P ranged from zero (for 12 immigrants and 19 Scandinavian-born offspring to immigrant founders) up to $F_P = 0.49$ for three Scandinavian-born siblings sampled after the population had experienced a prolonged period of isolation. The number of generations of pedigree known for each individual is given in Table S1.

We performed whole-genome resequencing of all wolves at a mean sequence read depth of 27.4 (s.d. = 10.3). After variant calling, we performed SNP filtering based on genotype qualities, read depth, deviation from Hardy-Weinberg genotype proportions, missing data, and minor allele frequency (MAF; see Methods). Mean MAF was 0.17 at 10,688,886 SNPs remaining before filtering based on allele frequency. After filtering based on allele frequency, the mean MAF was 0.26 (s.d. = 0.13) at 6,701,147 SNPs. Given that almost one hundred individuals were sequenced, the number of detected variants is low for a large mammalian genome. However, low genetic diversity is expected given the small population size and limited number of founders. Moreover, nucleotide diversity estimated from the 12 immigrants was 0.001, which is in the lower end of what has been reported among other vertebrates.

Runs of homozygosity (ROH)

We identified ROH (putative IBD chromosome segments) in the whole-genome resequencing data using a likelihood ratio-based sliding window method which accounts for SNP allele frequencies and sequencing errors[29,38]. We detected a total of 269,309 ROH among the 97 wolves, ranging from 0 to 76.6 cM in genetic map length, and from 2,695 bp to 95.8 Mb in physical length (Figure 1, Figure S1. Describing ROH by genetic map length is motivated by the fact that recombination determines the size of IBD segments. Additionally, our theoretical understanding of the expected lengths of ROH, and of the variance of F around pedigree expectations is in terms of ROH genetic map lengths[12,17,39]. The choice of using physical versus genetic mapping coordinates of ROH had nearly no effect on genomic estimates of inbreeding (Figure S2). Notably, many individuals had ROH spanning either entire or nearly entire chromosomes, giving extreme

patterns with a complete lack of heterozygosity over large parts of the genome (Figures 2-3, Supplementary File 1).

Though there were many strikingly large ROH (Figures 2-3), most were very short. Specifically, more than 50% of ROH were less than 0.02 cM long (Figure 1) and these represent IBD segments that generally arise from ancestors in deep history. We estimated the number of generations (g) back to the common ancestor of the two homologous sequence copies for each ROH based on its map length. The very short ROH (≤0.02 cM long) are expected to arise an average from ancestors ≥2,500 generations ago (i.e., g = 2,500 for a 0.02 cM ROH; see Methods); 2,500 generations correspond to 10,000 years assuming a four-year generation interval for wolves. Yet, the highly abundant, short ROH contributed little to total IBD. For example, segments shorter than 0.02 cM represented only 1.3% of all IBD chromosome regions in the 97 wolves (Figure 1, Supplementary File 1). In contrast, the less frequent but very long ROH arising from recent ancestors accounted for the majority of total IBD sequence.

Genomic measures of inbreeding

We measured individual inbreeding as the proportion of the genome that was in ROH (F_{ROH}) identified in the whole-genome resequencing data. F_{ROH} is an estimator of the realized IBD fraction of the genome and was obtained using only long ROH (i.e., ROH with small g-values). We conducted separate analyses using different maximum values of g (10, 25, 50, and 100 generations) for the ROH included in estimates of F_{ROH}. This ensured that we measured inbreeding due to recent ancestors while also allowing us to evaluate the sensitivity of the results to different maximum values of g. Including very short ROH would have meant that F_{ROH} captured inbreeding due to distant ancestors, which is less likely to be important to inbreeding depression because at least some deleterious alleles are expected to be purged over long time spans[38,40].

There was a large range of F_{ROH} in the population. F_{ROH} measured using ROH with $g \leq 10$ ranged from 0.01 to 0.54 (mean = 0.27, σ^2 =0.02) among Scandinavian-

195 born wolves (Figure 4, Figure S3). Unexpectedly, F_{ROH} of immigrants ranged from
196 0.01 up to 0.15 (mean = 0.045, σ^2 = 0.022) (Figure S3). This demonstrates that
197 some immigrants had relatively high inbreeding (the expected *F* of offspring
198 from half-sib mating is 0.125). For example, two immigrants that appeared in
199 northern Sweden in 2013 and were translocated by management authorities to
200 the wolf breeding range in southern Sweden were both inbred (F_{ROH} = 0.10 and
201 0.15, respectively). These translocated immigrants bred with each other the
202 same year and were clearly closely related since two of their offspring that were
203 sequenced had F_{ROH} = 0.26 and 0.24, respectively (suggesting that their parents
204 were related at approximately the level of full siblings). Excluding these two
205 related individuals, mean F_{ROH} of immigrants was 0.029 (σ^2 = 0.028). Emigration
206 from a small peripheral wolf population in Russia or Finland may explain the
207 non-zero inbreeding of immigrants into Scandinavia.
208
209 The non-zero F_{ROH} of immigrants is counter to the assumptions of unrelated and
210 non-inbred founders and immigrants in standard pedigree analysis of
211 inbreeding. Related pedigree founders mean that F_P fails to capture all of the
212 inbreeding that is due to recent common ancestors of parents not included in the
213 pedigree. Having inbred founders also means that F_P fails to capture inbreeding
214 due to IBD segments in the founders.
215
216 We used *MLH* as a second genomic measure of individual inbreeding. *MLH*
217 estimates the realized fraction of heterozygous SNPs across the genome (*H*), and
218 is related to *F* according to the expression $H = H_0(1-F)$, where H_0 is the genome-
219 wide heterozygosity of a hypothetical non-inbred individual[7,41]. *MLH* was
220 strongly correlated with F_{ROH} (r^2 = 0.91) (Figure S4). A perfect correlation
221 between F_{ROH} and *MLH* is not expected because F_{ROH} accounts only for IBD
222 segments that are detected; the very shortest IBD segments arising from
223 ancestors in deep history are likely to go undetected because they contain too
224 few SNPs to reliably differentiate from non-IBD[29]. Unlike F_{ROH}, *MLH* captures
225 variation in *F* due to all IBD segments, arising from recent ancestors as well as
226 the most distant ancestors.
227

Performance of F_P and molecular measures of individual inbreeding

We used linear regression to evaluate the statistical performance of F_P, F_{ROH}, and *MLH* as predictors of realized individual inbreeding. F_{ROH} measured with the whole genome is equivalent to *F*, and the same applies to *MLH* with respect to *H*. F_P was strongly correlated (r^2 = 0.86-0.87) with F_{ROH} (Figure 4). The linear regression of F_P versus F_{ROH} had a slope of 1.0 and an intercept of –0.03 when F_{ROH} was measured with only the longest ROH ($g \leq 10$). The negative intercept shows that F_P was a downwardly biased measure of F_{ROH}, and the slope of 1.0 shows that the size of the downward bias was constant on average across the range of observed F_{ROH} values (Figure 4). The correlations between F_P and F_{ROH} were only slightly weaker (r^2 = 0.83 to 0.84), and the slopes and intercepts were unchanged, when immigrants were excluded from this analysis (Figure S5). The choice of a maximum value of *g* for the ROH included in the measurement of F_{ROH} did not substantively affect the correlation between F_P and F_{ROH}, but the magnitude of the downward bias in F_P increased with higher values of threshold of *g* (Figure 4). This makes sense as F_{ROH} calculated using ROH with larger values of *g* captures inbreeding due to more distant ancestors.

The high variation in F_{ROH} among individuals with F_P = 0 weakened the precision of F_P. Specifically, a combination of some highly inbred individuals and individuals with F_{ROH} near zero clearly decreased the variance in F_{ROH} explained by F_P (Figure 4). F_P is likely to have higher precision in populations with less variation in F_{ROH} among founders and immigrants. An obvious strength of genomic measures of individual inbreeding is that they do not require making *a priori* assumptions regarding the inbreeding or relatedness of any individuals.

The relatively high precision of F_P as a measure of individual inbreeding observed here (compared to previous simulation results[27]) is expected. Theoretical and simulation-based investigations have shown that the precision of F_P as a measure of *F* depends strongly on the number of chromosomes, the recombination rate, and the distribution of recombination events across the genome [5,12,14,16,39]. Canids have a large number of chromosomes (38 autosomes).

Thus, F_P is expected to be more precise in wolves compared to species with fewer chromosomes, as long as pedigrees are deep and complete enough to capture the great majority of recent common ancestors of parents. The high variance in individual inbreeding in this study also must have contributed to the high r^2 from a regression of F_P versus F_{ROH}. We sampled from throughout the range of F_P values observed in the population, which resulted in a higher variance in F_P among the selected wolves ($\sigma^2 = 0.026$) relative to the population as a whole ($\sigma^2 = 0.006$). This is expected to have increased the correlation of realized genomic inbreeding with F_P and the molecular inbreeding measures based on subsampled SNPs in the sampled wolves compared to the population as a whole. All else equal, a lower correlation of F with F_P and molecular measures of inbreeding is expected in populations with lower variance in F[24,42].

Performance of MLH as a measure of individual inbreeding

To evaluate the precision of *MLH* as a measure of *H*, we randomly subsampled between 50 and 20,000 SNPs from the genome. For each subsampled set of loci, a linear regression model with *MLH* measured from the subsampled loci was fitted as the response variable, and *MLH* measured with the whole genome as the predictor variable. We then used r^2 from these regression models as a measure of the precision of *MLH*. To ensure that the subsamples were drawn as independently as possible from the genome, no locus was used in more than one of the 100 subsamples for each number of loci analyzed.

The mean r^2 between *MLH* based on subsampled loci and *MLH* from the whole genome was 0.88 when 500 SNPs were used, and ≥ 0.94 when 1,000 or more SNPs were used (Figure 5). *MLH* and other measures of individual inbreeding are expected to have high precision when the variance in F is as high as it was in this study[22]. The correlation between *MLH* based on subsampled loci and *MLH* measured with the whole genome matches theoretical expectations remarkably well. For example, the expected correlation between *MLH* (estimated with 500 loci) and realized genome-wide heterozygosity is 0.87 according to the analytical results of Miller et al.[22], very close to the observed r^2 of 0.88. This is highly

encouraging for studies of natural populations where pedigrees, mapped loci, and large-scale SNP genotyping arrays or whole-genome resequencing data are unavailable. This is also empirical evidence that individual inbreeding can be more precisely measured with a modest number of molecular markers than with pedigrees[14,27].

Performance of F_{ROH} as a measure of individual inbreeding

We used the same subsampling and regression approach applied above for *MLH* to evaluate the performance of F_{ROH}. However, for F_{ROH}, we used subsamples of 10,000 SNPs and larger, and the predictor variable in the regression models was F_{ROH} measured with the whole genome. F_{ROH} estimated with as few as 10,000 SNPs was strongly correlated with F_{ROH} estimated with the whole genome (mean $r^2 = 0.97$ [s.d. = 0.003] among 100 replicates, Figure S6). F_{ROH} estimated with subsampled SNPs was slightly upwardly biased (Figure S7). This bias was likely caused by overestimating the length of real IBD segments, or by incorrectly calling an ROH where no true IBD segment existed when using relatively few loci. We therefore urge caution when interpreting results of ROH analyses (e.g., for estimating individual inbreeding or mapping loci responsible for inbreeding depression) when only tens of thousands of loci are used.

Detecting genomic regions that may contribute to inbreeding depression

Alleles that strongly reduce fitness when homozygous (i.e., either strongly deleterious recessive or overdominant alleles) are likely to cause ROH to be absent or exceptionally rare in the local chromosomal vicinity [7,43,44]. We quantified the abundance of ROH with values of $g \leq 50$ in non-overlapping 100 kb windows across all 38 autosomes and used a permutation approach to test for regions with lower than expected abundance of ROH given a random distribution of ROH across the genome (see Methods for details). Ten such regions were found on chromosomes 3, 11, 14, 16, 20, 21, and 22 (Figure 6, Table S2). Thus, it appears that several genomic regions likely contained loci with strong enough deleterious fitness effects when homozygous to substantially reduce the

frequency of individuals carrying IBD segments in these regions. Repeating this analysis with different ROH length thresholds ($g \leq 10$, 20 or 100) did not substantively change the results (results not shown). As in many types of genomic analysis, it is possible that technical artifacts such as genome assembly errors or incorrectly mapped sequence reads could have contributed to some of the regions with low ROH abundance. These genomic regions should therefore be analyzed in further detail, including genotyping or sequencing of larger population samples.

Discussion

This study is one of the first large-scale examples of the power of genome resequencing to record the genomic consequences of inbreeding in a population of conservation concern. The combination of a huge number of SNPs resulting from the whole-genome resequencing of 97 individuals and a high-quality genome assembly enabled us to precisely delineate IBD chromosome segments as ROH, and to quantify realized genomic inbreeding, and to identify genomic regions that likely contributed substantially to inbreeding depression in this vulnerable population of Scandinavian wolves. In many individuals, the signatures of inbreeding were remarkably visible as entire or nearly entire chromosomes were completely homozygous (Figures 2 and 3).

Our results demonstrate that the vast majority of IBD segments in a recently bottlenecked population are actually very short and originate from common ancestors in the far past. However, quantitatively these short IBD segments contributed little to individual F_{ROH}, which was primarily governed by more limited numbers of very long segments resulting from common ancestors of parents less than 10 generations ago. Still, while F_P correlated well with F_{ROH} over a range of time spans to common ancestors, it becomes an increasingly downward biased estimator of F_{ROH} as older IBD segments are taken into account.

Our results also provide empirical evidence based on large-scale whole-genome resequencing that inbreeding is better measured with molecular genetic data than with F_P estimated from an extensive pedigree. . While several previous studies have assessed correlations among molecular measures of inbreeding and F_p[25,26,28], none have rigorously evaluated the performance of F_p and molecular measures of inbreeding because the true realized genomic inbreeding was unknown[7]. To our knowledge, this is the first study to carry out such an analysis. F_P has been the standard measure of individual inbreeding for decades[10]. While pedigrees are clearly still useful for estimating inbreeding (e.g., in species with many chromosomes[12]), and for many other purposes[10], molecular measures of *F* are more powerful as they account for related and inbred pedigree founders and immigrants, and the stochastic effects of linkage and Mendelian segregation. Additionally, molecular approaches allow mapping of loci contributing to inbreeding depression[5,44]. An interesting question that arises from our observations and that should be investigated further is the overall phenotypic consequences of individuals within a population being IBD for different haplotypes of very large chromosome segments. One might expect that this will disclose 'hidden' phenotypic variation encoded by rare variants or variation that is otherwise rarely seen due to dominance effects.

The demonstration of inbreeding and relatedness among immigrants has important implications for population viability and the design of management programs. In the case of the Scandinavian wolf population, having inbred and related immigrants means that animals are on average more inbred than it appears based on pedigree information alone (Figure 4). This emphasizes the importance of immigration into the population to limit inbreeding and inbreeding depression. Also, it highlights the importance of taking the genetic status (i.e., the degree of inbreeding and relatedness arising from finite population size and population fragmentation) of a larger metapopulation into account. Importantly, a similar situation may apply to many other species of conservation concern where a fragmented population structure increases the likelihood for inbreeding and close relatedness among immigrants[26].

Identifying regions of the genome with exceptionally low abundance of ROH is an important step towards understanding the genetic basis of inbreeding depression in Scandinavian wolves. These genomic regions are likely to contain loci with overdominant or deleterious recessive alleles strongly contributing to inbreeding depression. Future mapping studies could be used to directly test for phenotypic effects of IBD in these regions. Ascertaining the loci underlying inbreeding depression and the magnitude of their phenotypic effects is crucial to advancing our understanding of the genetic basis of inbreeding depression and the potential for purging to lessen the genetic load.

Methods

Study population and DNA samples

As in many other parts of the world[45], the wolf experienced a significant population decline in Scandinavia during the last centuries. Once common and spread over the entire Scandinavian peninsula, hunt and persecution eventually led to the functional extinction of wolves in the 1960-70s[35]. The closest surviving populations were found in eastern Finland (where it was rare) and western Russia. The Scandinavian population was subsequently re-established in the early 1980s by a single mating pair that are likely to have had an eastern origin[32,36]. The founder female was killed in 1985 and the founding male disappeared one year later. Subsequent breeding 1987-1990 consisted of successive mating between sibling and parent-offspring pairs resulting in severe inbreeding[30,33,34]. A third (male) founder immigrated and reproduced in the population in 1991-1993 but no further successful immigration occurred until 2008, after which five reproductively successful immigrants have entered Scandinavia from the Finnish-Russian population[30,36,46]. Before the arrival of the third founder, there was only one reproducing pack and likely no more than 10 wolves in the population. The immigrant male in 1991 had very high reproductive success and the population subsequently grew to around 365 (estimated range 300-443) by the winter season 2014/2015[45].

426 *Parentage assignment and pedigree construction*

427

428 To determine parental identities, we used a two-step process based on the variation at
429 19-36 microsatellite loci (see Åkesson *et al*.[30]) and field observations (Liberg *et al*.[34]
430 Åkesson *et al.*[30]). First, parents were determined by genetic exclusion of putative
431 parental pairs, i.e. a pair of identified individuals that were known to have scent-
432 marked in the same territory. If all putative parental pairs could be excluded assuming
433 no more than two Mendelian mismatches, we used parental assignment in CERVUS
434 v3.0 using the entire database of individuals identified between 1983 and 2016. The
435 genealogy of >99% of the breeding individuals in the population could be
436 reconstructed. For a more detailed description of the reconstruction of the pedigree,
437 see Åkesson *et al.*[30].

438

439 *Sample collection and DNA extraction*

440

441 We selected 97 DNA-samples collected invasively from live caught (blood or skin
442 tissue) or dead (tissue) wolves in Scandinavia. Captures, handling and collaring
443 of wolves[31] was in accordance with ethical requirements and have been
444 approved by the Swedish Animal Welfare Agency (Permit Number: C 281/6) and
445 the Norwegian Experimental Animal Ethics Committee (permit number:
446 2014/284738-1).

447

448 The individuals used in the study were chosen based on a sampling scheme
449 consisting of (i) all wolves sampled before 1991 and (ii) wolves distributed in
450 predefined individual categories (Table S1) characterizing five inbreeding
451 classes ($0 \leq F_P < 0.1$, $0.1 \leq F_P < 0.2$, $0.2 \leq F_P < 0.3$, $0.3 \leq F_P < 0.4$, $0.4 \leq F_P < 0.5$) and three
452 temporal classes (sampling year period 1991-1998, 1999-2006, 2007-2014). The
453 representation from each category varied depending on the availability of
454 individuals. Genomic DNA from tissue and blood was isolated using standard

phenol/chloroform–isoamylalcohol extraction and the precipitate was solved in 20-100 µl distilled water.

Whole-genome resequencing and variant calling and filtering

Library construction and 150 basepair paired-end sequencing was performed on an Illumina HiSeqX with standard procedures. Sequencing reads were mapped to the dog genome build CanFam3.1, using BWA v0.7.13[47]. The resulting BAM files were sorted using SAMtools v1.3[48], duplicate marked using Picard v1.118 (http://broadinstitute.github.io/picard/), and locally realigned around indels using GATK v3.3.0[49,50]. Read information was updated in the bam files with Picard FixMateInformation.

A first round of variant calling was performed with GATK HaplotypeCaller and the whole cohort genotyped using GATK GenotypeGVCFs. The resulting variant list was filtered for low quality variants with low allele frequency using bcftools v1.3 (http://samtools.github.io/bcftools/) (filtering criteria: INFO/AF < 0.01 && INFO/MQRankSum<-0.2). The variants passing this filter were used as a true positive set of variant sites for BQSR, performed with GATK. Variant calling was repeated for the recalibrated bam files and then the whole cohort re-genotyped using GATK.

We applied several SNP filters to ensure high quality of the data. First, all tri-allelic loci, loci with only heterozygous or only homozygous genotypes, and loci with mean read depth (among all 97 individuals) of less than 10 or greater than 52 (twice the mean sequence read depth genome-wide) were removed. Second, genotypes with Phred-scaled genotype quality scores of less than 20 and loci that had missing genotypes in >15 individuals were discarded. We then removed loci where the *P*-value was <0.001 in a test for an excess of heterozygotes relative to Hardy-Weinberg genotype proportions using the --*hardy* function in VCFtools[50]. Finally, we retained only loci with minor allele frequency ≥0.05. The heterozygote excess and read depth filters were successful at removing SNPs in regions with poor read mapping (Figures S8-S9).

Inferring SNP linkage map positions

The genetic map position (in cM) of each SNP in the wolf whole-genome resequencing data were inferred from a recent sex-averaged high-density domestic dog linkage map[51]. This was done by first identifying the closest upstream and downstream SNP included in the dog map. We then interpolated the genetic position of the focal SNP while assuming that the recombination rate was constant between the two flanking linkage-mapped SNPs[29].

Quantifying individual inbreeding

The pedigree was determined using parentage information derived from field observations and microsatellite-based parentage assignments as described previously[30,34]. F_P was calculated using CFC v1.0 software[52]. To estimate F_{ROH}, we identified ROH using a likelihood ratio method[17,29,38]. First, we split each chromosome up into sliding windows that each included 100 adjacent SNPs, using a step size of 10 SNPs. For each 100-SNP window *i*, and individual *j*, we calculated the probability (Pr) of the genotype at each SNP *k* (G_k) assuming the SNP was IBD, and separately assuming the SNP was non-IBD. We then calculated a LOD score by summing the $\log_{10}$ of the ratio of these probabilities across all loci within the window:

$$LOD(j,i) = \sum_{k=1}^{k_i} \log_{10}\left(\frac{\Pr(G_k \mid IBD)}{\Pr(G_k \mid non-IBD)}\right)$$

The genotype probabilities under IBD and non-IBD were calculated according to Wang *et al.*[17], accounting for occasional heterozygous positions within ROH resulting from sequencing errors, read mapping errors (e.g., due to segmental duplications), and occasional mutations. Specifically, we accepted that 2% of SNPs would be heterozygous within IBD segments. Using shorter window sizes (40- and 60-SNP windows) resulted in obvious IBD segments being artificially broken in regions with poor mapping of sequence reads (results not shown). Likewise, assuming fewer heterozygous positions within IBD segments (e.g.,

0.1%) artificially broke up obvious IBD segments in regions with apparent poor mapping of sequence reads.

We estimated g for each ROH in order to include only IBD segments arising from recent ancestors when estimating F_{ROH}. For each ROH, we solved for g in the equation $l = 100/2g$ cM, where l is the length of the ROH in cM[39]. We estimated the map length of each ROH in cM by interpolating the mapping positions of each SNP in the genome from a recent high-density linkage map of the domestic dog genome[51], assuming that the recombination rate is conserved between domestic dogs and wolves.

Permutation test for regions with exceptionally low ROH abundance

We used a permutation (randomization) approach to simulate the null distribution of ROH abundance in 100 kb windows. For each of 5,000,000 permutations, we first randomly sampled 97 individuals with replacement from the sequenced wolves. We then randomly selected a 100 kb chromosome segment from the genome of each individual independently. We then quantified ROH abundance for the segment as the sum of the lengths of all IBD parts of the 97 sampled chromosome segments (in kb) divided by the length of the segment (100 kb). A *P*-value for each 100 kb segment in the genome was calculated as the proportion of the 5,000,000 permuted ROH abundance estimates that were smaller than the observed ROH abundance. The *P*-value was set to 1/5,000,001 for segments where none of the 5,000,000 permutation repetitions produced an ROH abundance ≤ the observed value. We used the Bonferroni method to correct for multiple testing. Specifically, the *P*-value below which a test was considered statistically significant was set to 0.05 divided by 22,055 (the number of analyzed 100 kb windows).

ROH abundance has previously been strongly related to the recombination rate and SNP density in other taxa (e.g., humans and birds), with low ROH abundance found in regions with high recombination rate and/or relatively low SNP density[29,43]. We tested for such effects in the present study to determine if

genome-wide variation in recombination rate or genetic diversity were likely explanations for the observed pattern of ROH abundance across the genome. We measured nucleotide diversity (π), ROH density (as described above), and the mean recombination rate (in cM/Mb from the domestic dog linkage map[58]) in 100 kb windows across the genome. We then fitted a regression model of ROH density versus π, then a separate regression model of ROH density versus recombation rate. ROH abundance was only very weakly correlated with nucleotide diversity (r^2 = 0.006, Figure S10) and recombination rate (r^2 = 0.0005, Figure S11). Thus, levels of genetic diversity and recombination rate do not appear to substantially affect the pattern of ROH abundance across the genome in this population of wolves.

Acknowledgements

Financial support was obtained from the Swedish Research Council, Swedish Research Council Formas, Swedish Environmental Protection Agency, Research Council of Norway, Norwegian Environment Agency, and Marie-Claire Cronstedts Foundation. We thank the National Veterinary Institute (Sweden), Norwegian Institute for Nature Research, Swedish Museum of Natural History, County Administrative Boards in Sweden, Wildlife Damage Centre at the Swedish University of Agricultural Sciences and Inland Norway University of Applied Sciences for contributing with samples. The preparation of samples was conducted by Anna Danielsson and Eva Hedmark at Grimsö Wildlife Research Station at the Swedish University of Agricultural Sciences. Bioinformatic computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX). We thank two anonymous reviewers for helpful comments on a previous version of the manuscript.

Author Contributions

Conceived the project: H.E. Initiated the project: M.K., M.Å., Ø.F., H.S., C.W. and H.E. Designed the project: H.E., M.K. and M.Å. Performed data analysis: M.K. and T.F. Original reconstruction of pedigree: O.L. Maintenance, updating and refinement of the pedigree: M.Å., Ø.F. and O.L. Calculations of F_P: M.Å. Coordinated field work and sampling: O.L., H.S., P.W. and C.W. Performed variant calling: P.O. The first draft of the manuscript was written by M.K. with input from H.E. and T.F. All authors contributed to discussing the results and editing the manuscript.

Data availability

Sequence data has been deposited to the European Nucleotide Archive (accession number PRJEB20635).

R scripts used to detect ROH and to infer genetic mapping positions of SNPs are available upon request.

References

1 Saccheri, I. *et al.* Inbreeding and extinction in a butterfly metapopulation. *Nature* **392**, 491-494 (1998).

2 Frankham, R. Genetics and extinction. *Biological Conservation* **126**, 131-140 (2005).

3 Keller, L. F. & Waller, D. M. Inbreeding effects in wild populations. *Trends in Ecology & Evolution* **17**, 230-241 (2002).

4 Fountain, T., Nieminen, M., Sirén, J., Wong, S. C. & Hanski, I. Predictable allele frequency changes due to habitat fragmentation in the Glanville fritillary butterfly. *Proceedings of the National Academy of Sciences, USA* **113**, 2678-2683 (2016).

5 Charlesworth, D. & Willis, J. H. The genetics of inbreeding depression. *Nature Reviews Genetics* **10**, 783-796 (2009).

6 Hedrick, P. W. & Garcia-Dorado, A. Understanding inbreeding depression, purging, and genetic rescue. *Trends in Ecology & Evolution* **31**, 940-952 (2016).

7 Kardos, M., Taylor, H. R., Ellegren, H., Luikart, G. & Allendorf, F. W. Genomics advances the study of inbreeding depression in the wild. *Evolutionary Applications* **9**, 1205-1218 (2016).

8 Malécot, G. *The Mathematics of Heredity*. (W. H. Freeman, 1970).
9 Wright, S. Coefficients of inbreeding and relationship. *The American Naturalist* **56**, 330-338 (1922).
10 Pemberton, J. M. Wild pedigrees: the way forward. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **275**, 613-621 (2008).
11 Taylor, H. R., Kardos, M. D., Ramstad, K. M. & Allendorf, F. W. Valid estimates of individual inbreeding coefficients from marker-based pedigrees are not feasible in wild populations with low allelic diversity. *Conservation Genetics* **16**, 901-913 (2015).
12 Franklin, I. R. The distribution of the proportion of the genome which is homozygous by descent in inbred individuals. *Theoretical Population Biology* **11**, 60-80 (1977).
13 Hedrick, P. W., Kardos, M., Peterson, R. O. & Vucetich, J. A. Genomic variation of inbreeding and ancestry in the remaining ywo Isle Royale wolves. *Journal of Heredity* **108**, 120-126 (2017).
14 Knief, U., Kempenaers, B. & Forstmeier, W. Meiotic recombination shapes precision of pedigree- and marker-based estimates of inbreeding. *Heredity* **118** (2017).
15 Forstmeier, W., Schielzeth, H., Mueller, J. C., Ellegren, H. & Kempenaers, B. Heterozygosity–fitness correlations in zebra finches: microsatellite markers can be better than their reputation. *Molecular Ecology* **21**, 3237-3249 (2012).
16 Fisher, R. A. *The Theory of Inbreeding*. (Oliver and Boyd, 1965).
17 Wang, J. Pedigrees or markers: Which are better in estimating relatedness and inbreeding coefficient? *Theoretical Population Biology* **107**, 4-13 (2016).
18 Leary, R. F., Allendorf, F. W. & Knudsen, K. L. Developmental stability and enzyme heterozygosity in rainbow trout. *Nature* **301**, 71-72 (1983).
19 Pierce, B. A. & Mitton, J. B. Allozyme heterozygosity and growth in the tiger salamander, Ambystoma tigrinum. *Journal of Heredity* **73**, 250-253 (1982).
20 Coltman, D. W., Pilkington, J. G., Smith, J. A. & Pemberton, J. M. Parasite-mediated selection against inbred Soay Sheep in a free-living, island population. *Evolution* **53**, 1259-1267 (1999).
21 Slate, J. *et al.* Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. *Heredity* **93**, 255-265 (2004).
22 Miller, J. M. *et al.* Estimating genome-wide heterozygosity: effects of demographic history and marker type. *Heredity* **112**, 240-247 (2014).
23 Balloux, F., Amos, W. & Coulson, T. Does heterozygosity estimate inbreeding in real populations? *Molecular Ecology* **13**, 3021-3031 (2004).
24 Szulkin, M., Bierne, N. & David, P. Heterozygosity-fitness correlations: A time for reapprisal. *Evolution* **64**, 1202-1217 (2010).
25 Huisman, J., Kruuk, L. E. B., Ellis, P. A., Clutton-Brock, T. & Pemberton, J. M. Inbreeding depression across the lifespan in a wild mammal population. *Proceedings of the National Academy of Sciences USA* **113**, 3585-3590 (2016).
26 Chen, N., Cosgrove, Elissa J., Bowman, R., Fitzpatrick, John W. & Clark, Andrew G. Genomic consequences of population decline in the endangered Florida Scrub-Jay. *Current Biology* **26**, 2974-2979 (2016)

27 Kardos, M., Luikart, G. & Allendorf, F. W. Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees. *Heredity* **115**, 63-72 (2015).
28 Hoffman, J. I. *et al.* High-throughput sequencing reveals inbreeding depression in a natural population. *Proceedings of the National Academy of Sciences, USA* **111**, 3775-3780 (2014).
29 Kardos, M., Qvarnström, A. & Ellegren, H. Inferring individual inbreeding and demographic history from segments of identity by descent in *Ficedula* flycatcher genome sequences. *Genetics* **205**, 1319-1334 (2017).
30 Åkesson, M. *et al.* Genetic rescue in a severely inbred wolf population. *Molecular Ecology* **25**, 4745-4756 (2016).
31 Bensch, S. *et al.* Selection for heterozygosity gives hope to a wild population of inbred wolves. *PLoS ONE* **1**, e72 (2006).
32 Flagstad, Ø. *et al.* Two centuries of the Scandinavian wolf population: patterns of genetic variability and migration during an era of dramatic decline. *Molecular Ecology* **12**, 869-880 (2003).
33 Vilà, C. *et al.* Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, 91-97 (2003).
34 Liberg, O. *et al.* Severe inbreeding depression in a wild wolf (*Canis lupus*) population. *Biology Letters* **1**, 17-20 (2005).
35 Haglund, B. De stora rovdjurens vintervanor II. *Viltrevy* **5**, 213-361 (1969).
36 Wabakken, P., Sand, H., Liberg, O. & Bjärvall, A. The recovery, distribution, and population dynamics of wolves on the Scandinavian peninsula, 1978-1998. *Canadian Journal of Zoology* **79**, 710-725 (2001).
37 Seddon, J. M., Sundqvist, A. K., Björnerfeldt, S. & Ellegren, H. Genetic identification of immigrants to the Scandinavian wolf population. *Conservation Genetics* **7**, 225-230 (2006).
38 Szpiech, Z. A. *et al.* Long runs of homozygosity are enriched for deleterious variation. *The American Journal of Human Genetics* **93**, 90-102 (2013).
39 Thompson, E. A. Identity by descent: variation in meiosis, across genomes, and in populations. *Genetics* **194**, 301-326 (2013).
40 García-Dorado, A. Understanding and predicting the fitness decline of shrunk populations: inbreeding, purging, mutation, and standard selection. *Genetics* **190**, 1461-1476 (2012).
41 Crow, J. F. & Kimura, M. *An Introduction to Population Genetics Theory*. (The Blackburn Press, 1970).
42 Kardos, M., Allendorf, F. W. & Luikart, G. Evaluating the role of inbreeding depression in heterozygosity-fitness correlations: how useful are tests for identity disequilibrium? *Molecular Ecology Resources* **14**, 519-530 (2014).
43 Pemberton, T. J. *et al.* Genomic patterns of homozygosity in worldwide human populations. *The American Journal of Human Genetics* **91**, 275-292 (2012).
44 Hedrick, P. W., Hellsten, U. & Grattapaglia, D. Examining the cause of high inbreeding depression: analysis of whole-genome sequence data in 28 selfed progeny of *Eucalyptus grandis*. *New Phytologist* **209**, 600-611 (2016).

45 Chapron, G. *et al.* Recovery of large carnivores in Europe's modern human-dominated landscapes. *Science* **346**, 1517 (2014).
46 Laikre, L., Olsson, F., Jansson, E., Hossjer, O. & Ryman, N. Metapopulation effective size and conservation genetic goals for the Fennoscandian wolf (*Canis lupus*) population. *Heredity* **117**, 279-289 (2016).
47 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
48 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
49 McKenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**, 1297-1303 (2010).
50 Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158 (2011).
51 Campbell, C. L., Bhérer, C., Morrow, B. E., Boyko, A. R. & Auton, A. A pedigree-based map of recombination in the domestic dog genome. *G3: Genes|Genomes|Genetics* **6**, 3517 (2016).
52 Sargolzaei, M., Iwaisaki, H. & Colleau, J. J. A fast algorithm for computing inbreeding coefficients in large populations. *Journal of Animal Breeding and Genetics* **122**, 325-331 (2005).

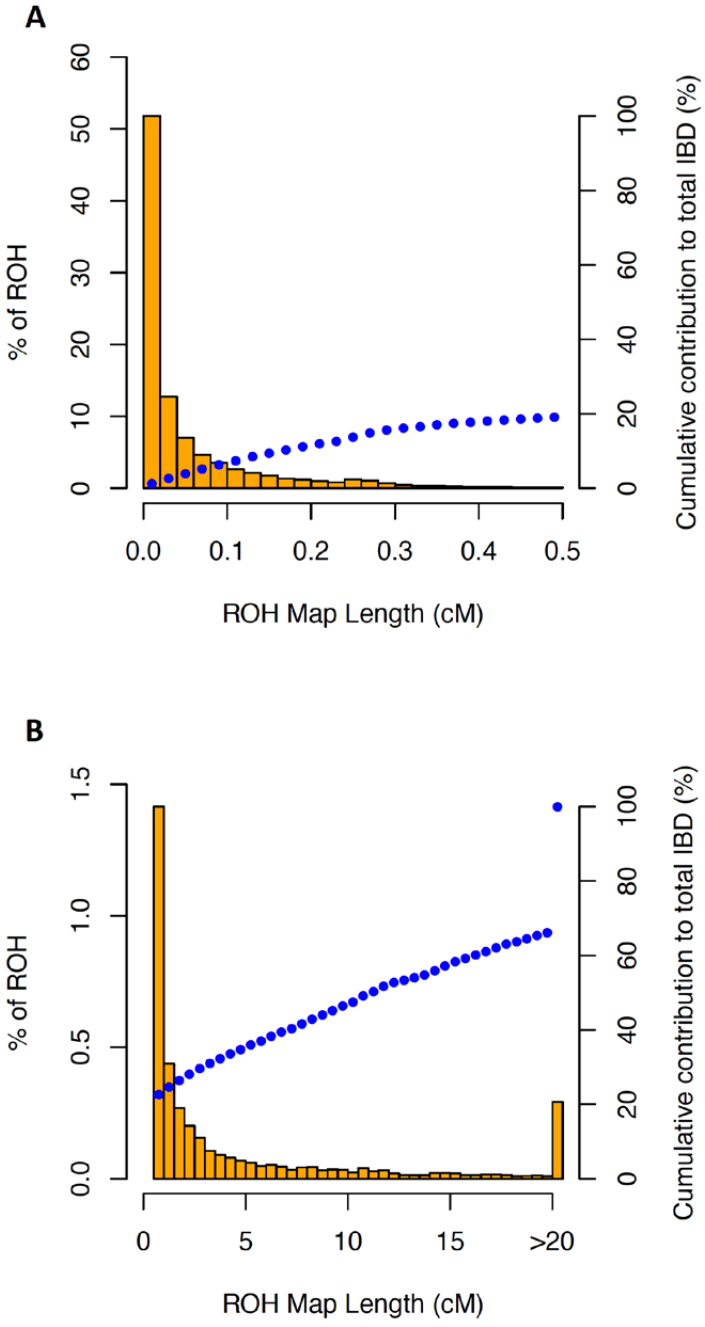


Figure 1. The length distribution of ROH shorter than 0.5 cM (**A**) and 0.5 cM or longe (**B**) identified in 97 Scandinavian wolf genomes. The blue points show the cumulative contribution of ROH of different lengths to the total length of IBD regions (right vertical axis). Note that A and B have different ranges on the y-axis.

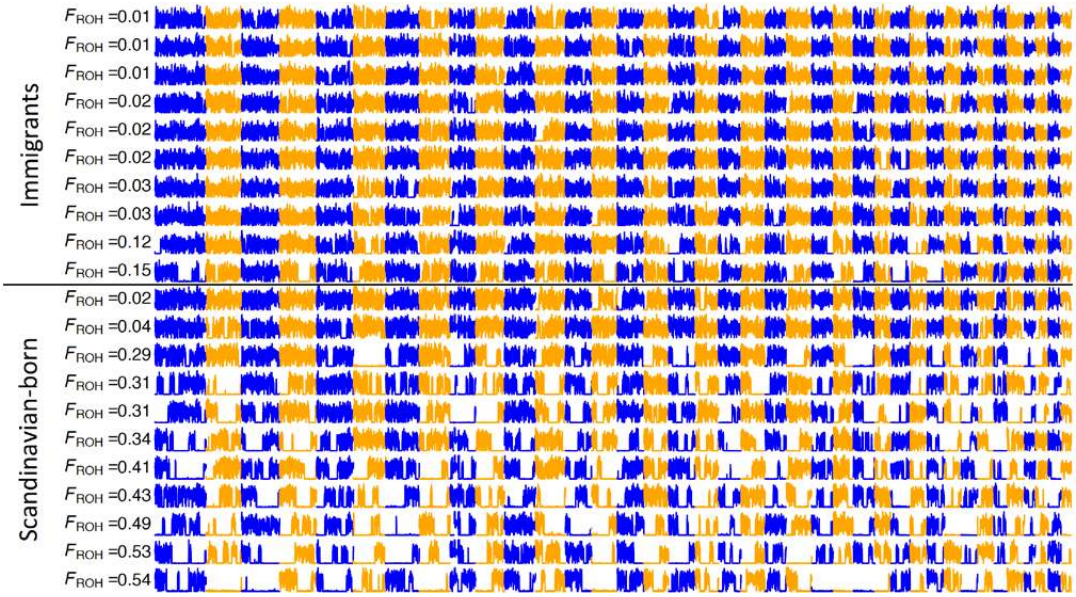
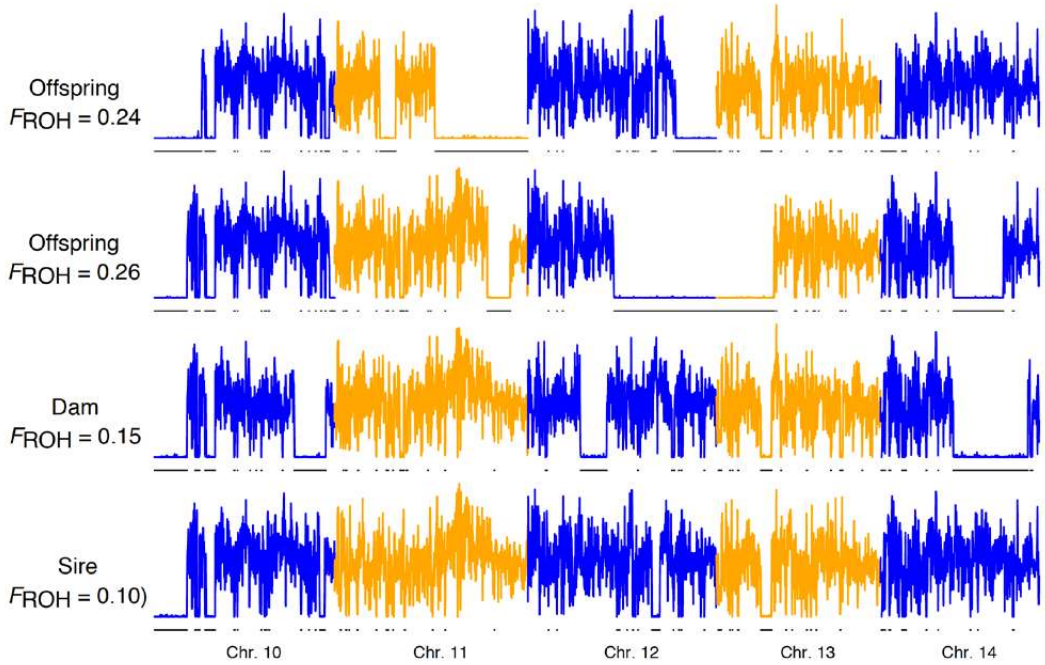


Figure 2. Heterozygosity across the 38 autosomes of 21 Scandinavian wolves. Heterozygosity was measured in non-overlapping 100 kb windows as the proportion of SNPs within each window that were heterozygous in the individual. The y-axis ranges from zero to one for each individual. The top ten individuals are immigrants into the population, followed by 11 Scandinavian-born individuals. The bottom individual is the most highly inbred wolf in the study. Chromosomes 1-38 are arranged left to right, with alternating blue and orange representing different chromosomes. Detailed plots of heterozygosity and identified ROH are provided for each chromosome in each of the 97 individuals in Supplementary File 1.



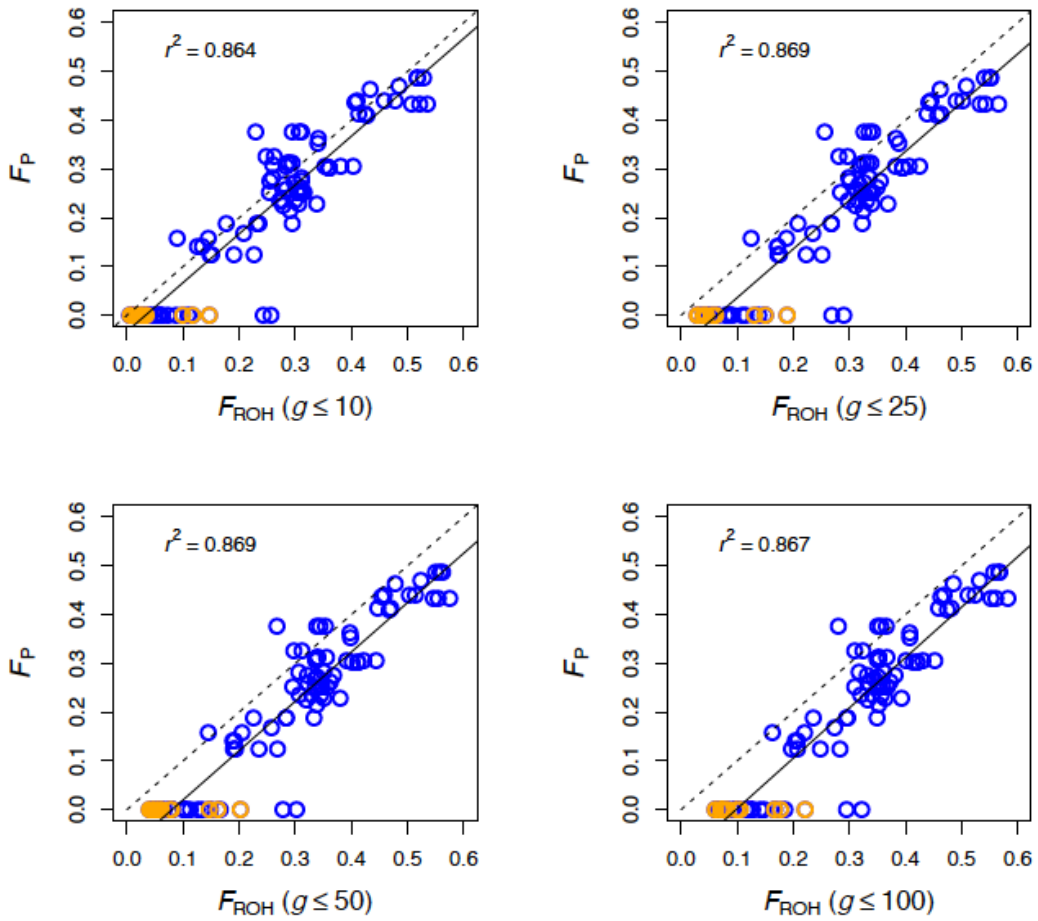
765

766

767

768 **Figure 3.** Heterozygosity across chromosomes 10-14 in two full siblings (top
769 two panels) and their parents (bottom two panels). The black lines show long
770 ROH (>100 kb) identified with the likelihood ratio-based sliding window
771 approach.

772



773
774
775

776 **Figure 4.** The relationship between F_P and F_{ROH} measured with the whole
777 genome in Scandinavian wolves. F_P is shown on the y-axis, and F_{ROH} measured
778 using ROH arising from relatively recent ancestors ($g \leq$ 10-100 generations) on
779 the x-axis. Immigrants are shown as orange points, and Scandinavian-born
780 individuals are shown as blue points. The dashed line has an intercept of zero
781 and a slope of one. Points below the line thus represent cases where F_P
782 underestimated F_{ROH}. The solid line is the fitted line from a regression of F_P
783 versus F_{ROH}.
784

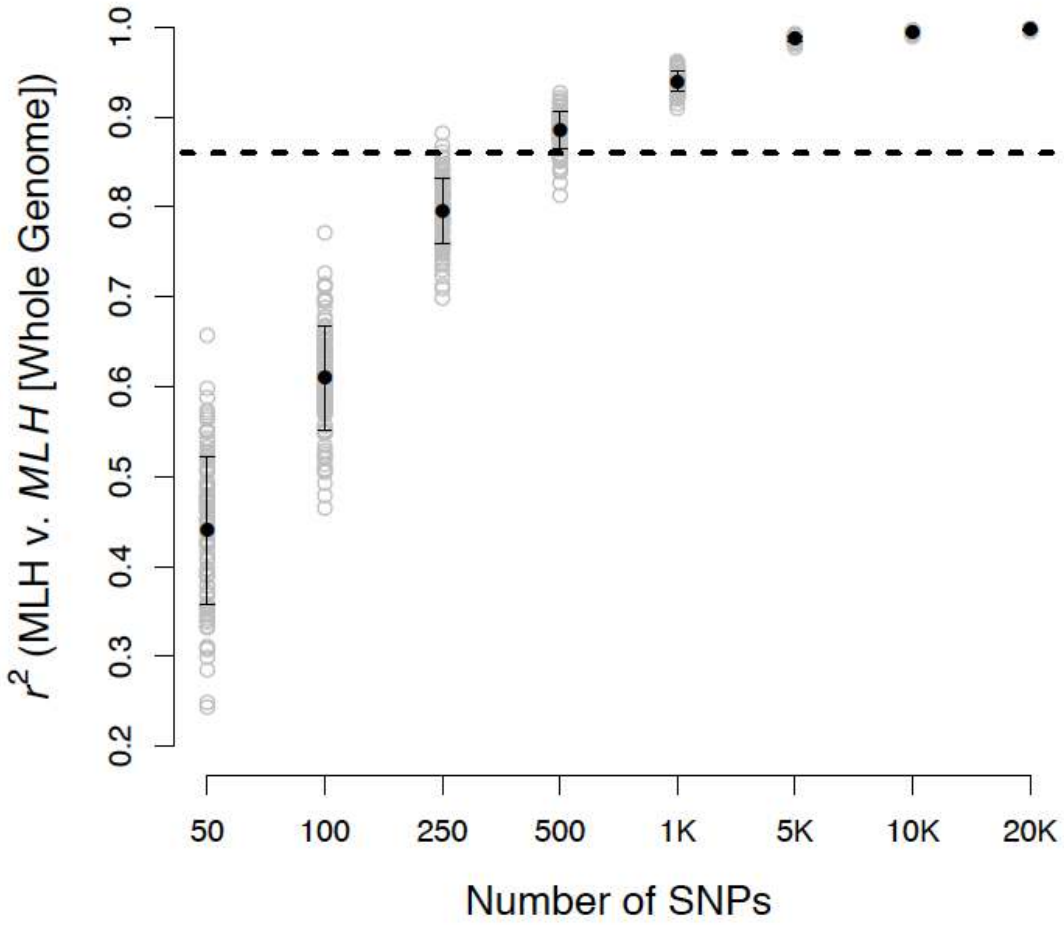


Figure 5. r^2 from regressions of *MLH* estimated with subsampled SNPs versus *MLH* calculated from the entire genome. The number of subsampled loci used to estimate *MLH* is shown on the x-axis. The black points represent the mean r^2 from analyses of 100 replicate subsamples of SNPs (+/- 1 sd). The dashed line represents the r^2 between the pedigree inbreeding coefficient and *F*. Note that the x-axis is not scaled.

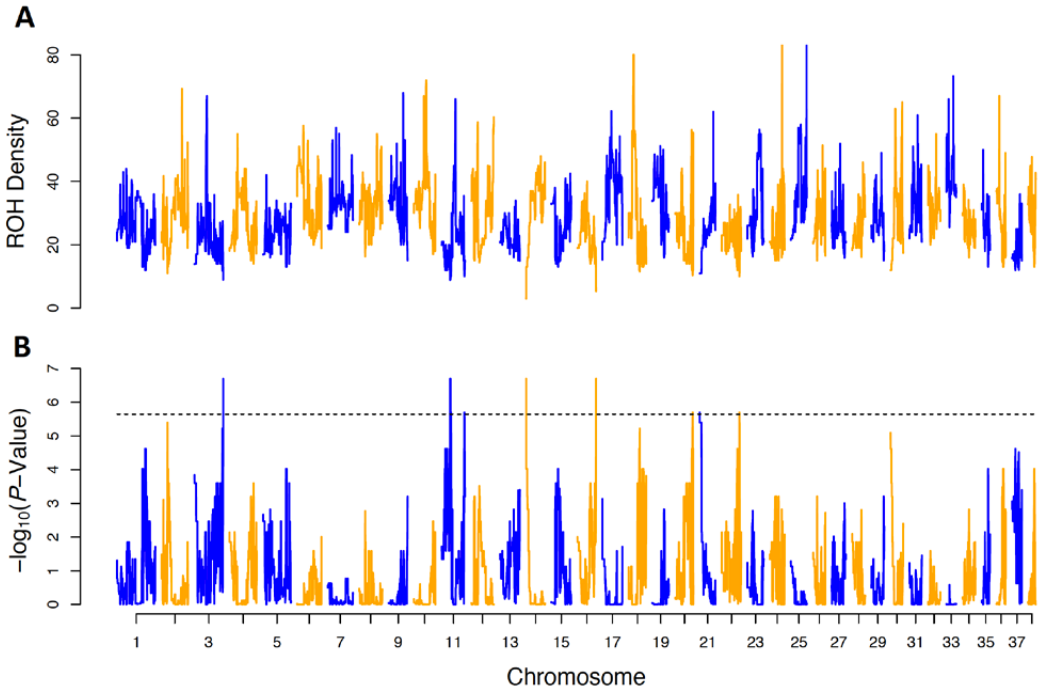


Figure 6. The density of ROH versus genomic position (**A**), and the $-\log_{10}$ of *P*-values from permutation tests for deficit of ROH abundance in non-overlapping 100 kb windows (**B**). Chromosomes are arranged 1 to 38 from left to right. ROH density was measured as the summed kb in ROH across all individuals divided by the window length. The horizontal dashed line in B represents the Bonferroni corrected threshold of statistical significance.

Table S1. Identity, sex (M, male; F, female), origin, pedigree-based inbreeding (F_P) and the longest ancestral path (i.e. highest number of generations to a founder) of the 97 wolves included in the study. The individuals were chosen based a sampling scheme consisting of (i) all wolves sampled before 1991 (temporal class 1983-1990) plus all immigrants and (ii) randomly selected wolves within five predefined inbreeding classes and three temporal classes.

ID	Sex	Origin	F_P	Longest ancestral path	Inbreeding class	Temporal class
D-05-18	M	Immigrant	0	-	- -	-
D-77-01	M	Immigrant	0	-	- -	-
D-79-01	F	Immigrant	0	-	- -	-
D-85-01*	F	Immigrant	0	-	- -	-
G23-13*	M	Immigrant	0	-	- -	-
G31-13*	F	Immigrant	0	-	- -	-
G82-10	F	Immigrant	0	-	- -	-
M-02-15	M	Immigrant	0	-	- -	-
M-05-01	M	Immigrant	0	-	- -	-
M-07-02	M	Immigrant	0	-	- -	-
M-09-03*	M	Immigrant	0	-	- -	-
M-10-10*	M	Immigrant	0	-	- -	-
D-84-03	M	Scandinavian born	0	1	0-0.1	1983-1990
D-85-02	M	Scandinavian born	0	1	0-0.1	1983-1990
D-86-01	M	Scandinavian born	0	1	0-0.1	1983-1990
D-89-01	M	Scandinavian born	0.25	2	0.2-0.3	1983-1990
D-91-01	F	Scandinavian born	0.25	2	0.2-0.3	1983-1990
D-89-03	F	Scandinavian born	0.25	2	0.2-0.3	1983-1990
D-92-05	M	Scandinavian born	0.25	2	0.2-0.3	1983-1990
D-93-02	F	Scandinavian born	0.25	2	0.2-0.3	1983-1990
D-93-03	F	Scandinavian born	0.375	3	0.3-0.4	1983-1990
M-98-02	M	Scandinavian born	0.375	3	0.3-0.4	1983-1990
D-92-06	M	Scandinavian born	0.375	3	0.3-0.4	1983-1990
D-94-01	F	Scandinavian born	0.375	3	0.3-0.4	1983-1990

ID	Sex	Origin	F_P	Longest ancestral path	Inbreeding class	Temporal class
D-92-01	M	Scandinavian born	0	2	0-0.1	1991-1998
D-99-02	M	Scandinavian born	0	2	0-0.1	1991-1998
D-93-01	M	Scandinavian born	0	2	0-0.1	1991-1998
D-92-02	M	Scandinavian born	0	2	0-0.1	1991-1998
D-96-01	M	Scandinavian born	0	2	0-0.1	1991-1998
M-98-03	F	Scandinavian born	0	2	0-0.1	1991-1998
M-98-08	M	Scandinavian born	0.125	4	0.1-0.2	1991-1998
M-00-09	M	Scandinavian born	0.125	4	0.1-0.2	1991-1998
D-00-15	M	Scandinavian born	0.125	4	0.1-0.2	1991-1998
D-01-18	M	Scandinavian born	0.125	4	0.1-0.2	1991-1998
D-05-23	M	Scandinavian born	0.234	5	0.2-0.3	1991-1998
M-01-10	F	Scandinavian born	0.234	5	0.2-0.3	1991-1998
M-03-07	F	Scandinavian born	0.234	5	0.2-0.3	1991-1998
M-98-01	F	Scandinavian born	0.281	5	0.2-0.3	1991-1998
M-03-06	M	Scandinavian born	0.188	6	0.1-0.2	1999-2006
M-09-17	M	Scandinavian born	0.188	6	0.1-0.2	1999-2006
G9-05	M	Scandinavian born	0.188	6	0.1-0.2	1999-2006
D-10-20	F	Scandinavian born	0.188	6	0.1-0.2	1999-2006
D-07-24	F	Scandinavian born	0.215	8	0.2-0.3	1999-2006
D-06-14	F	Scandinavian born	0.261	7	0.2-0.3	1999-2006
M-09-05	M	Scandinavian born	0.227	7	0.2-0.3	1999-2006
D-00-12	F	Scandinavian born	0.281	5	0.2-0.3	1999-2006
D-10-29	M	Scandinavian born	0.227	7	0.2-0.3	1999-2006

ID	Sex	Origin	F_P	Longest ancestral path	Inbreeding class	Temporal class
D-10-30	M	Scandinavian born	0.27	6	0.2-0.3	1999-2006
M-00-10	M	Scandinavian born	0.313	5	0.3-0.4	1999-2006
M-06-03	M	Scandinavian born	0.302	7	0.3-0.4	1999-2006
D-06-16	F	Scandinavian born	0.305	5	0.3-0.4	1999-2006
M-06-04	F	Scandinavian born	0.302	7	0.3-0.4	1999-2006
D-11-17	F	Scandinavian born	0.324	7	0.3-0.4	1999-2006
M-01-06	F	Scandinavian born	0.305	5	0.3-0.4	1999-2006
M-05-07	F	Scandinavian born	0.302	7	0.3-0.4	1999-2006
D-08-20	M	Scandinavian born	0.324	7	0.3-0.4	1999-2006
D-07-09	F	Scandinavian born	0.438	8	0.4-0.5	1999-2006
D-07-17	M	Scandinavian born	0.438	8	0.4-0.5	1999-2006
D-08-08	F	Scandinavian born	0.438	8	0.4-0.5	1999-2006
M-07-06	M	Scandinavian born	0.437	8	0.4-0.5	1999-2006
D-08-10	F	Scandinavian born	0.434	6	0.4-0.5	1999-2006
D-07-28	M	Scandinavian born	0.434	6	0.4-0.5	1999-2006
G67-15	F	Scandinavian born	0	1	0-0.1	2007-2014
M-10-04	F	Scandinavian born	0	8	0-0.1	2007-2014
G47-11	F	Scandinavian born	0	9	0-0.1	2007-2014
G37-10	M	Scandinavian born	0	8	0-0.1	2007-2014
M-11-02	F	Scandinavian born	0	8	0-0.1	2007-2014
G100-14	M	Scandinavian born	0	1	0-0.1	2007-2014
D-10-53	M	Scandinavian born	0	9	0-0.1	2007-2014
G106-13	M	Scandinavian born	0.166	10	0.1-0.2	2007-2014

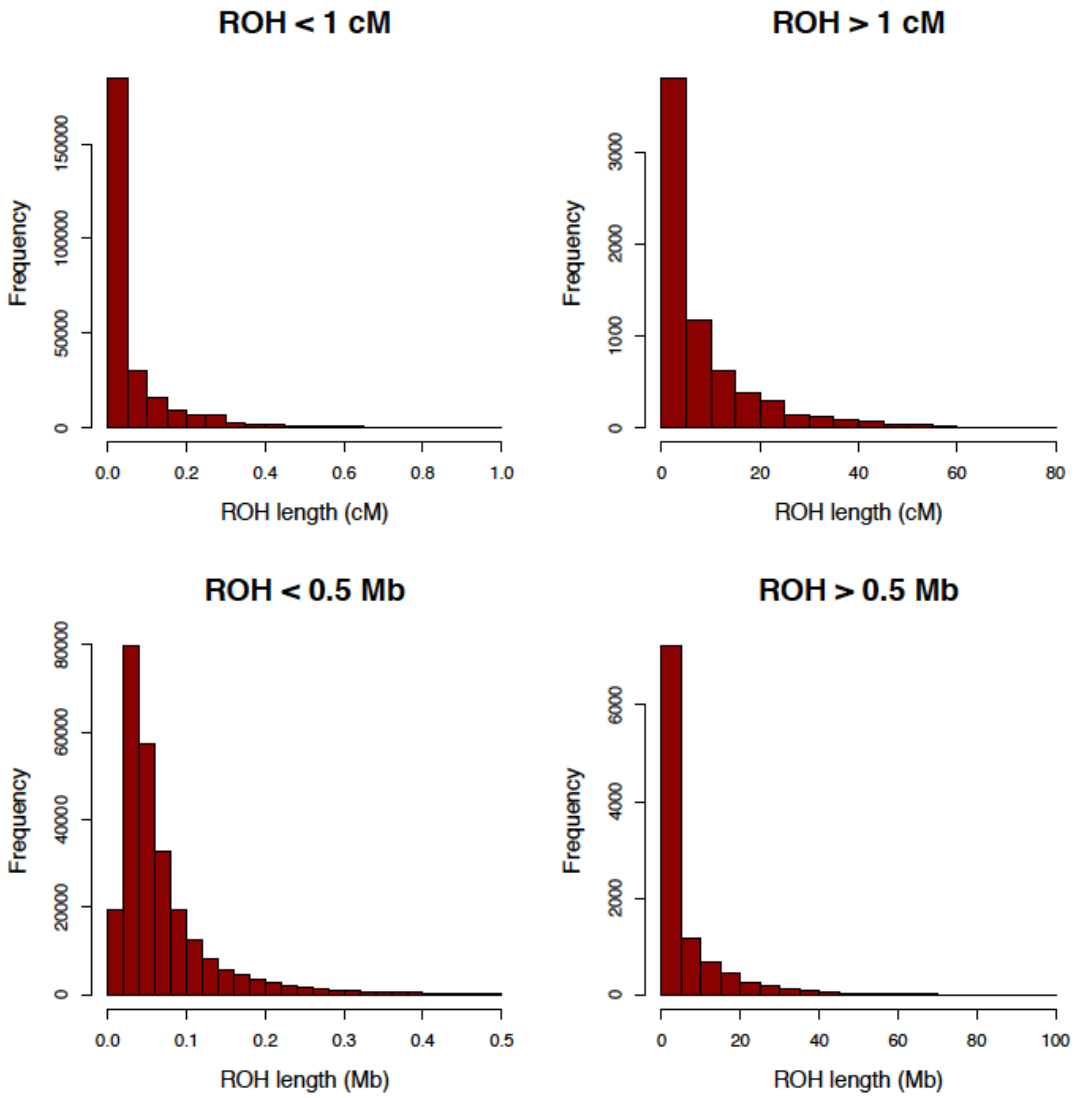
ID	Sex	Origin	F_P	Longest ancestral path	Inbreeding class	Temporal class
G111-14	M	Scandinavian born	0.158	10	0.1-0.2	2007-2014
G139-12	F	Scandinavian born	0.158	10	0.1-0.2	2007-2014
G109-11	M	Scandinavian born	0.139	9	0.1-0.2	2007-2014
G100-12	M	Scandinavian born	0.139	9	0.1-0.2	2007-2014
D-08-21	M	Scandinavian born	0.223	8	0.2-0.3	2007-2014
G87-12	F	Scandinavian born	0.274	9	0.2-0.3	2007-2014
D-11-58	F	Scandinavian born	0.257	8	0.2-0.3	2007-2014
D-08-19	F	Scandinavian born	0.267	8	0.2-0.3	2007-2014
D-10-50	M	Scandinavian born	0.275	8	0.2-0.3	2007-2014
G32-12	M	Scandinavian born	0.306	8	0.3-0.4	2007-2014
D-10-68	M	Scandinavian born	0.311	8	0.3-0.4	2007-2014
G32-15	F	Scandinavian born	0.361	11	0.3-0.4	2007-2014
G126-13	M	Scandinavian born	0.307	9	0.3-0.4	2007-2014
G97-13	M	Scandinavian born	0.306	8	0.3-0.4	2007-2014
G174-13	F	Scandinavian born	0.311	10	0.3-0.4	2007-2014
D-11-22	F	Scandinavian born	0.352	9	0.3-0.4	2007-2014
D-07-16	M	Scandinavian born	0.434	6	0.4-0.5	2007-2014
G58-15	F	Scandinavian born	0.462	9	0.4-0.5	2007-2014
G50-12	M	Scandinavian born	0.413	9	0.4-0.5	2007-2014
G175-13	M	Scandinavian born	0.413	9	0.4-0.5	2007-2014
D-10-15	F	Scandinavian born	0.486	9	0.4-0.5	2007-2014
D-10-23	F	Scandinavian born	0.486	9	0.4-0.5	2007-2014
G34-10	F	Scandinavian born	0.47	8	0.4-0.5	2007-2014

ID	Sex	Origin	F_P	Longest ancestral path	Inbreeding class	Temporal class
D-10-44	M	Scandinavian born	0.486	9	0.4-0.5	2007-2014
G110-11	F	Scandinavian born	0.41	8	0.4-0.5	2007-2014

*founders

Table S2. ROH regions with statistically significant deficit of ROH abundance. Shown are the chromosome with start and stop position (in Mb), segment lengths (Mb), mean recombination rate (cM/Mb), nucleotide diversity (π), the number of SNPs present in the segment, and ROH density. Note that adjacent 100 kb windows with statistically significant permutation tests were joined in the table.

Chr.	Start (Mb)	End (Mb)	Length (Mb)	Mean Rec. Rate. (cM/Mb)	π	SNPs	ROH Density
3	91.8	91.89	0.09	0.39	0.0001	40	8.9
11	26.7	28.2	1.5	0.6	0.0008	3,486	9
11	28.3	29.3	1	0.55	0.0012	3,545	10.1
11	72.4	73.1	0.7	1.69	0.0025	5,202	10.1
14	0	0.2	0.2	0.57	0.0005	271	3.5
16	59.3	59.6	0.3	1.27	0.0001	122	7.2
20	54.7	54.9	0.2	0.55	0.0005	312	10.3
20	55.7	55.9	0.2	2	0.0012	678	10.5
21	0	0.1	0.1	11.53	0.0004	164	10.9
22	55.8	56.1	0.3	0.86	0.0017	1,474	10



839

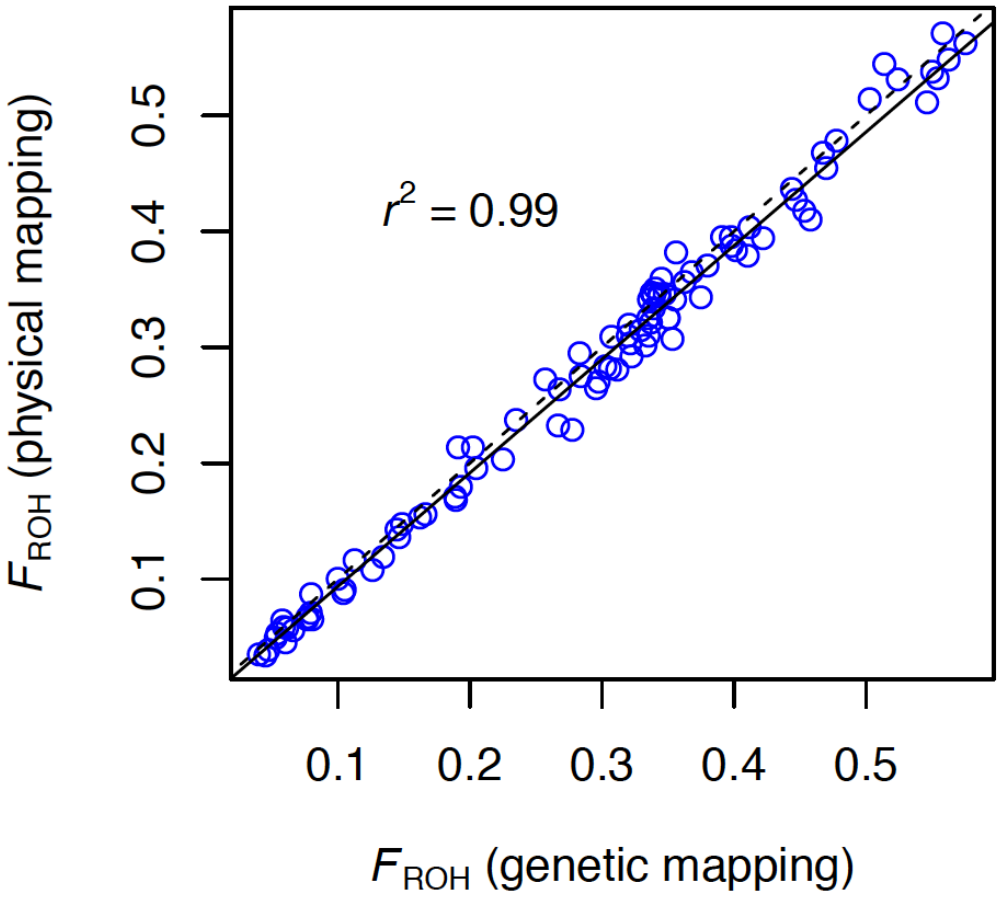
840

841 **Figure S1.** Distribution of the lengths of ROH. The genetic map lengths (in cM) of

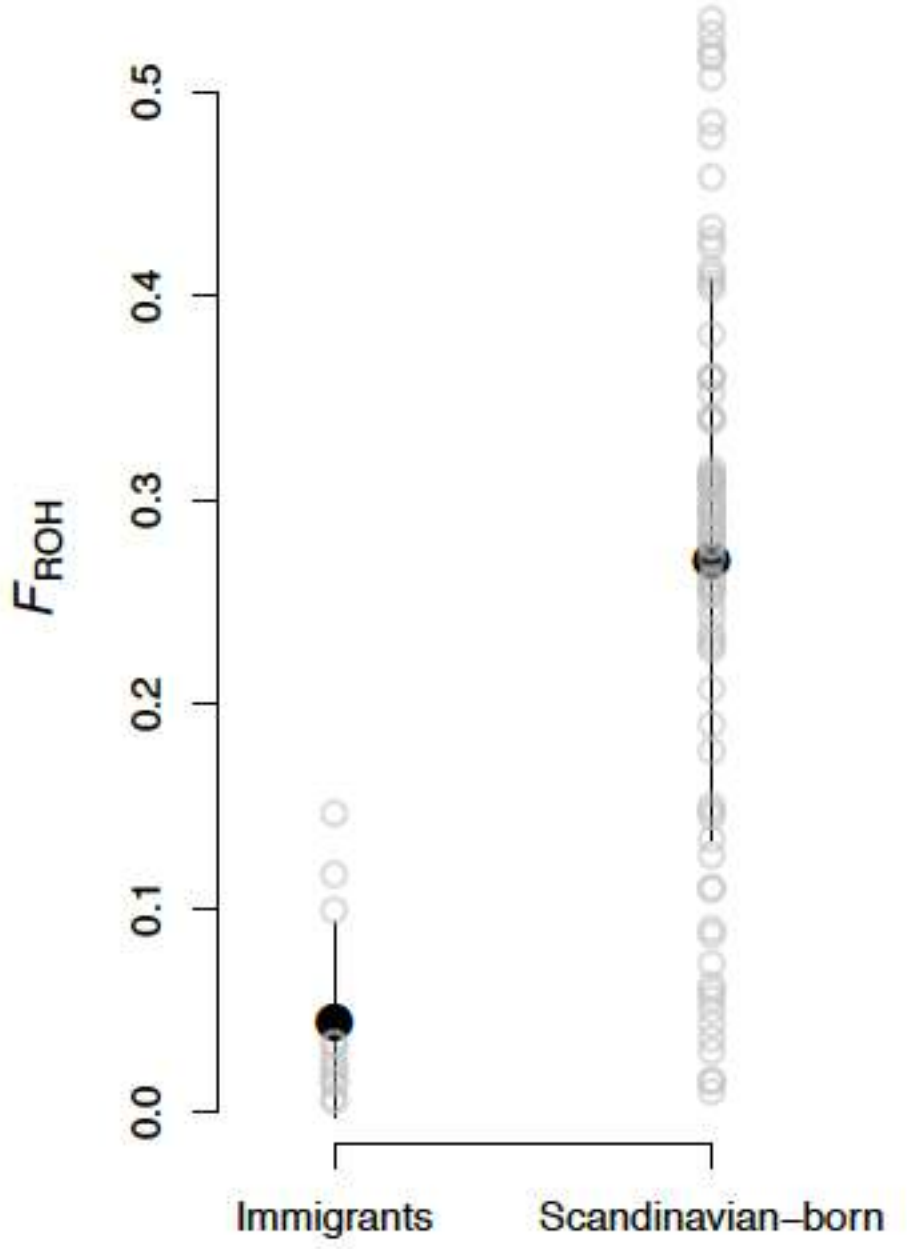
842 ROH are shown in the two upper panels. The lower two panels show the physical

843 lengths (in Mb) of ROH.

844



845
846 **Figure S2.** Scatterplot of FROH measured using physical mapping of ROH versus
847 FROH measured using genetic mapping of ROH. The ROH shown are those with
848 values of g ≤ 50. The results were essentially identical using other thresholds of
849 ROH length, and when using all ROH (data not shown). The r^2 value and the fitted
850 solid line are from a linear regression model. The dashed line has an intercept of
851 zero and a slope of one.
852



853
854
855 **Figure S3.** The distribution of F_{ROH} for immigrants and Scandinavian-born
856 wolves calculated using only ROH with $g \leq 10$ generations. The filled point and
857 error bars represent the mean +/- one standard deviation.

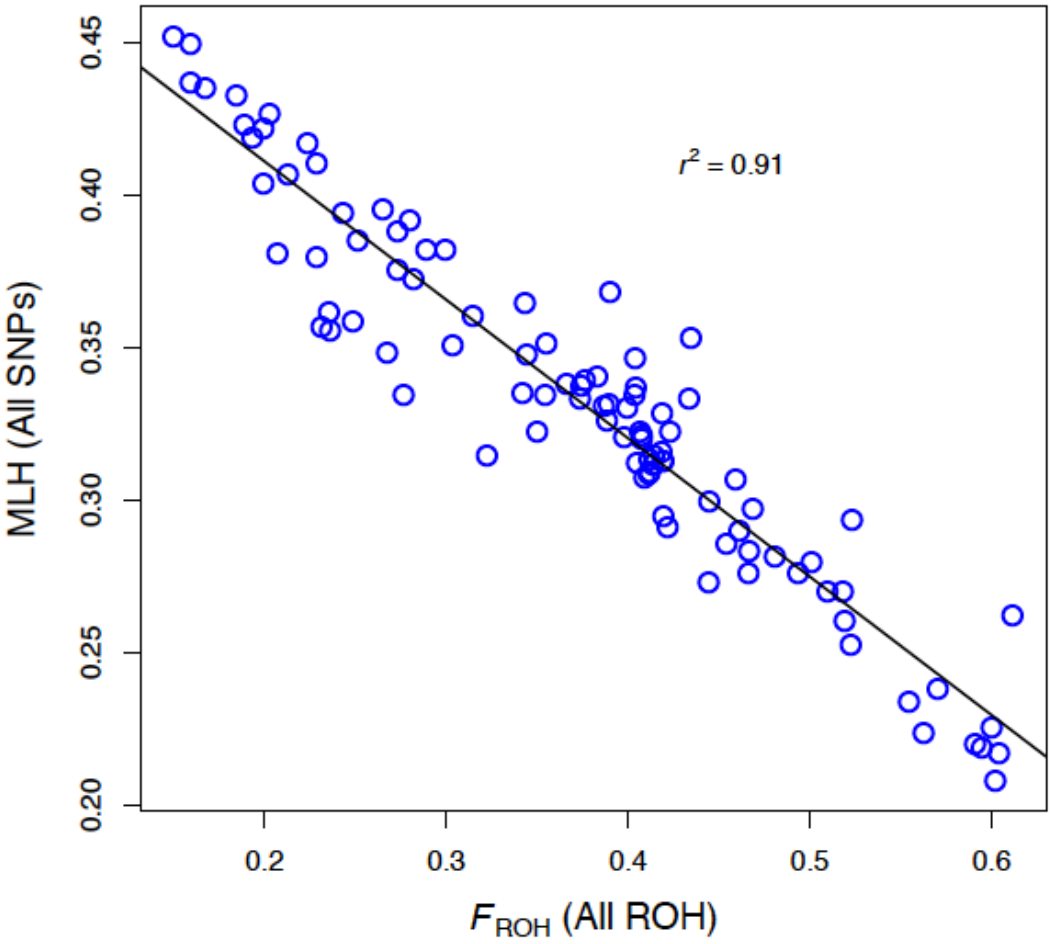


Figure S4. The relationship between MLH and F_{ROH} measured with the whole genome in Scandinavian wolves.

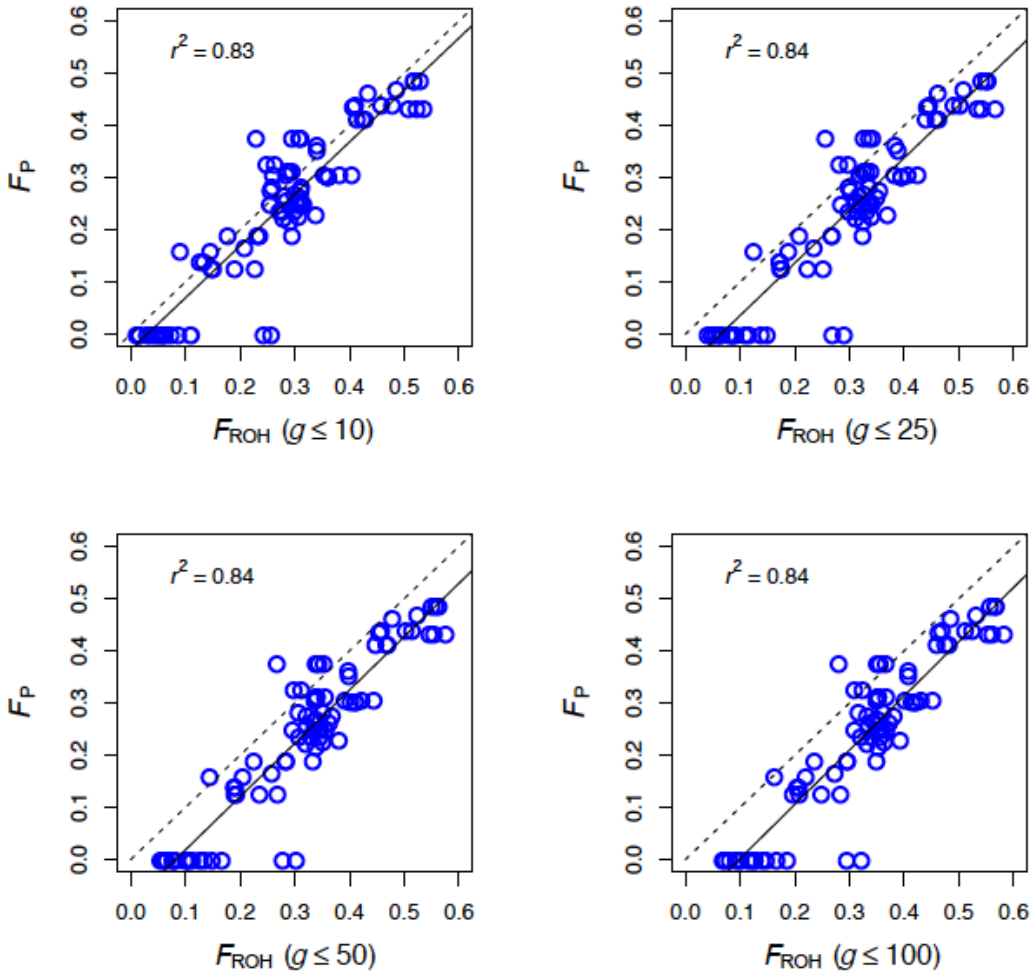


Figure S5. The relationship between F_P and F_{ROH} measured with the whole genome in Scandinavian wolves, after excluding immigrants from the data. F_P is shown on the y-axis, and F_{ROH} measured using ROH arising from relatively recent ancestors ($g \leq$ 10-100 generations). The dashed line has an intercept of zero and a slope of one. Points below the line thus represent cases where F_P underestimated F_{ROH}. The solid line is the fitted line from a regression of F_P versus F_{ROH}.

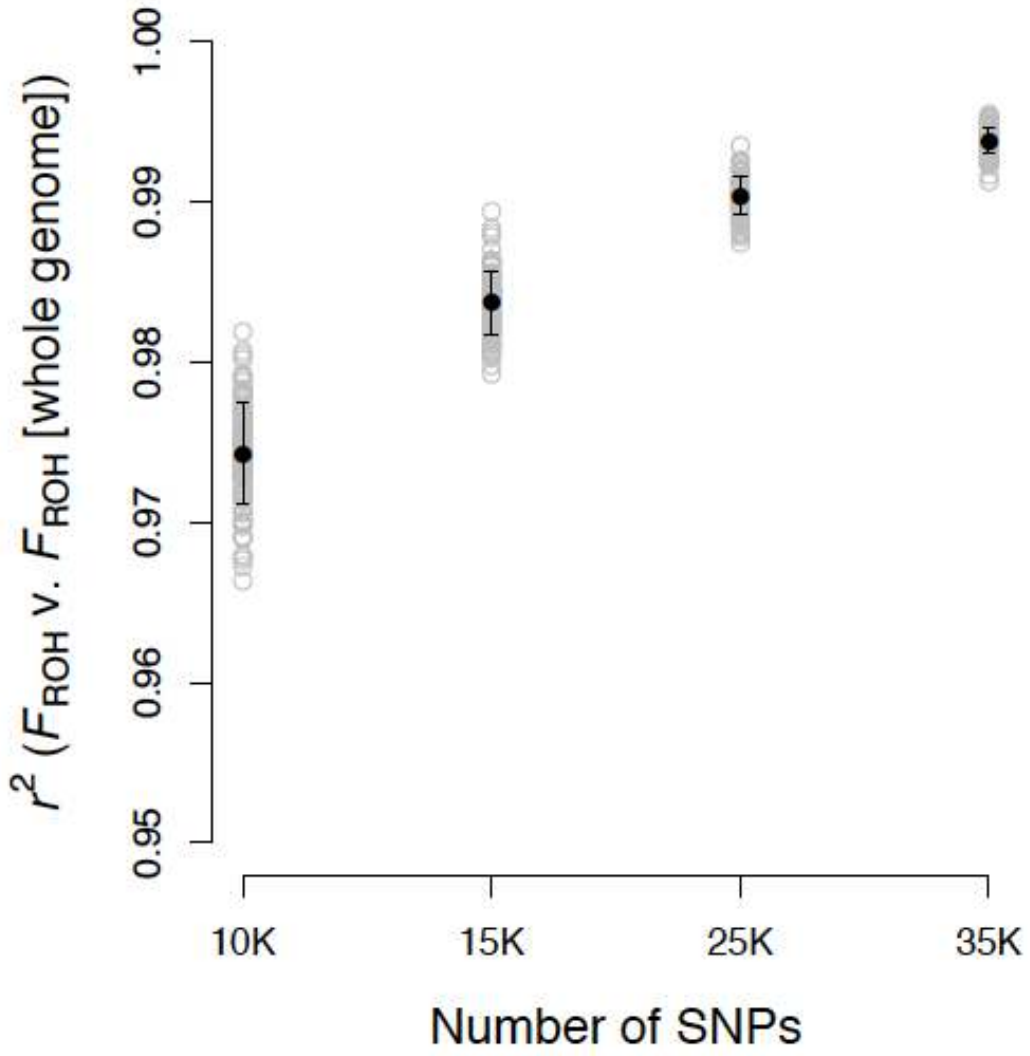


Figure S6. The precision of F_{ROH} as a function of the number of SNPs (x10^3) analyzed. The r^2 from regressions of F_{ROH} (estimated with subsamples of SNPs) versus F_{ROH} calculated from the whole genome (using ROH with $g \leq 10$) is shown on the y-axis. The number of subsampled SNPs is shown on the x-axis. The gray points represent the r^2 values from each of the 100 analyses for each number of SNPs. The black points and error bars represent the mean r^2 +/- one standard deviation of r^2 for each number of SNPs analyzed.

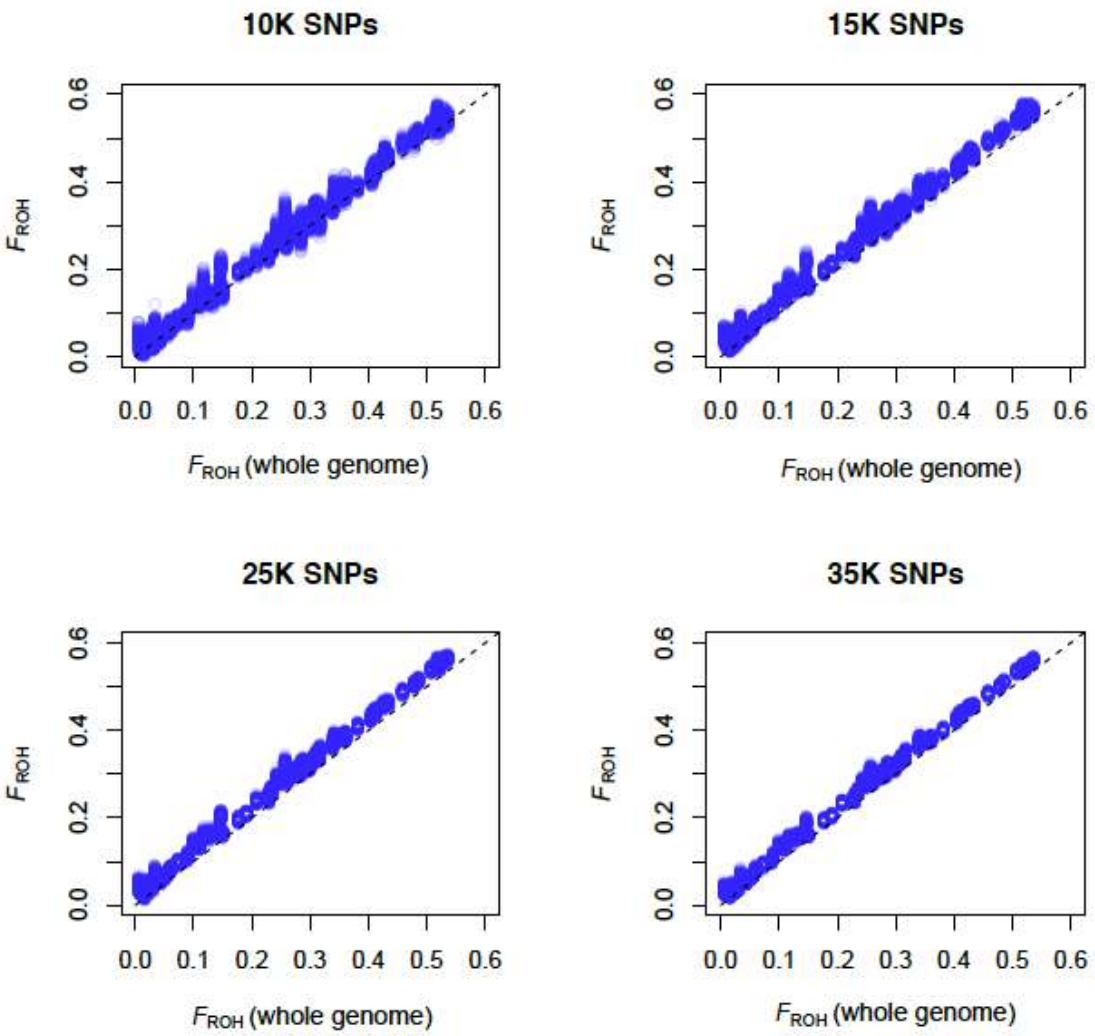


Figure S7. F_{ROH} measured with 10,000-35,000 subsampled SNPs plotted against F_{ROH} measured with the whole genome in Scandinavian wolves.

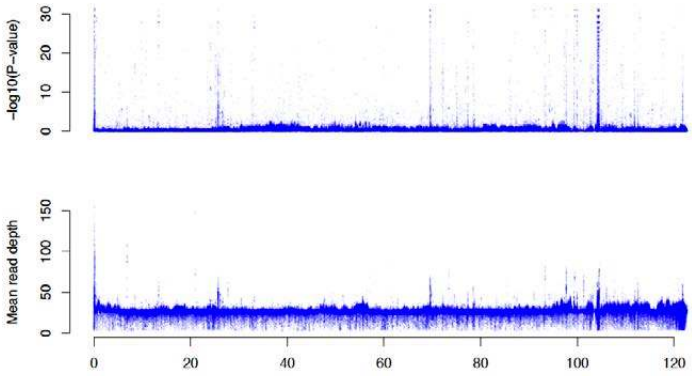


Figure S8. *P*-values from tests for excess of heterozygotes relative to Hardy-Weinberg proportions (top panel), and mean sequence read depth (lower panel), across all SNPs on chromosome 1.

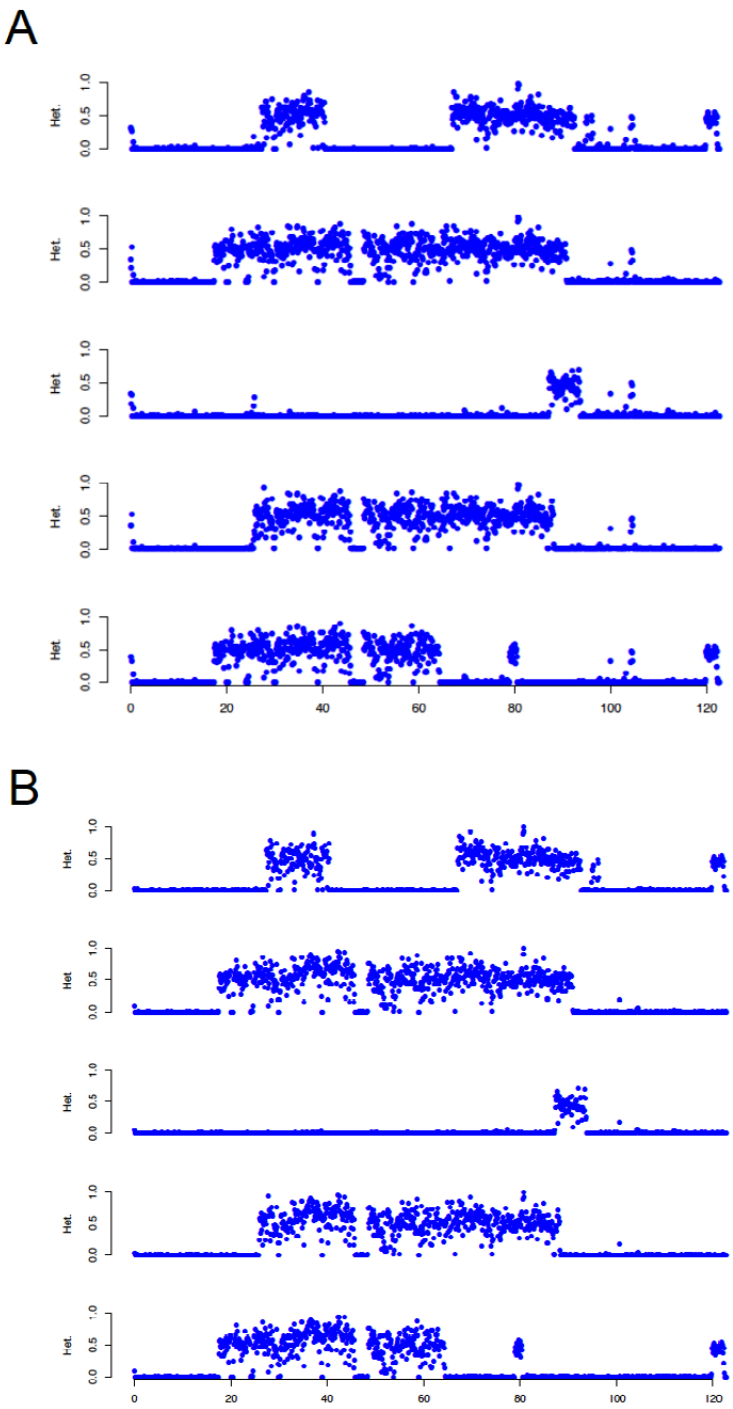


Figure S9. Heterozygosity across chromosome 1 in 100 kb windows for five wolves. The results shown are from analysis of data not filtered based on sequence read depth and deviation from Hardy-Weinberg proportions (**A**), and after filtering SNPs based on sequence read depth and deviation from Hardy-Weinberg proportions as described in the main text (**B**).

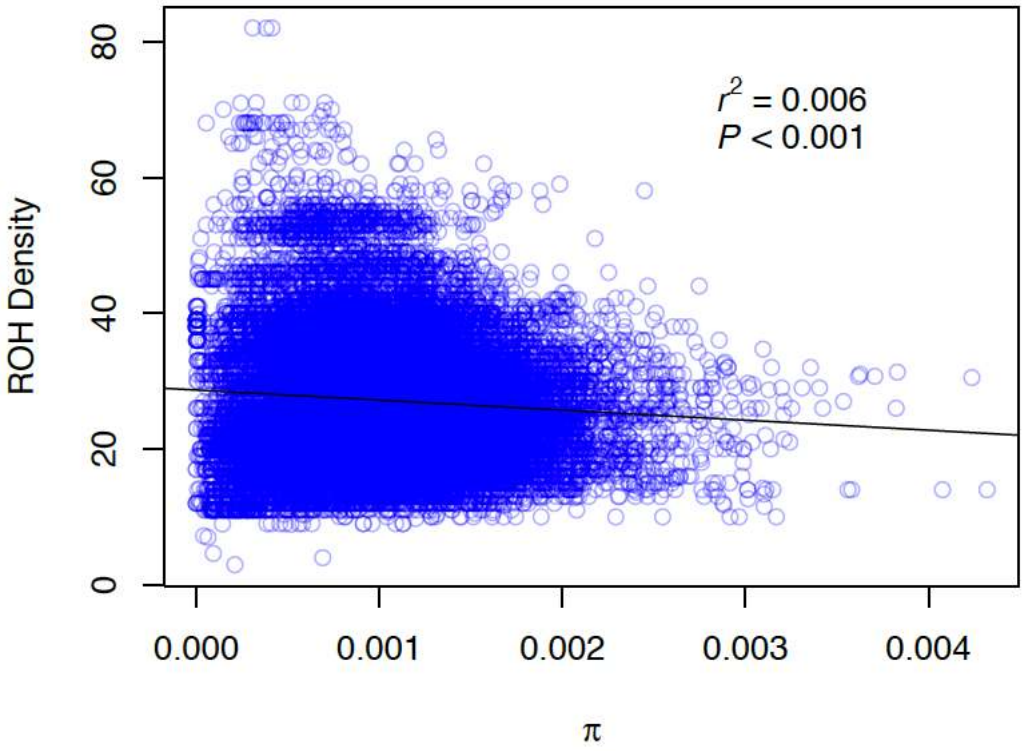


Figure S10. ROH density versus π measured in 100 kb windows of the 97 wolf genomes. The *P*-value, r^2, and solid black fitted line are from a linear regression model.

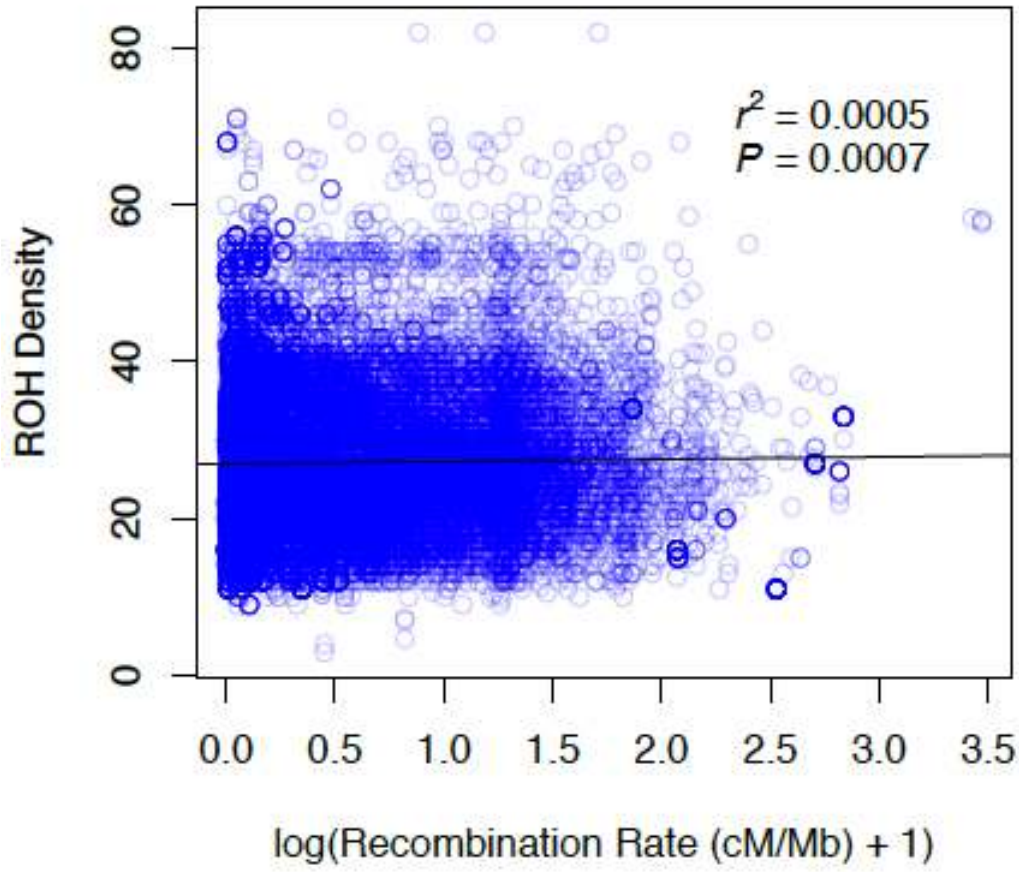


Figure S11. Scatterplot of ROH density versus the log-transformed recombination rate (cM/Mb) in 100 kb windows of the 97 wolf genomes. The *P*-value, r^2, and solid black fitted line are from a linear regression model.