1 # Genomic consequences of intensive inbreeding in an

2 # isolated wolf population

3

4

5 **Marty Kardos[1,2], Mikael Åkesson[3], Toby Fountain[1], Øystein Flagstad[4], Olof**

6 **Liberg[3], Pall Olason[5], Håkan Sand[3], Petter Wabakken[6], Camilla Wikenros[3],**

7 **and Hans Ellegren[1],***

8

9

10 1. Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala

11 University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden

12

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

14 59860

15

16 3. Department of Ecology, Grimsö Wildlife Research Station, Swedish University

17 of Agricultural Sciences, SE-730 91 Riddarhyttan, Sweden.

18

19 4. Norwegian Institute for Nature Research, P.O. Box 5685 Sluppen, Trondheim,

20 NO-7485, Norway.

21

22 5. Wallenberg Advanced Bioinformatics Infrastructure (WABI), Science for Life

23 Laboratory, Uppsala University, 75123 Uppsala, Sweden

24

25 6. Faculty of Applied Ecology and Agricultural Sciences, Campus Evenstad, Inland

26 Norway University of Applied Sciences, NO-2480, Norway

27

28

29

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

31

32

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.

63   Small populations are particularly vulnerable to extinction due to demographic

64   stochasticity, reduced genetic variation, and inbreeding depression[1-4].

65   Inbreeding (mating between relatives) in small populations can lead to

66   decreased individual fitness and population growth rate, owing to the expression

67   of deleterious recessive alleles and increased homozygosity at loci with

68   heterozygous advantage[3,5]. While inbreeding depression has long interested

69   biologists, its strength and genetic basis in the wild are still not well

70   understood[6,7]. A major challenge has been accurately measuring individual

71   inbreeding in natural populations.

72

73   Individual inbreeding has classically been estimated with the pedigree

74   inbreeding coefficient (F_P) for an individual using path analysis on a known

75   pedigree[3,8,9]. F_P predicts F, the fraction of an individual's genome that is identical-

76   by-descent (IBD), assuming that the pedigree founders, and any subsequent

77   immigrants are non-inbred and unrelated. However, not only are the necessary

78   multi-generation pedigrees difficult to obtain for most natural populations[10,11],

79   but F_P often imprecisely measures F because of the stochastic effects of

80   Mendelian segregation and linkage[7,12-17].

81

82   An alternative approach is to measure individual inbreeding indirectly by using

83   genetic markers to estimate multi-locus heterozygosity (MLH)[18-21], as the major

84   effect of inbreeding is to reduce the genome-wide heterozygosity of the

85   offspring[5]. This reduction occurs because related parents pass on IBD

86   chromosome segments that arise from a single chromosome copy in a shared

87   ancestor, with these segments characterized by long stretches of homozygous

88   genotypes (i.e., runs of homozygosity, ROH)[7]. MLH and similar statistics have the

89   advantage of not requiring a pedigree, but suffer from low precision when using

90   few loci[21-24].

91

92   High-throughput sequencing technologies can make it possible to measure

93   genome-wide heterozygosity using thousands of genetic markers[25-28].

94   Importantly, whole-genome resequencing in species with high quality genome

95   assemblies should facilitate the identification of IBD chromosome segments as

96   ROH, allowing the measurement of F as the fraction of the genome in long ROH

97   (F_{ROH}) with very little error[29]. Additionally, whole-genome resequencing of many

98   individuals from natural populations where high quality pedigrees are available

99   would allow rigorous empirical evaluation of how well F_P, MLH and F_{ROH} based

100 on smaller number of loci perform as estimators of F.

101

102 Here, we resequenced 97 genomes sampled from a semi-isolated and

103 bottlenecked wolf population in Scandinavia, This population is of high

104 conservation concern and has been subject to long-term studies of inbreeding,

105 inbreeding depression and genetic rescue[30-34]. Importantly, the population

106 represents a rare example of having a nearly complete pedigree available[30]. First,

107 we sought to identify IBD chromosome segments and to quantify F among

108 individuals in the population. Second, we evaluated the statistical performance of

109 F_P, MLH, and F_{ROH} as measures of F. Finally, we searched for regions of the

110 genome that may harbor alleles with large phenotypic effects contributing to

111 inbreeding depression by scanning for chromosome segments where ROH were

112 exceptionally rare or absent. To our knowledge, this is the first study combining

113 whole-genome resequencing and pedigree information to study individual

114 inbreeding in the wild.

115

116

117 **Results**

118

119 *Study population, pedigree, and whole-genome resequencing*

120

121 After a long period of population decline, wolves became functionally extinct

122 from the Scandinavian Peninsula in 1960-1970s[35]. The contemporary

123 Scandinavian wolf population was founded by two individuals in the early

124 1980s[33,36], and is characterized by prolonged periods of isolation with only rare

125 reproductively successful immigrants[30,37]. We sampled 97 wolves from

126 Scandinavia between 1977 and 2015, including 12 immigrants of which five

127 were founders of the population. These individuals were chosen to represent the

128 range of observed F_P values in the population, which were derived from a

129   pedigree extending back to the first breeding event in 1983[34]. F_P ranged from

130   zero (for 12 immigrants and 19 Scandinavian-born offspring to immigrant

131   founders) up to F_P = 0.49 for three Scandinavian-born siblings sampled after the

132   population had experienced a prolonged period of isolation. The number of

133   generations of pedigree known for each individual is given in Table S1.

134

135   We performed whole-genome resequencing of all wolves at a mean sequence

136   read depth of 27.4 (s.d. = 10.3). After variant calling, we performed SNP filtering

137   based on genotype qualities, read depth, deviation from Hardy-Weinberg

138   genotype proportions, missing data, and minor allele frequency (MAF; see

139   Methods). Mean MAF was 0.17 at 10,688,886 SNPs remaining before filtering

140   based on allele frequency. After filtering based on allele frequency, the mean

141   MAF was 0.26 (s.d. = 0.13) at 6,701,147 SNPs. Given that almost one hundred

142   individuals were sequenced, the number of detected variants is low for a large

143   mammalian genome. However, low genetic diversity is expected given the small

144   population size and limited number of founders. Moreover, nucleotide diversity

145   estimated from the 12 immigrants was 0.001, which is in the lower end of what

146   has been reported among other vertebrates.

147

148   *Runs of homozygosity (ROH)*

149

150   We identified ROH (putative IBD chromosome segments) in the whole-genome

151   resequencing data using a likelihood ratio-based sliding window method which

152   accounts for SNP allele frequencies and sequencing errors[29,38]. We detected a

153   total of 269,309 ROH among the 97 wolves, ranging from 0 to 76.6 cM in genetic

154   map length, and from 2,695 bp to 95.8 Mb in physical length (Figure 1, Figure S1.

155   Describing ROH by genetic map length is motivated by the fact that

156   recombination determines the size of IBD segments. Additionally, our theoretical

157   understanding of the expected lengths of ROH, and of the variance of F around

158   pedigree expectations is in terms of ROH genetic map lengths[12,17,39]. The choice

159   of using physical versus genetic mapping coordinates of ROH had nearly no

160   effect on genomic estimates of inbreeding (Figure S2). Notably, many individuals

161   had ROH spanning either entire or nearly entire chromosomes, giving extreme

162   patterns with a complete lack of heterozygosity over large parts of the genome

163   (Figures 2-3, Supplementary File 1).

164

165   Though there were many strikingly large ROH (Figures 2-3), most were very

166   short. Specifically, more than 50% of ROH were less than 0.02 cM long (Figure 1)

167   and these represent IBD segments that generally arise from ancestors in deep

168   history. We estimated the number of generations (g) back to the common

169   ancestor of the two homologous sequence copies for each ROH based on its map

170   length. The very short ROH (≤0.02 cM long) are expected to arise an average

171   from ancestors ≥2,500 generations ago (i.e., g = 2,500 for a 0.02 cM ROH; see

172   Methods); 2,500 generations correspond to 10,000 years assuming a four-year

173   generation interval for wolves. Yet, the highly abundant, short ROH contributed

174   little to total IBD. For example, segments shorter than 0.02 cM represented only

175   1.3% of all IBD chromosome regions in the 97 wolves (Figure 1, Supplementary

176   File 1). In contrast, the less frequent but very long ROH arising from recent

177   ancestors accounted for the majority of total IBD sequence.

178

179   *Genomic measures of inbreeding*

180

181   We measured individual inbreeding as the proportion of the genome that was in

182   ROH (F_{ROH}) identified in the whole-genome resequencing data. F_{ROH} is an

183   estimator of the realized IBD fraction of the genome and was obtained using only

184   long ROH (i.e., ROH with small g-values). We conducted separate analyses using

185   different maximum values of g (10, 25, 50, and 100 generations) for the ROH

186   included in estimates of F_{ROH}. This ensured that we measured inbreeding due to

187   recent ancestors while also allowing us to evaluate the sensitivity of the results

188   to different maximum values of g. Including very short ROH would have meant

189   that F_{ROH} captured inbreeding due to distant ancestors, which is less likely to be

190   important to inbreeding depression because at least some deleterious alleles are

191   expected to be purged over long time spans[38,40].

192

193   There was a large range of F_{ROH} in the population. F_{ROH} measured using ROH with

194   g ≤ 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

7

228 *Performance of F_P and molecular measures of individual inbreeding*

229

230 We used linear regression to evaluate the statistical performance of F_P, F_{ROH}, and

231 *MLH* as predictors of realized individual inbreeding. F_{ROH} measured with the

232 whole genome is equivalent to F, and the same applies to *MLH* with respect to H.

233 F_P was strongly correlated (r^2 = 0.86-0.87) with F_{ROH} (Figure 4). The linear

234 regression of F_P versus F_{ROH} had a slope of 1.0 and an intercept of –0.03 when

235 F_{ROH} was measured with only the longest ROH ($g \leq 10$). The negative intercept

236 shows that F_P was a downwardly biased measure of F_{ROH}, and the slope of 1.0

237 shows that the size of the downward bias was constant on average across the

238 range of observed F_{ROH} values (Figure 4). The correlations between F_P and F_{ROH}

239 were only slightly weaker (r^2 = 0.83 to 0.84), and the slopes and intercepts were

240 unchanged, when immigrants were excluded from this analysis (Figure S5). The

241 choice of a maximum value of g for the ROH included in the measurement of F_{ROH}

242 did not substantively affect the correlation between F_P and F_{ROH}, but the

243 magnitude of the downward bias in F_P increased with higher values of threshold

244 of g (Figure 4). This makes sense as F_{ROH} calculated using ROH with larger values

245 of g captures inbreeding due to more distant ancestors.

246

247 The high variation in F_{ROH} among individuals with F_P = 0 weakened the precision

248 of F_P. Specifically, a combination of some highly inbred individuals and

249 individuals with F_{ROH} near zero clearly decreased the variance in F_{ROH} explained

250 by F_P (Figure 4). F_P is likely to have higher precision in populations with less

251 variation in F_{ROH} among founders and immigrants. An obvious strength of

252 genomic measures of individual inbreeding is that they do not require making *a*

253 *priori* assumptions regarding the inbreeding or relatedness of any individuals.

254

255 The relatively high precision of F_P as a measure of individual inbreeding

256 observed here (compared to previous simulation results[27]) is expected.

257 Theoretical and simulation-based investigations have shown that the precision

258 of F_P as a measure of F depends strongly on the number of chromosomes, the

259 recombination rate, and the distribution of recombination events across the

260 genome [5,12,14,16,39]. Canids have a large number of chromosomes (38 autosomes).

261 Thus, F_P is expected to be more precise in wolves compared to species with

262 fewer chromosomes, as long as pedigrees are deep and complete enough to

263 capture the great majority of recent common ancestors of parents. The high

264 variance in individual inbreeding in this study also must have contributed to the

265 high r^2 from a regression of F_P versus F_{ROH}. We sampled from throughout the

266 range of F_P values observed in the population, which resulted in a higher

267 variance in F_P among the selected wolves ($\sigma^2 = 0.026$) relative to the population

268 as a whole ($\sigma^2 = 0.006$). This is expected to have increased the correlation of

269 realized genomic inbreeding with F_P and the molecular inbreeding measures

270 based on subsampled SNPs in the sampled wolves compared to the population as

271 a whole. All else equal, a lower correlation of F with F_P and molecular measures

272 of inbreeding is expected in populations with lower variance in F[24,42].

273

274 *Performance of MLH as a measure of individual inbreeding*

275

276 To evaluate the precision of *MLH* as a measure of *H*, we randomly subsampled

277 between 50 and 20,000 SNPs from the genome. For each subsampled set of loci, a

278 linear regression model with *MLH* measured from the subsampled loci was fitted

279 as the response variable, and *MLH* measured with the whole genome as the

280 predictor variable. We then used r^2 from these regression models as a measure

281 of the precision of *MLH*. To ensure that the subsamples were drawn as

282 independently as possible from the genome, no locus was used in more than one

283 of the 100 subsamples for each number of loci analyzed.

284

285 The mean r^2 between *MLH* based on subsampled loci and *MLH* from the whole

286 genome was 0.88 when 500 SNPs were used, and ≥ 0.94 when 1,000 or more

287 SNPs were used (Figure 5). *MLH* and other measures of individual inbreeding are

288 expected to have high precision when the variance in *F* is as high as it was in this

289 study[22]. The correlation between *MLH* based on subsampled loci and *MLH*

290 measured with the whole genome matches theoretical expectations remarkably

291 well. For example, the expected correlation between *MLH* (estimated with 500

292 loci) and realized genome-wide heterozygosity is 0.87 according to the analytical

293 results of Miller et al.[22], very close to the observed r^2 of 0.88. This is highly

9

294   encouraging for studies of natural populations where pedigrees, mapped loci,

295   and large-scale SNP genotyping arrays or whole-genome resequencing data are

296   unavailable. This is also empirical evidence that individual inbreeding can be

297   more precisely measured with a modest number of molecular markers than with

298   pedigrees[14,27].

299

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

301

302   We used the same subsampling and regression approach applied above for *MLH*

303   to evaluate the performance of F_{ROH}. However, for F_{ROH}, we used subsamples of

304   10,000 SNPs and larger, and the predictor variable in the regression models was

305   F_{ROH} measured with the whole genome. F_{ROH} estimated with as few as 10,000

306   SNPs was strongly correlated with F_{ROH} estimated with the whole genome (mean

307   $r^2 = 0.97$ [s.d. = 0.003] among 100 replicates, Figure S6). F_{ROH} estimated with

308   subsampled SNPs was slightly upwardly biased (Figure S7). This bias was likely

309   caused by overestimating the length of real IBD segments, or by incorrectly

310   calling an ROH where no true IBD segment existed when using relatively few loci.

311   We therefore urge caution when interpreting results of ROH analyses (e.g., for

312   estimating individual inbreeding or mapping loci responsible for inbreeding

313   depression) when only tens of thousands of loci are used.

314

315   *Detecting genomic regions that may contribute to inbreeding depression*

316

317   Alleles that strongly reduce fitness when homozygous (i.e., either strongly

318   deleterious recessive or overdominant alleles) are likely to cause ROH to be

319   absent or exceptionally rare in the local chromosomal vicinity [7,43,44]. We

320   quantified the abundance of ROH with values of $g \leq 50$ in non-overlapping 100 kb

321   windows across all 38 autosomes and used a permutation approach to test for

322   regions with lower than expected abundance of ROH given a random distribution

323   of ROH across the genome (see Methods for details). Ten such regions were

324   found on chromosomes 3, 11, 14, 16, 20, 21, and 22 (Figure 6, Table S2). Thus, it

325   appears that several genomic regions likely contained loci with strong enough

326   deleterious fitness effects when homozygous to substantially reduce the

327 frequency of individuals carrying IBD segments in these regions. Repeating this

328 analysis with different ROH length thresholds ($g \leq 10$, 20 or 100) did not

329 substantively change the results (results not shown). As in many types of

330 genomic analysis, it is possible that technical artifacts such as genome assembly

331 errors or incorrectly mapped sequence reads could have contributed to some of

332 the regions with low ROH abundance. These genomic regions should therefore

333 be analyzed in further detail, including genotyping or sequencing of larger

334 population samples.

335

336

337 **Discussion**

338

339 This study is one of the first large-scale examples of the power of genome

340 resequencing to record the genomic consequences of inbreeding in a population

341 of conservation concern. The combination of a huge number of SNPs resulting

342 from the whole-genome resequencing of 97 individuals and a high-quality

343 genome assembly enabled us to precisely delineate IBD chromosome segments

344 as ROH, and to quantify realized genomic inbreeding, and to identify genomic

345 regions that likely contributed substantially to inbreeding depression in this

346 vulnerable population of Scandinavian wolves. In many individuals, the

347 signatures of inbreeding were remarkably visible as entire or nearly entire

348 chromosomes were completely homozygous (Figures 2 and 3).

349

350 Our results demonstrate that the vast majority of IBD segments in a recently

351 bottlenecked population are actually very short and originate from common

352 ancestors in the far past. However, quantitatively these short IBD segments

353 contributed little to individual F_{ROH}, which was primarily governed by more

354 limited numbers of very long segments resulting from common ancestors of

355 parents less than 10 generations ago. Still, while F_P correlated well with F_{ROH}

356 over a range of time spans to common ancestors, it becomes an increasingly

357 downward biased estimator of F_{ROH} as older IBD segments are taken into

358 account.

359

360 Our results also provide empirical evidence based on large-scale whole-genome

361 resequencing that inbreeding is better measured with molecular genetic data

362 than with F_P estimated from an extensive pedigree. . While several previous

363 studies have assessed correlations among molecular measures of inbreeding and

364 F_p[25,26,28], none have rigorously evaluated the performance of F_p and molecular

365 measures of inbreeding because the true realized genomic inbreeding was

366 unknown[7]. To our knowledge, this is the first study to carry out such an analysis.

367 F_P has been the standard measure of individual inbreeding for decades[10]. While

368 pedigrees are clearly still useful for estimating inbreeding (e.g., in species with

369 many chromosomes[12]), and for many other purposes[10], molecular measures of F

370 are more powerful as they account for related and inbred pedigree founders and

371 immigrants, and the stochastic effects of linkage and Mendelian segregation.

372 Additionally, molecular approaches allow mapping of loci contributing to

373 inbreeding depression[5,44]. An interesting question that arises from our

374 observations and that should be investigated further is the overall phenotypic

375 consequences of individuals within a population being IBD for different

376 haplotypes of very large chromosome segments. One might expect that this will

377 disclose 'hidden' phenotypic variation encoded by rare variants or variation that

378 is otherwise rarely seen due to dominance effects.

379

380 The demonstration of inbreeding and relatedness among immigrants has

381 important implications for population viability and the design of management

382 programs. In the case of the Scandinavian wolf population, having inbred and

383 related immigrants means that animals are on average more inbred than it

384 appears based on pedigree information alone (Figure 4). This emphasizes the

385 importance of immigration into the population to limit inbreeding and

386 inbreeding depression. Also, it highlights the importance of taking the genetic

387 status (i.e., the degree of inbreeding and relatedness arising from finite

388 population size and population fragmentation) of a larger metapopulation into

389 account. Importantly, a similar situation may apply to many other species of

390 conservation concern where a fragmented population structure increases the

391 likelihood for inbreeding and close relatedness among immigrants[26].

392

393 Identifying regions of the genome with exceptionally low abundance of ROH is an
394 important step towards understanding the genetic basis of inbreeding
395 depression in Scandinavian wolves. These genomic regions are likely to contain
396 loci with overdominant or deleterious recessive alleles strongly contributing to
397 inbreeding depression. Future mapping studies could be used to directly test for
398 phenotypic effects of IBD in these regions. Ascertaining the loci underlying
399 inbreeding depression and the magnitude of their phenotypic effects is crucial to
400 advancing our understanding of the genetic basis of inbreeding depression and
401 the potential for purging to lessen the genetic load.
402
403
404 **Methods**
405
406 *Study population and DNA samples*
407
408 As in many other parts of the world[45], the wolf experienced a significant
409 population decline in Scandinavia during the last centuries. Once common and
410 spread over the entire Scandinavian peninsula, hunt and persecution eventually
411 led to the functional extinction of wolves in the 1960-70s[35]. The closest surviving
412 populations were found in eastern Finland (where it was rare) and western
413 Russia. The Scandinavian population was subsequently re-established in the
414 early 1980s by a single mating pair that are likely to have had an eastern
415 origin[32,36]. The founder female was killed in 1985 and the founding male
416 disappeared one year later. Subsequent breeding 1987-1990 consisted of
417 successive mating between sibling and parent-offspring pairs resulting in severe
418 inbreeding[30,33,34]. A third (male) founder immigrated and reproduced in the
419 population in 1991-1993 but no further successful immigration occurred until
420 2008, after which five reproductively successful immigrants have entered
421 Scandinavia from the Finnish-Russian population[30,36,46]. Before the arrival of the
422 third founder, there was only one reproducing pack and likely no more than 10
423 wolves in the population. The immigrant male in 1991 had very high
424 reproductive success and the population subsequently grew to around 365
425 (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

14

455   phenol/chloroform–isoamylalcohol extraction and the precipitate was solved in

456   20-100 μl distilled water.

457

458   *Whole-genome resequencing and variant calling and filtering*

459

460   Library construction and 150 basepair paired-end sequencing was performed on

461   an Illumina HiSeqX with standard procedures. Sequencing reads were mapped to

462   the dog genome build CanFam3.1, using BWA v0.7.13[47]. The resulting BAM files

463   were sorted using SAMtools v1.3[48], duplicate marked using Picard v1.118

464   (http://broadinstitute.github.io/picard/), and locally realigned around indels

465   using GATK v3.3.0[49,50]. Read information was updated in the bam files with

466   Picard FixMateInformation.

467

468   A first round of variant calling was performed with GATK HaplotypeCaller and

469   the whole cohort genotyped using GATK GenotypeGVCFs. The resulting variant

470   list was filtered for low quality variants with low allele frequency using bcftools

471   v1.3 (http://samtools.github.io/bcftools/) (filtering criteria: INFO/AF < 0.01 &&

472   INFO/MQRankSum<-0.2). The variants passing this filter were used as a true

473   positive set of variant sites for BQSR, performed with GATK. Variant calling was

474   repeated for the recalibrated bam files and then the whole cohort re-genotyped

475   using GATK.

476

477   We applied several SNP filters to ensure high quality of the data. First, all tri-

478   allelic loci, loci with only heterozygous or only homozygous genotypes, and loci

479   with mean read depth (among all 97 individuals) of less than 10 or greater than

480   52 (twice the mean sequence read depth genome-wide) were removed. Second,

481   genotypes with Phred-scaled genotype quality scores of less than 20 and loci that

482   had missing genotypes in >15 individuals were discarded. We then removed loci

483   where the *P*-value was <0.001 in a test for an excess of heterozygotes relative to

484   Hardy-Weinberg genotype proportions using the --*hardy* function in VCFtools[50].

485   Finally, we retained only loci with minor allele frequency ≥0.05. The

486   heterozygote excess and read depth filters were successful at removing SNPs in

487   regions with poor read mapping (Figures S8-S9).

Inferring SNP linkage map positions

489

490   The genetic map position (in cM) of each SNP in the wolf whole-genome

491   resequencing data were inferred from a recent sex-averaged high-density

492   domestic dog linkage map[51]. This was done by first identifying the closest

493   upstream and downstream SNP included in the dog map. We then interpolated

494   the genetic position of the focal SNP while assuming that the recombination rate

495   was constant between the two flanking linkage-mapped SNPs[29].

496

497   *Quantifying individual inbreeding*

498

499   The pedigree was determined using parentage information derived from field

500   observations and microsatellite-based parentage assignments as described

501   previously[30,34]. F_P was calculated using CFC v1.0 software[52]. To estimate F_{ROH}, we

502   identified ROH using a likelihood ratio method[17,29,38]. First, we split each

503   chromosome up into sliding windows that each included 100 adjacent SNPs,

504   using a step size of 10 SNPs. For each 100-SNP window i, and individual j, we

505   calculated the probability (Pr) of the genotype at each SNP k (G_k) assuming the

506   SNP was IBD, and separately assuming the SNP was non-IBD. We then calculated

507   a LOD score by summing the $\log_{10}$ of the ratio of these probabilities across all loci

508   within the window:

509

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

510

511

512   The genotype probabilities under IBD and non-IBD were calculated according to

513   Wang *et al.*[17], accounting for occasional heterozygous positions within ROH

514   resulting from sequencing errors, read mapping errors (e.g., due to segmental

515   duplications), and occasional mutations. Specifically, we accepted that 2% of

516   SNPs would be heterozygous within IBD segments. Using shorter window sizes

517   (40- and 60-SNP windows) resulted in obvious IBD segments being artificially

518   broken in regions with poor mapping of sequence reads (results not shown).

519   Likewise, assuming fewer heterozygous positions within IBD segments (e.g.,

520 0.1%) artificially broke up obvious IBD segments in regions with apparent poor

521 mapping of sequence reads.

522

523 We estimated g for each ROH in order to include only IBD segments arising from

524 recent ancestors when estimating F_{ROH}. For each ROH, we solved for g in the

525 equation $l = 100/2g$ cM, where l is the length of the ROH in cM[39]. We estimated

526 the map length of each ROH in cM by interpolating the mapping positions of each

527 SNP in the genome from a recent high-density linkage map of the domestic dog

528 genome[51], assuming that the recombination rate is conserved between domestic

529 dogs and wolves.

530

531 ***Permutation test for regions with exceptionally low ROH abundance***

532

533 We used a permutation (randomization) approach to simulate the null

534 distribution of ROH abundance in 100 kb windows. For each of 5,000,000

535 permutations, we first randomly sampled 97 individuals with replacement from

536 the sequenced wolves. We then randomly selected a 100 kb chromosome

537 segment from the genome of each individual independently. We then quantified

538 ROH abundance for the segment as the sum of the lengths of all IBD parts of the

539 97 sampled chromosome segments (in kb) divided by the length of the segment

540 (100 kb). A P-value for each 100 kb segment in the genome was calculated as the

541 proportion of the 5,000,000 permuted ROH abundance estimates that were

542 smaller than the observed ROH abundance. The P-value was set to 1/5,000,001

543 for segments where none of the 5,000,000 permutation repetitions produced an

544 ROH abundance ≤ the observed value. We used the Bonferroni method to correct

545 for multiple testing. Specifically, the P-value below which a test was considered

546 statistically significant was set to 0.05 divided by 22,055 (the number of

547 analyzed 100 kb windows).

548

549 ROH abundance has previously been strongly related to the recombination rate

550 and SNP density in other taxa (e.g., humans and birds), with low ROH abundance

551 found in regions with high recombination rate and/or relatively low SNP

552 density[29,43]. We tested for such effects in the present study to determine if

17

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.

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).

625   8   Malécot, G. *The Mathematics of Heredity*. (W. H. Freeman, 1970).
626   9   Wright, S. Coefficients of inbreeding and relationship. *The American*
627   *Naturalist* **56**, 330-338 (1922).
628   10   Pemberton, J. M. Wild pedigrees: the way forward. *Proceedings of the*
629   *Royal Society of London. Series B: Biological Sciences* **275**, 613-621
630   (2008).
631   11   Taylor, H. R., Kardos, M. D., Ramstad, K. M. & Allendorf, F. W. Valid
632   estimates of individual inbreeding coefficients from marker-based pedigrees are
633   not feasible in wild populations with low allelic diversity. *Conservation Genetics*
634   **16**, 901-913 (2015).
635   12   Franklin, I. R. The distribution of the proportion of the genome which is
636   homozygous by descent in inbred individuals. *Theoretical Population Biology* **11**,
637   60-80 (1977).
638   13   Hedrick, P. W., Kardos, M., Peterson, R. O. & Vucetich, J. A. Genomic
639   variation of inbreeding and ancestry in the remaining ywo Isle Royale wolves.
640   *Journal of Heredity* **108**, 120-126 (2017).
641   14   Knief, U., Kempenaers, B. & Forstmeier, W. Meiotic recombination shapes
642   precision of pedigree- and marker-based estimates of inbreeding. *Heredity* **118**
643   (2017).
644   15   Forstmeier, W., Schielzeth, H., Mueller, J. C., Ellegren, H. & Kempenaers, B.
645   Heterozygosity–fitness correlations in zebra finches: microsatellite markers can
646   be better than their reputation. *Molecular Ecology* **21**, 3237-3249 (2012).
647   16   Fisher, R. A. *The Theory of Inbreeding*. (Oliver and Boyd, 1965).
648   17   Wang, J. Pedigrees or markers: Which are better in estimating relatedness
649   and inbreeding coefficient? *Theoretical Population Biology* **107**, 4-13 (2016).
650   18   Leary, R. F., Allendorf, F. W. & Knudsen, K. L. Developmental stability and
651   enzyme heterozygosity in rainbow trout. *Nature* **301**, 71-72 (1983).
652   19   Pierce, B. A. & Mitton, J. B. Allozyme heterozygosity and growth in the
653   tiger salamander, Ambystoma tigrinum. *Journal of Heredity* **73**, 250-253 (1982).
654   20   Coltman, D. W., Pilkington, J. G., Smith, J. A. & Pemberton, J. M. Parasite-
655   mediated selection against inbred Soay Sheep in a free-living, island population.
656   *Evolution* **53**, 1259-1267 (1999).
657   21   Slate, J. *et al.* Understanding the relationship between the inbreeding
658   coefficient and multilocus heterozygosity: theoretical expectations and empirical
659   data. *Heredity* **93**, 255-265 (2004).
660   22   Miller, J. M. *et al.* Estimating genome-wide heterozygosity: effects of
661   demographic history and marker type. *Heredity* **112**, 240-247 (2014).
662   23   Balloux, F., Amos, W. & Coulson, T. Does heterozygosity estimate
663   inbreeding in real populations? *Molecular Ecology* **13**, 3021-3031 (2004).
664   24   Szulkin, M., Bierne, N. & David, P. Heterozygosity-fitness correlations: A
665   time for reapprisal. *Evolution* **64**, 1202-1217 (2010).
666   25   Huisman, J., Kruuk, L. E. B., Ellis, P. A., Clutton-Brock, T. & Pemberton, J. M.
667   Inbreeding depression across the lifespan in a wild mammal population.
668   *Proceedings of the National Academy of Sciences USA* **113**, 3585-3590 (2016).
669   26   Chen, N., Cosgrove, Elissa J., Bowman, R., Fitzpatrick, John W. & Clark,
670   Andrew G. Genomic consequences of population decline in the endangered
671   Florida Scrub-Jay. *Current Biology* **26**, 2974-2979 (2016)

672   27   Kardos, M., Luikart, G. & Allendorf, F. W. Measuring individual inbreeding
673   in the age of genomics: marker-based measures are better than pedigrees.
674   *Heredity* **115**, 63-72 (2015).
675   28   Hoffman, J. I. *et al.* High-throughput sequencing reveals inbreeding
676   depression in a natural population. *Proceedings of the National Academy of*
677   *Sciences, USA* **111**, 3775-3780 (2014).
678   29   Kardos, M., Qvarnström, A. & Ellegren, H. Inferring individual inbreeding
679   and demographic history from segments of identity by descent in *Ficedula*
680   flycatcher genome sequences. *Genetics* **205**, 1319-1334 (2017).
681   30   Åkesson, M. *et al.* Genetic rescue in a severely inbred wolf population.
682   *Molecular Ecology* **25**, 4745-4756 (2016).
683   31   Bensch, S. *et al.* Selection for heterozygosity gives hope to a wild
684   population of inbred wolves. *PLoS ONE* **1**, e72 (2006).
685   32   Flagstad, Ø. *et al.* Two centuries of the Scandinavian wolf population:
686   patterns of genetic variability and migration during an era of dramatic decline.
687   *Molecular Ecology* **12**, 869-880 (2003).
688   33   Vilà, C. *et al.* Rescue of a severely bottlenecked wolf (*Canis lupus*)
689   population by a single immigrant. *Proceedings of the Royal Society of London.*
690   *Series B: Biological Sciences* **270**, 91-97 (2003).
691   34   Liberg, O. *et al.* Severe inbreeding depression in a wild wolf (*Canis lupus*)
692   population. *Biology Letters* **1**, 17-20 (2005).
693   35   Haglund, B. De stora rovdjurens vintervanor II. *Viltrevy* **5**, 213-361
694   (1969).
695   36   Wabakken, P., Sand, H., Liberg, O. & Bjärvall, A. The recovery, distribution,
696   and population dynamics of wolves on the Scandinavian peninsula, 1978-1998.
697   *Canadian Journal of Zoology* **79**, 710-725 (2001).
698   37   Seddon, J. M., Sundqvist, A. K., Björnerfeldt, S. & Ellegren, H. Genetic
699   identification of immigrants to the Scandinavian wolf population. *Conservation*
700   *Genetics* **7**, 225-230 (2006).
701   38   Szpiech, Z. A. *et al.* Long runs of homozygosity are enriched for
702   deleterious variation. *The American Journal of Human Genetics* **93**, 90-102
703   (2013).
704   39   Thompson, E. A. Identity by descent: variation in meiosis, across genomes,
705   and in populations. *Genetics* **194**, 301-326 (2013).
706   40   García-Dorado, A. Understanding and predicting the fitness decline of
707   shrunk populations: inbreeding, purging, mutation, and standard selection.
708   *Genetics* **190**, 1461-1476 (2012).
709   41   Crow, J. F. & Kimura, M. *An Introduction to Population Genetics Theory*.
710   (The Blackburn Press, 1970).
711   42   Kardos, M., Allendorf, F. W. & Luikart, G. Evaluating the role of inbreeding
712   depression in heterozygosity-fitness correlations: how useful are tests for
713   identity disequilibrium? *Molecular Ecology Resources* **14**, 519-530 (2014).
714   43   Pemberton, T. J. *et al.* Genomic patterns of homozygosity in worldwide
715   human populations. *The American Journal of Human Genetics* **91**, 275-292
716   (2012).
717   44   Hedrick, P. W., Hellsten, U. & Grattapaglia, D. Examining the cause of high
718   inbreeding depression: analysis of whole-genome sequence data in 28 selfed
719   progeny of *Eucalyptus grandis*. *New Phytologist* **209**, 600-611 (2016).

720 45 Chapron, G. *et al.* Recovery of large carnivores in Europe's modern
721 human-dominated landscapes. *Science* **346**, 1517 (2014).
722 46 Laikre, L., Olsson, F., Jansson, E., Hossjer, O. & Ryman, N. Metapopulation
723 effective size and conservation genetic goals for the Fennoscandian wolf (*Canis*
724 *lupus*) population. *Heredity* **117**, 279-289 (2016).
725 47 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–
726 Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
727 48 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools.
728 *Bioinformatics* **25**, 2078-2079 (2009).
729 49 McKenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework
730 for analyzing next-generation DNA sequencing data. *Genome Research* **20**, 1297-
731 1303 (2010).
732 50 Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**,
733 2156-2158 (2011).
734 51 Campbell, C. L., Bhérer, C., Morrow, B. E., Boyko, A. R. & Auton, A. A
735 pedigree-based map of recombination in the domestic dog genome. *G3:*
736 *Genes|Genomes|Genetics* **6**, 3517 (2016).
737 52 Sargolzaei, M., Iwaisaki, H. & Colleau, J. J. A fast algorithm for computing
738 inbreeding coefficients in large populations. *Journal of Animal Breeding and*
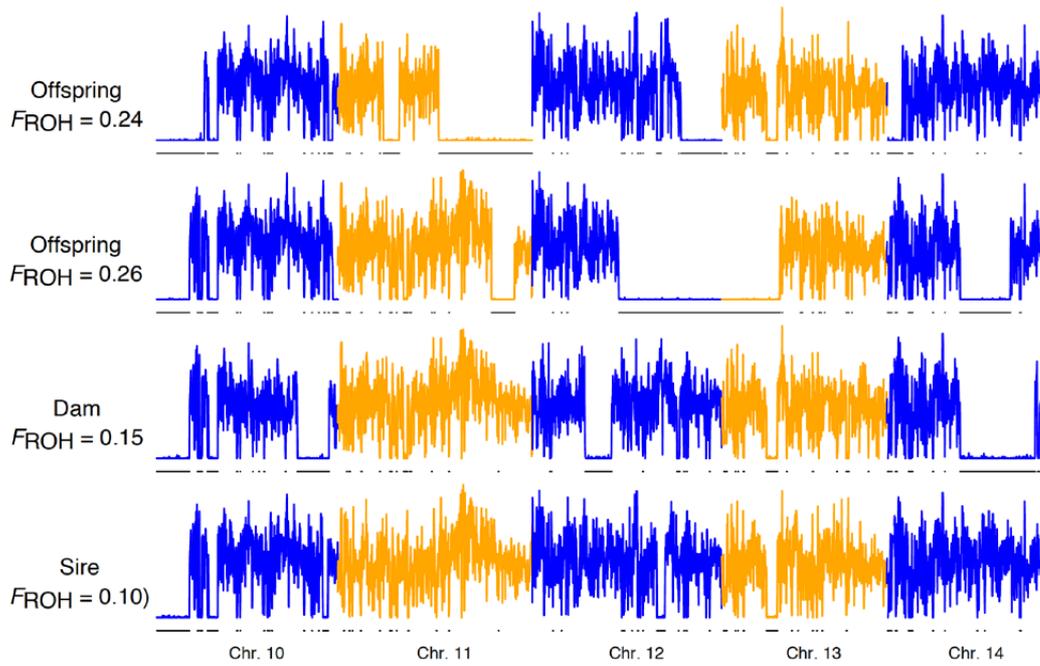739 *Genetics* **122**, 325-331 (2005).
740
741

742
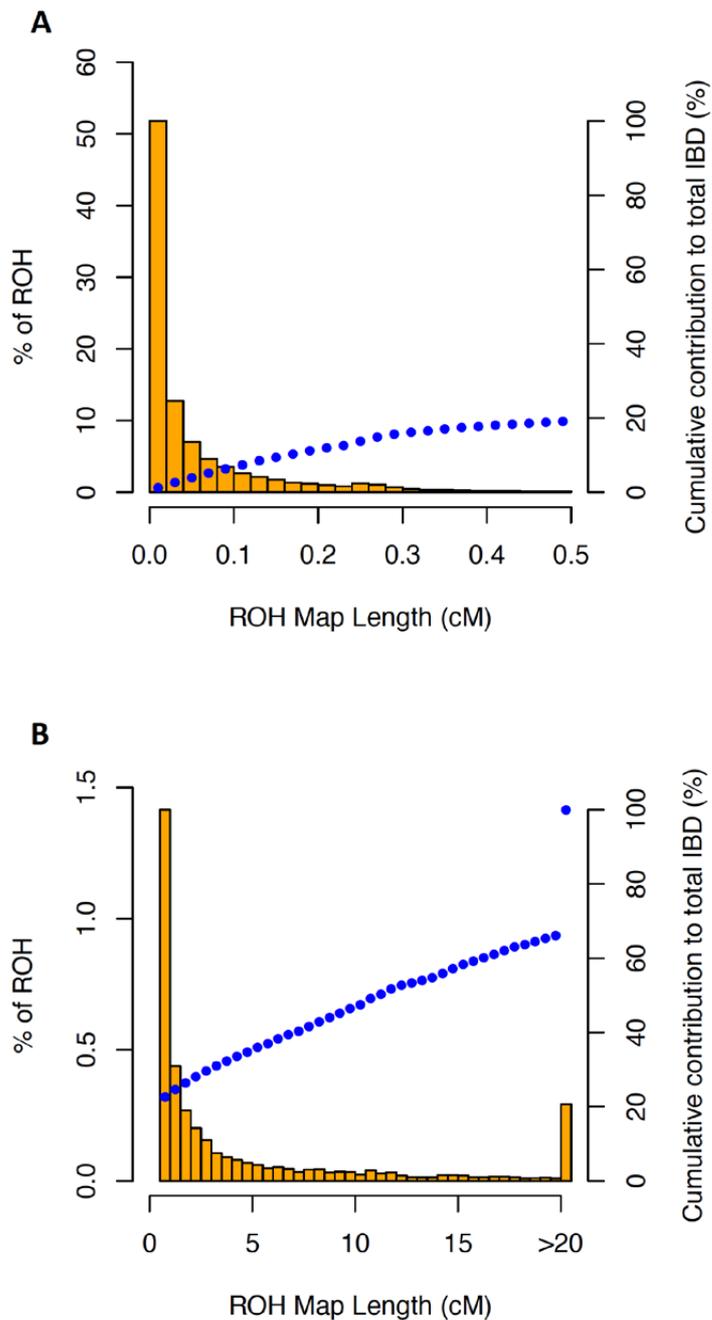
743
744

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

750

751

752

753
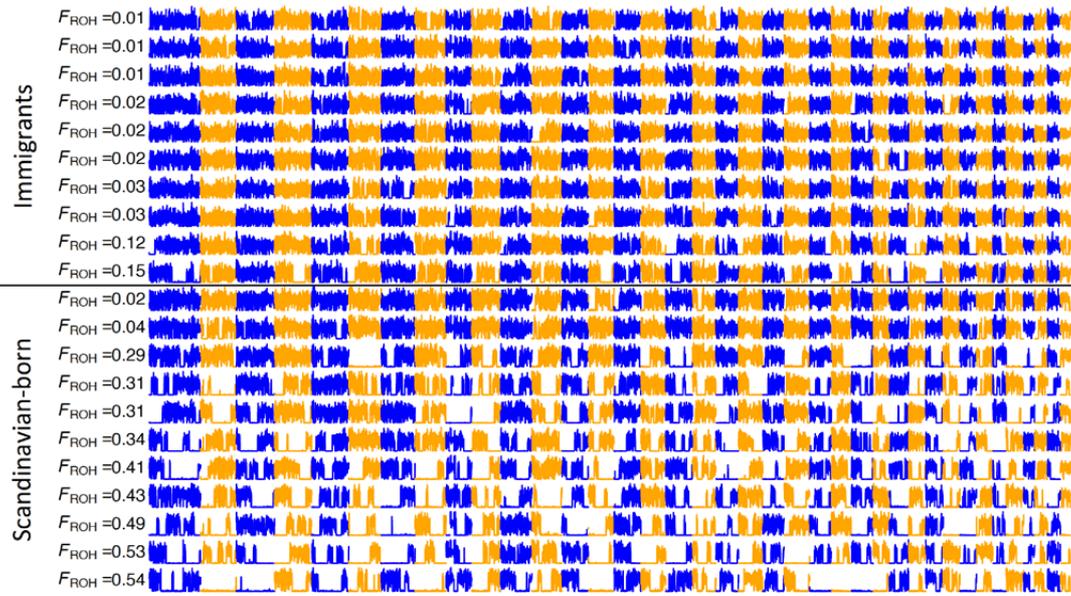
754 **Figure 2.** Heterozygosity across the 38 autosomes of 21 Scandinavian wolves.

755 Heterozygosity was measured in non-overlapping 100 kb windows as the

756 proportion of SNPs within each window that were heterozygous in the

757 individual. The y-axis ranges from zero to one for each individual. The top ten

758 individuals are immigrants into the population, followed by 11 Scandinavian-

759 born individuals. The bottom individual is the most highly inbred wolf in the

760 study. Chromosomes 1-38 are arranged left to right, with alternating blue and

761 orange representing different chromosomes. Detailed plots of heterozygosity

762 and identified ROH are provided for each chromosome in each of the 97

763 individuals in Supplementary File 1.

764

Offspring
$F_{ROH} = 0.24$

Offspring
$F_{ROH} = 0.26$

Dam
$F_{ROH} = 0.15$

Sire
$F_{ROH} = 0.10)$

Chr. 10   Chr. 11   Chr. 12   Chr. 13   Chr. 14

Figure 3. Heterozygosity across chromosomes 10-14 in two full siblings (top two panels) and their parents (bottom two panels). The black lines show long ROH (>100 kb) identified with the likelihood ratio-based sliding window approach.
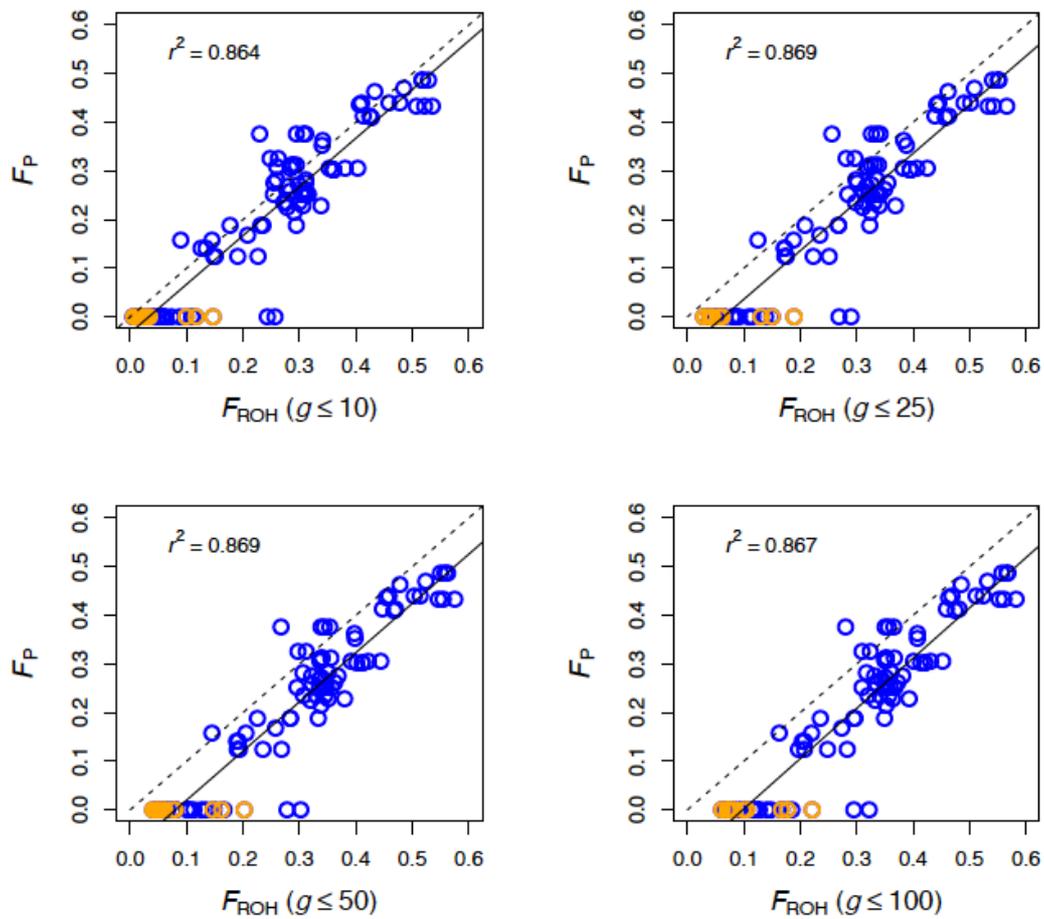
Figure 4. The relationship between F_P and F_{ROH} measured with the whole genome in Scandinavian wolves. F_P is shown on the y-axis, and F_{ROH} measured using ROH arising from relatively recent ancestors ($g \leq 10\text{-}100$ generations) on the x-axis. Immigrants are shown as orange points, and Scandinavian-born individuals are shown as blue points. 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}.
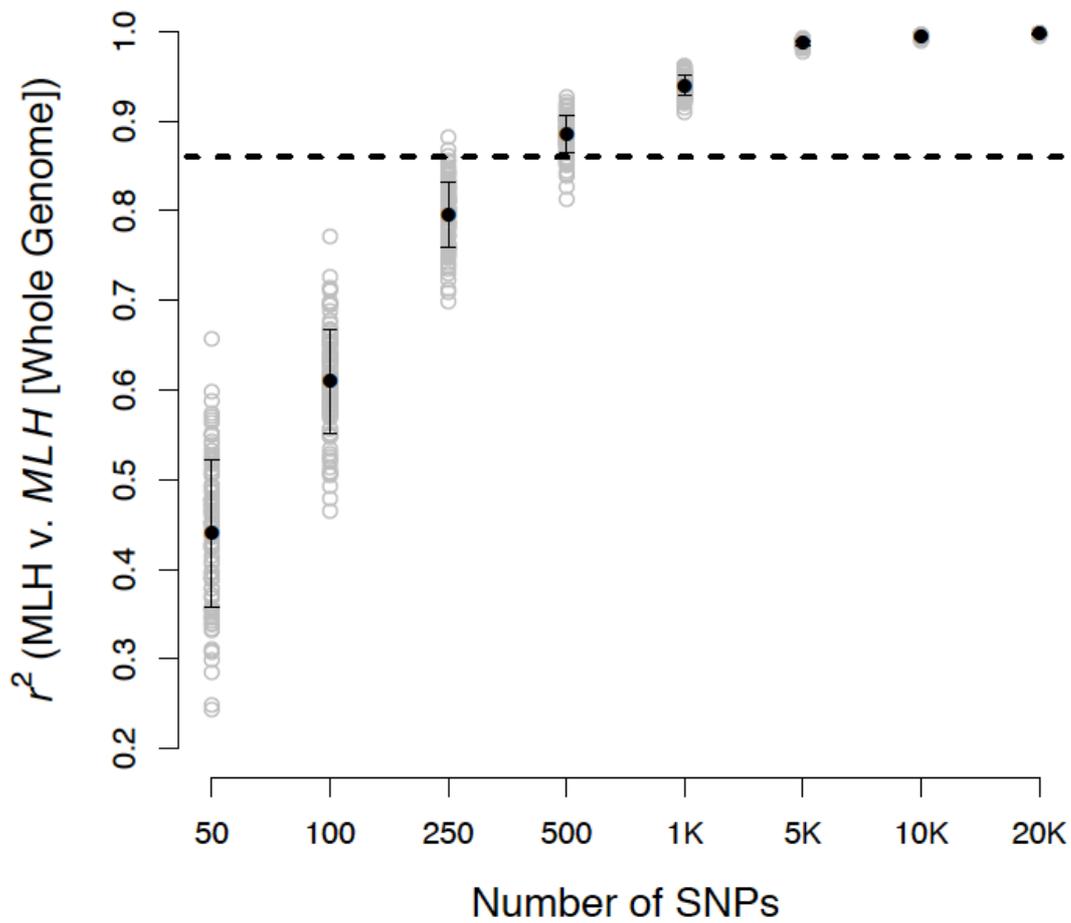
Figure 5. r^2 from regressions of *MLH* estimated with subsampled SNPs versus *MLH* calculated from the entire genome. The number of subsampled loci used to estimate *MLH* is shown on the x-axis. The black points represent the mean r^2 from analyses of 100 replicate subsamples of SNPs (+/- 1 sd). The dashed line represents the r^2 between the pedigree inbreeding coefficient and *F*. Note that the x-axis is not scaled.
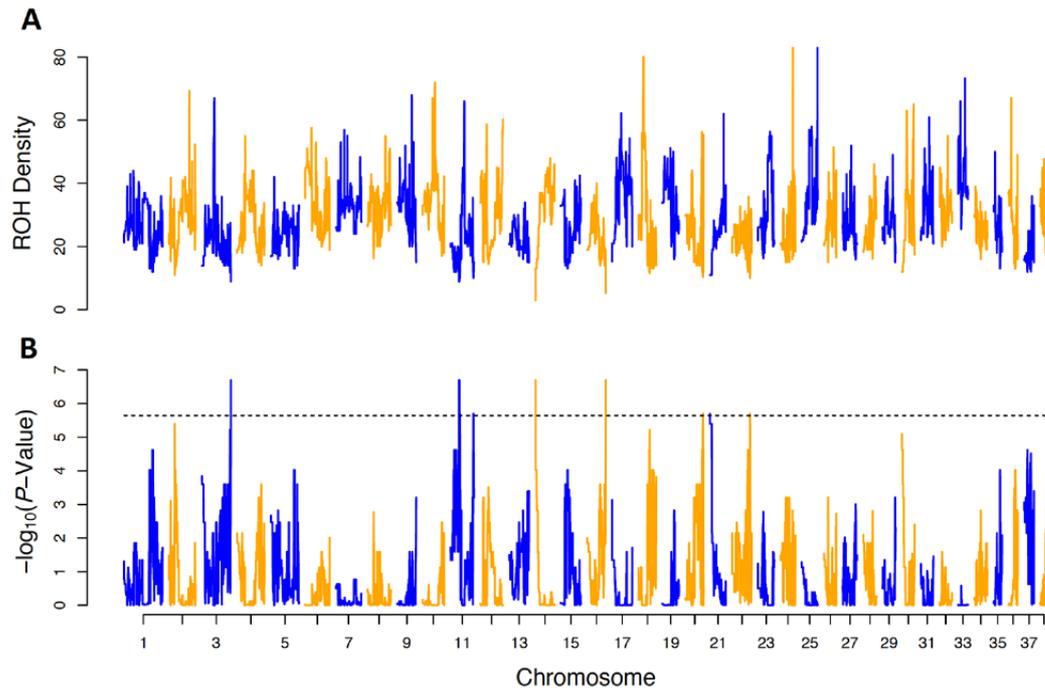
Figure 6. The density of ROH versus genomic position (**A**), and the −log$_{10}$ of *P*-values from permutation tests for deficit of ROH abundance in non-overlapping 100 kb windows (**B**). Chromosomes are arranged 1 to 38 from left to right. ROH density was measured as the summed kb in ROH across all 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

819

820 *founders

821

822

823

824 **Table S2.** ROH regions with statistically significant deficit of ROH abundance.
825 Shown are the chromosome with start and stop position (in Mb), segment
826 lengths (Mb), mean recombination rate (cM/Mb), nucleotide diversity (π), the
827 number of SNPs present in the segment, and ROH density. Note that adjacent 100
828 kb windows with statistically significant permutation tests were joined in the
829 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

830
831

832

833

834

835

836

837

838
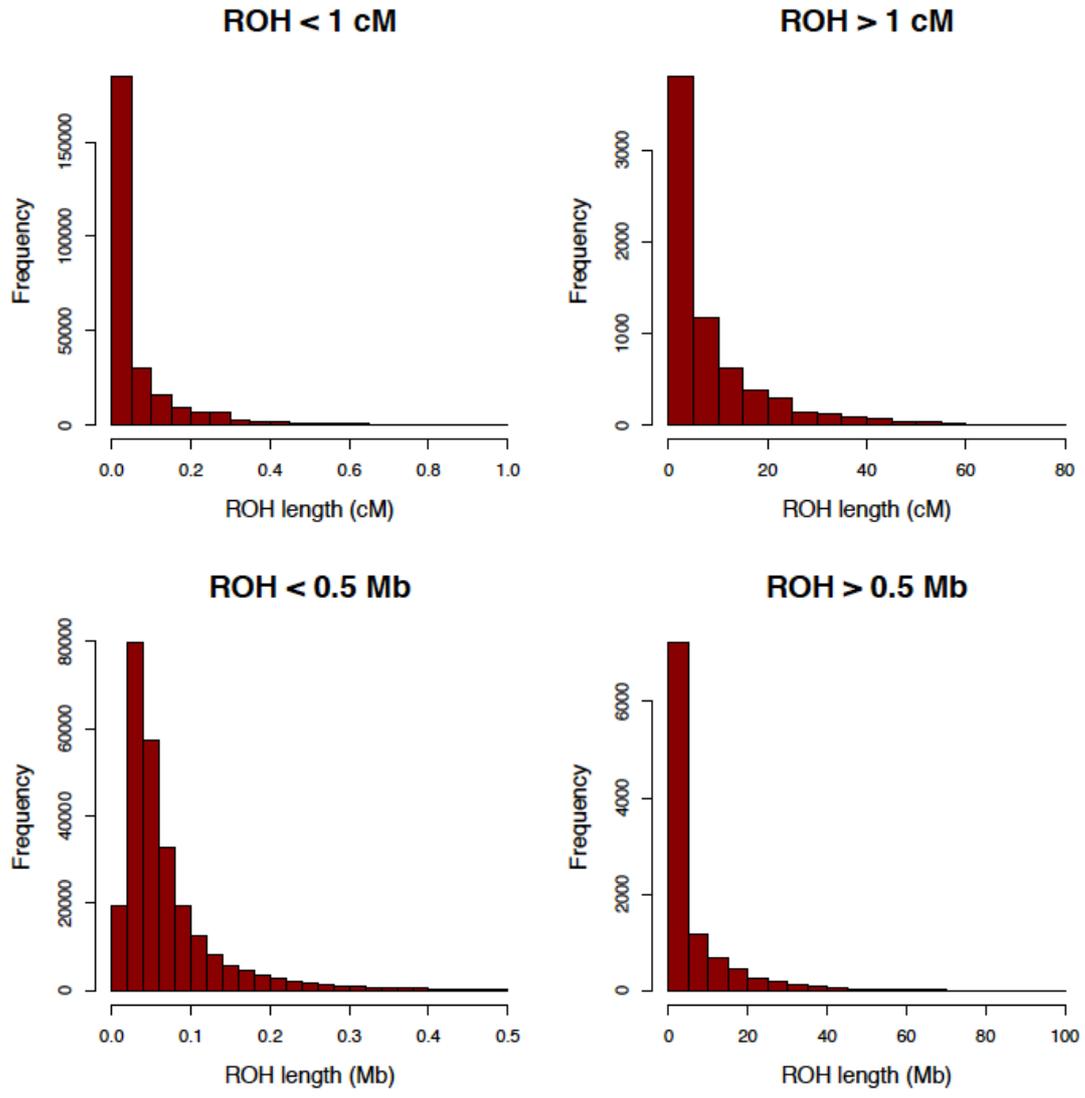
Figure S1. Distribution of the lengths of ROH. The genetic map lengths (in cM) of ROH are shown in the two upper panels. The lower two panels show the physical lengths (in Mb) of ROH.

$r^2 = 0.99$

F_{ROH} (physical mapping)

F_{ROH} (genetic mapping)
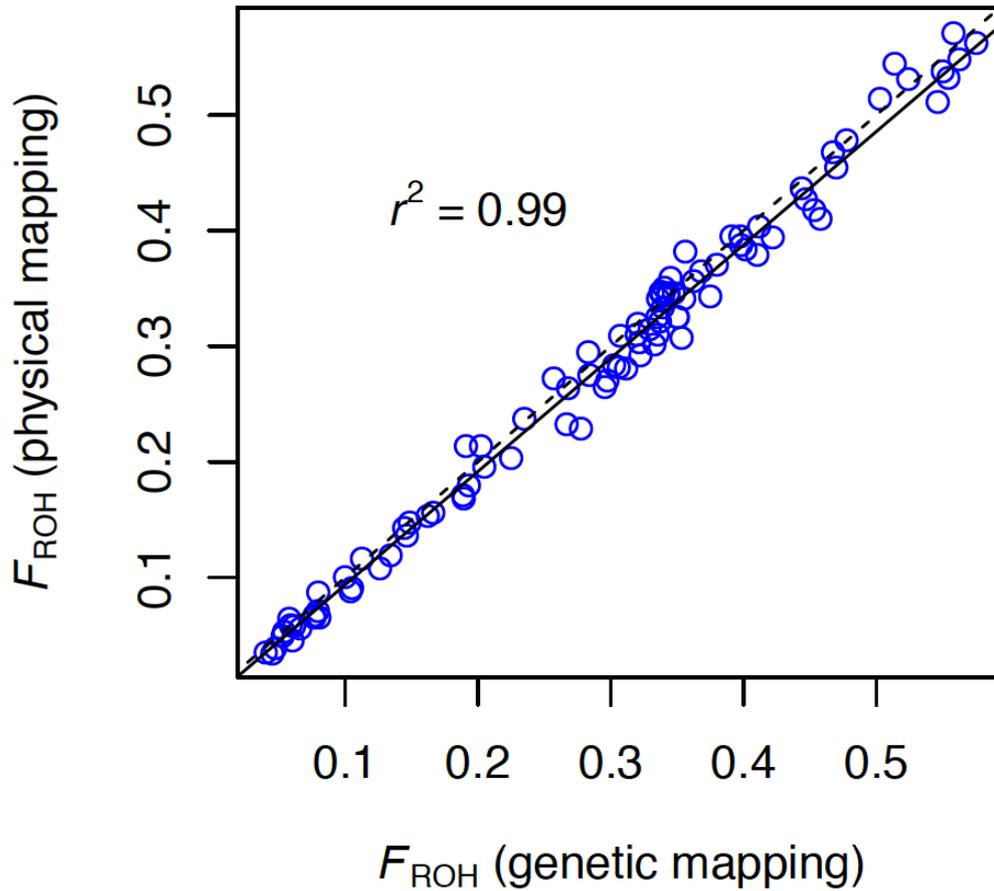
845
846 **Figure S2.** Scatterplot of FROH measured using physical mapping of ROH versus
847 FROH measured using genetic mapping of ROH. The ROH shown are those with
848 values of g ≤ 50. The results were essentially identical using other thresholds of
849 ROH length, and when using all ROH (data not 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
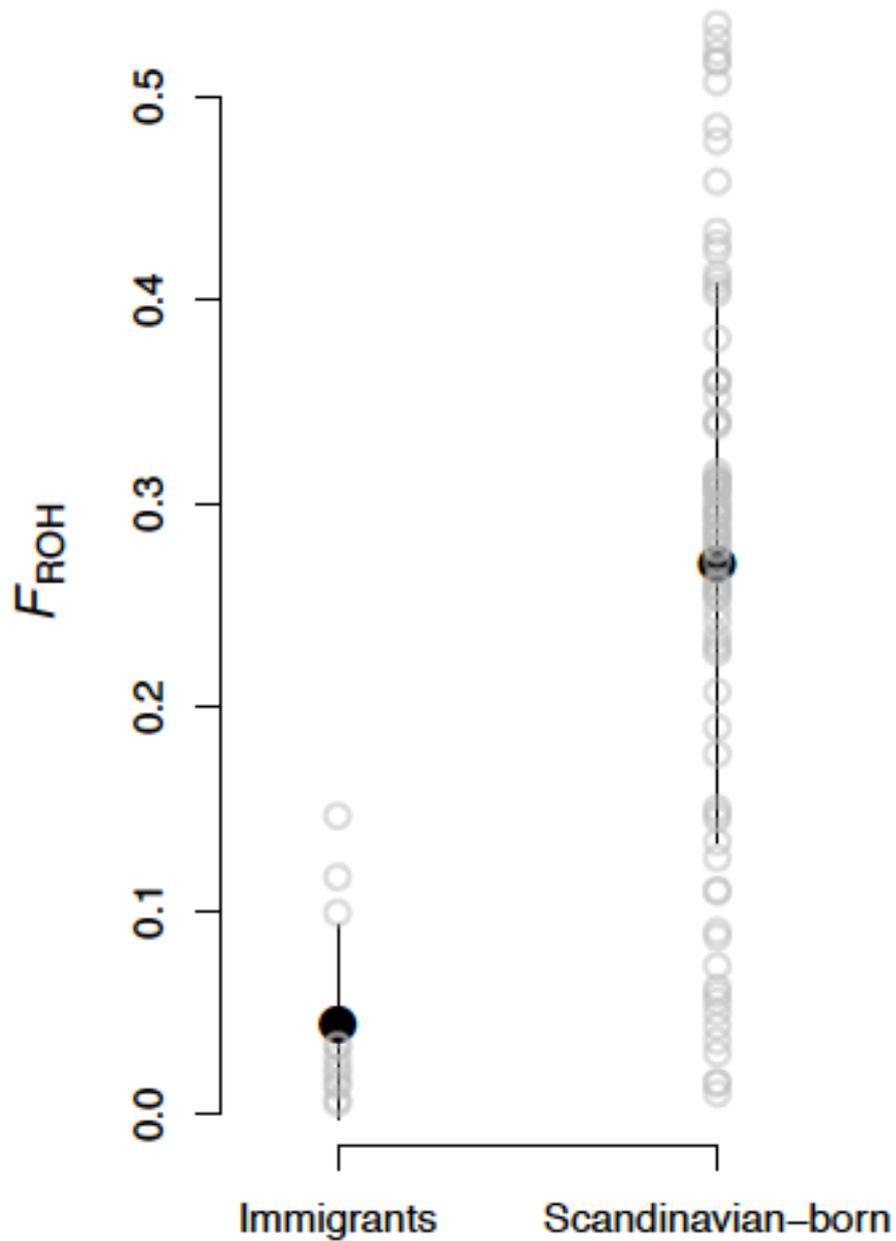
855 **Figure S3.** The distribution of F_{ROH} for immigrants and Scandinavian-born

856 wolves calculated using only ROH with $g \leq 10$ generations. The filled point and

857 error bars represent the mean +/- one standard deviation.

858
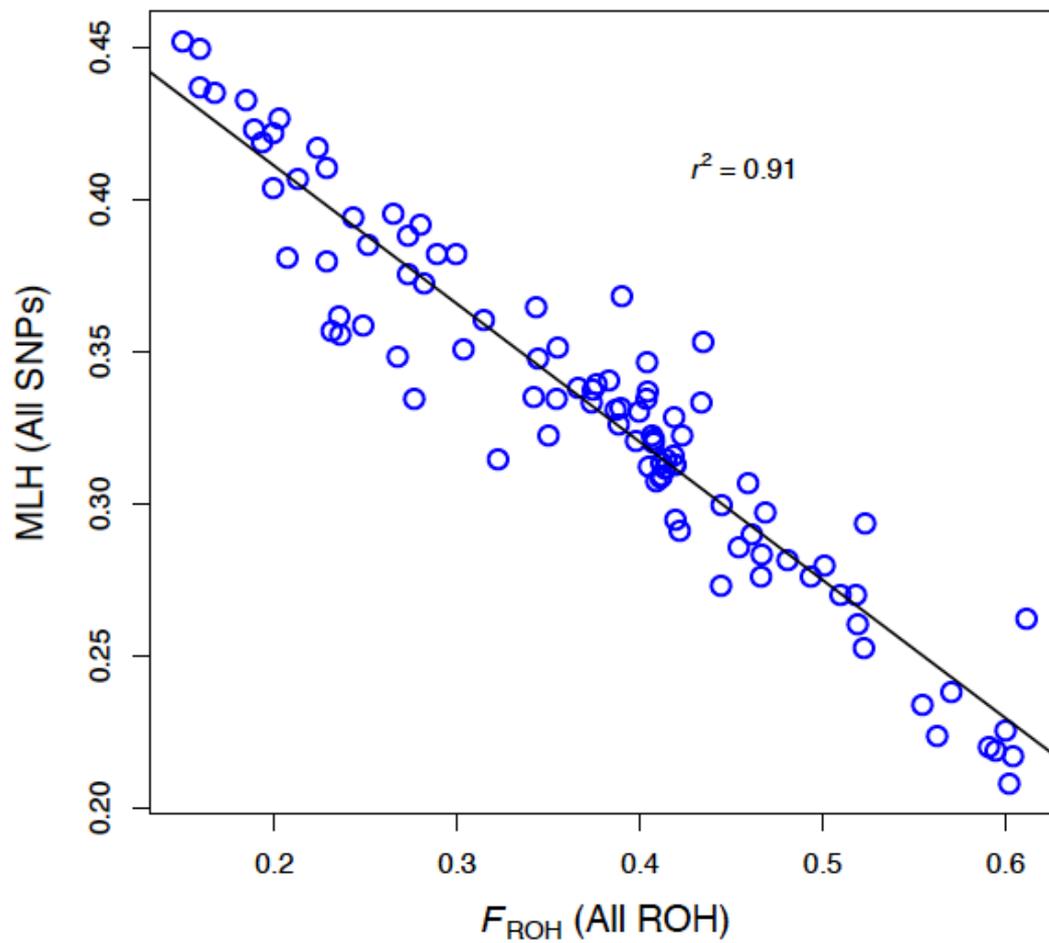
Figure S4. The relationship between MLH and F_{ROH} measured with the whole genome in Scandinavian wolves.

859
860
861
862
863
864
865
866
867
868
869
870
871
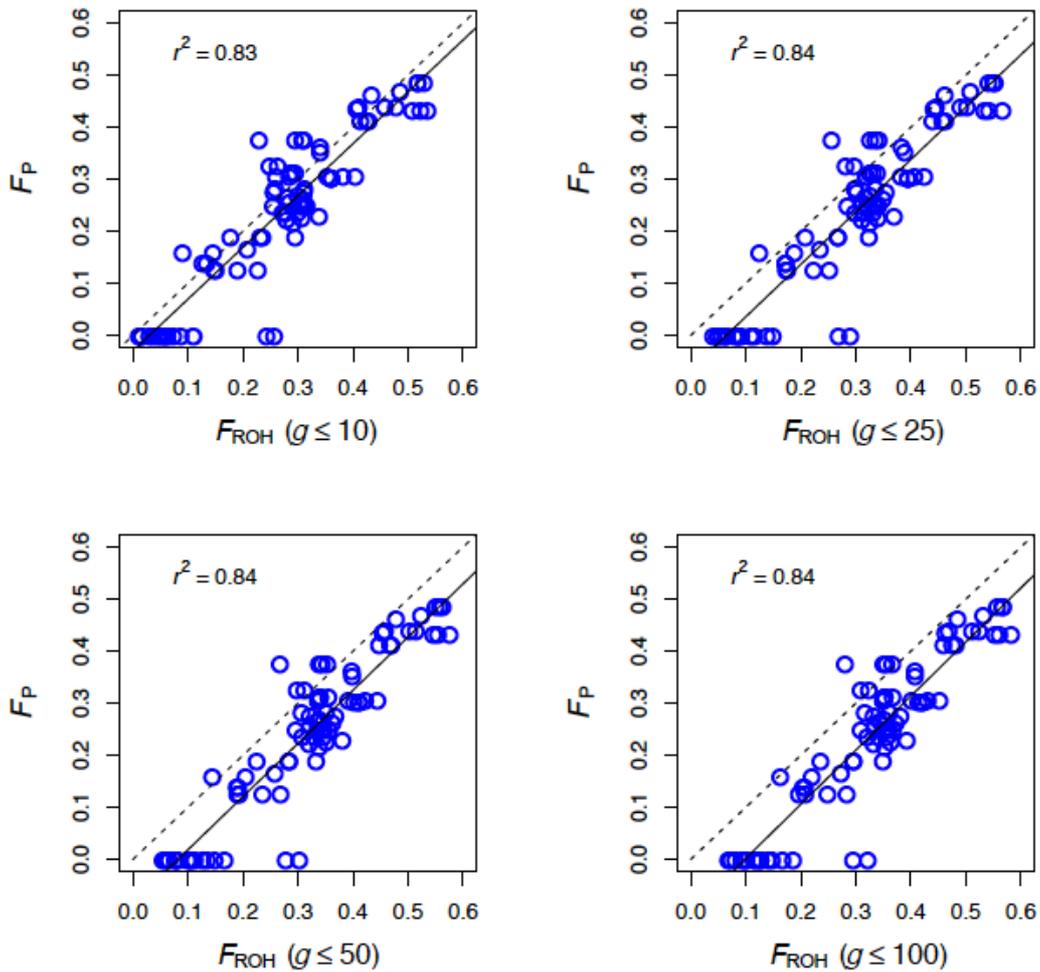
874
875
876
877 **Figure S5.** The relationship between F_P and F_{ROH} measured with the whole
878 genome in Scandinavian wolves, after excluding immigrants from the data. F_P is
879 shown on the y-axis, and F_{ROH} measured using ROH arising from relatively recent
880 ancestors ($g \leq 10\text{-}100$ generations). The dashed line has an intercept of zero and
881 a slope of one. Points below the line thus represent cases where F_P
882 underestimated F_{ROH}. The solid line is the fitted line from a regression of F_P
883 versus F_{ROH}.
884
885

886

887

888
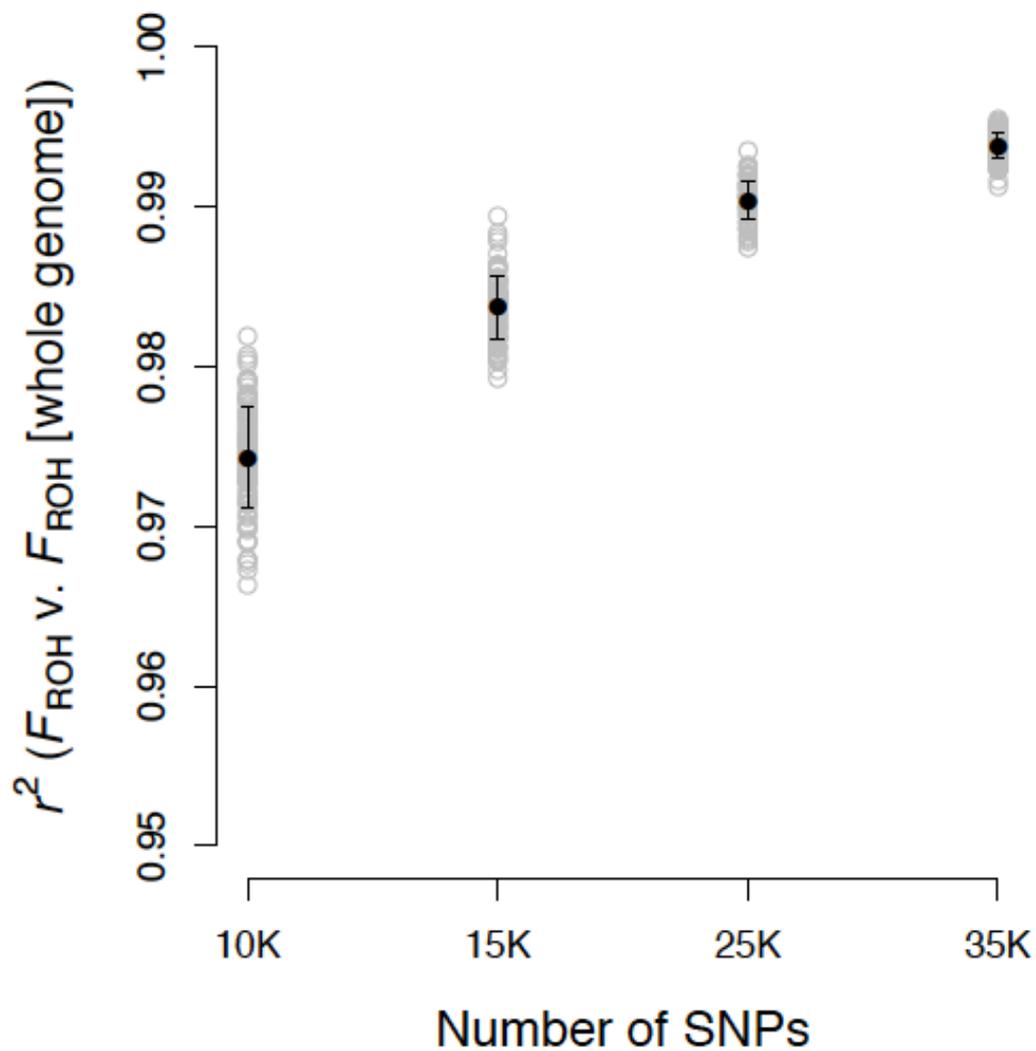
889 **Figure S6.** The precision of F_{ROH} as a function of the number of SNPs ($x10^3$)

890 analyzed. The r^2 from regressions of F_{ROH} (estimated with subsamples of SNPs)

891 versus F_{ROH} calculated from the whole genome (using ROH with $g \leq 10$) is shown

892 on the y-axis. The number of subsampled SNPs is shown on the x-axis. The gray

893 points represent the r^2 values from each of the 100 analyses for each number of

894 SNPs. The black points and error bars represent the mean r^2 +/- one standard

895 deviation of r^2 for each number of SNPs analyzed.

896

10K SNPs

15K SNPs

25K SNPs

35K SNPs
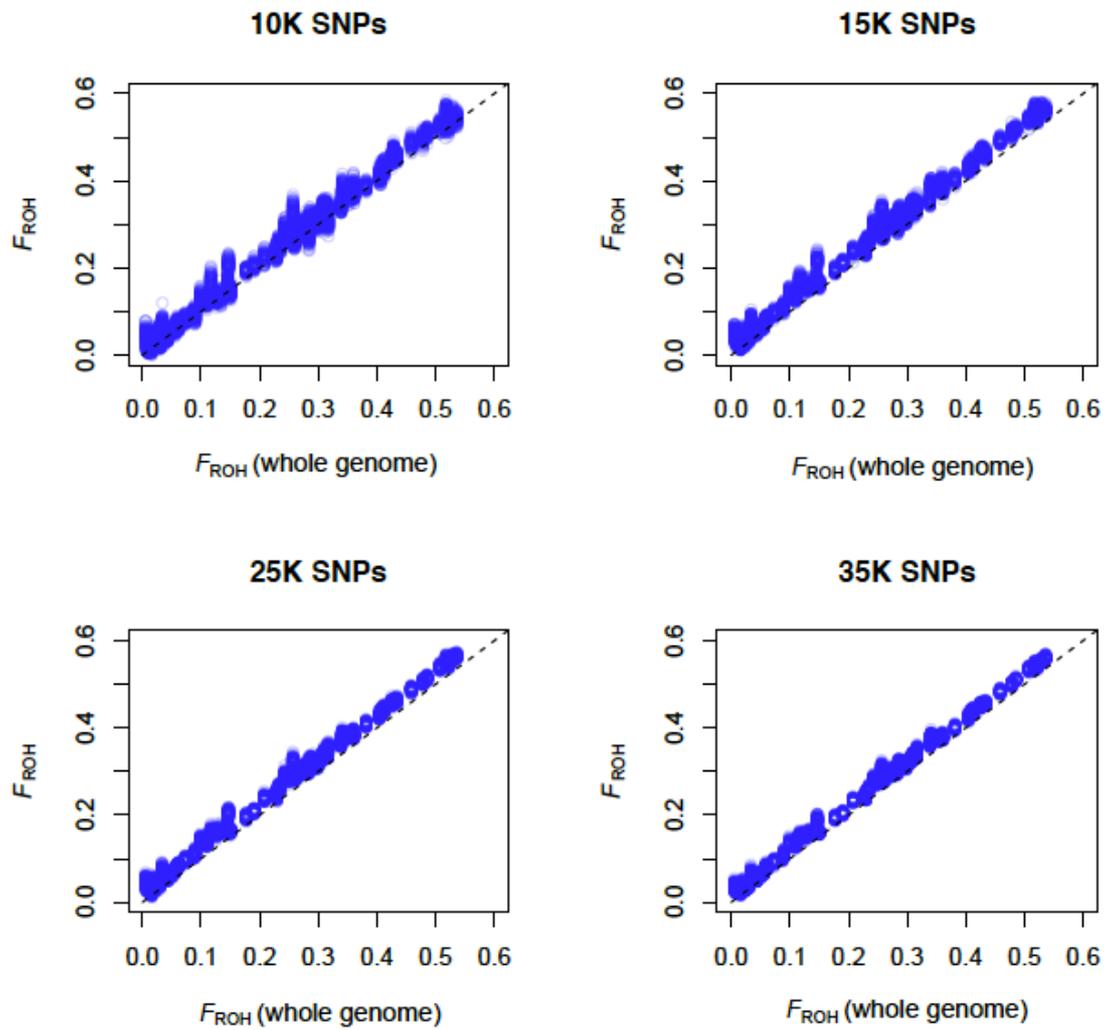
897

898

899

900 **Figure S7.** F_{ROH} measured with 10,000-35,000 subsampled SNPs plotted against

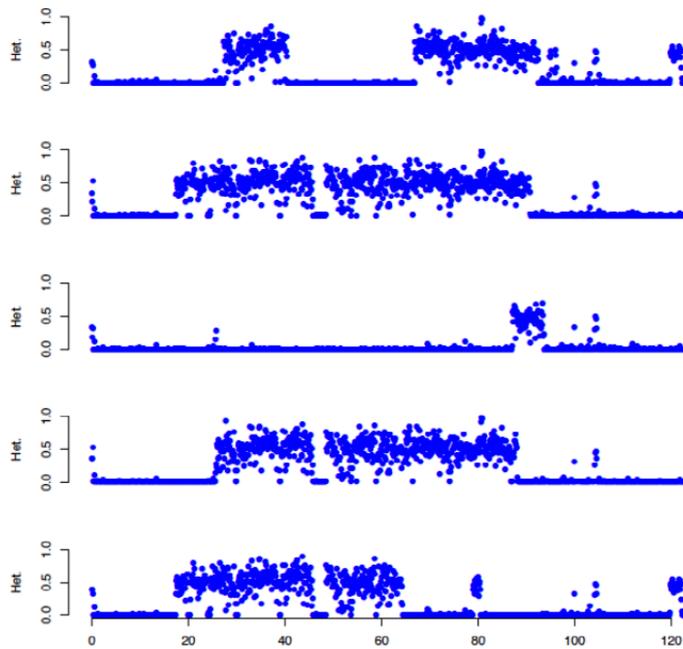901 F_{ROH} measured with the whole genome in Scandinavian wolves.
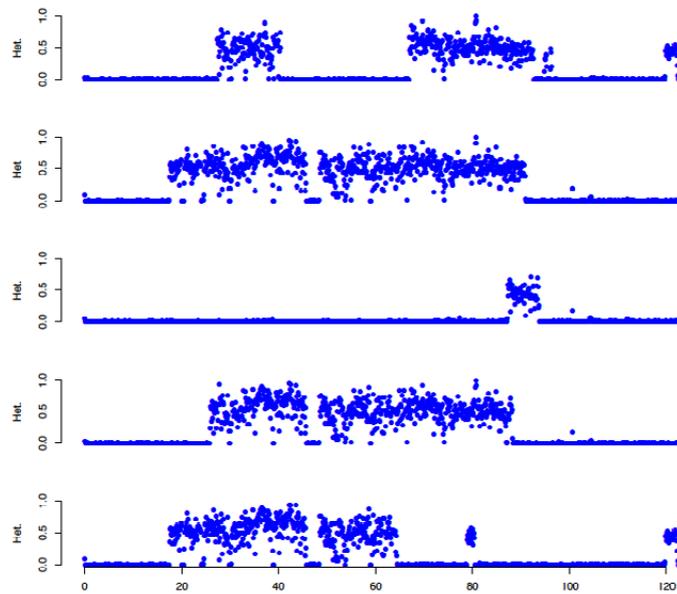
902

903
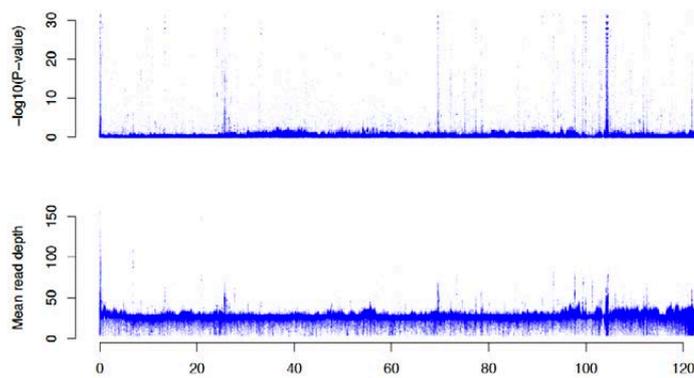
904

905

906

907 **Figure S8.** *P*-values from tests for excess of heterozygotes relative to Hardy-

908 Weinberg proportions (top panel), and mean sequence read depth (lower panel),

909 across all SNPs on chromosome 1.

910

911

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**).

917

918
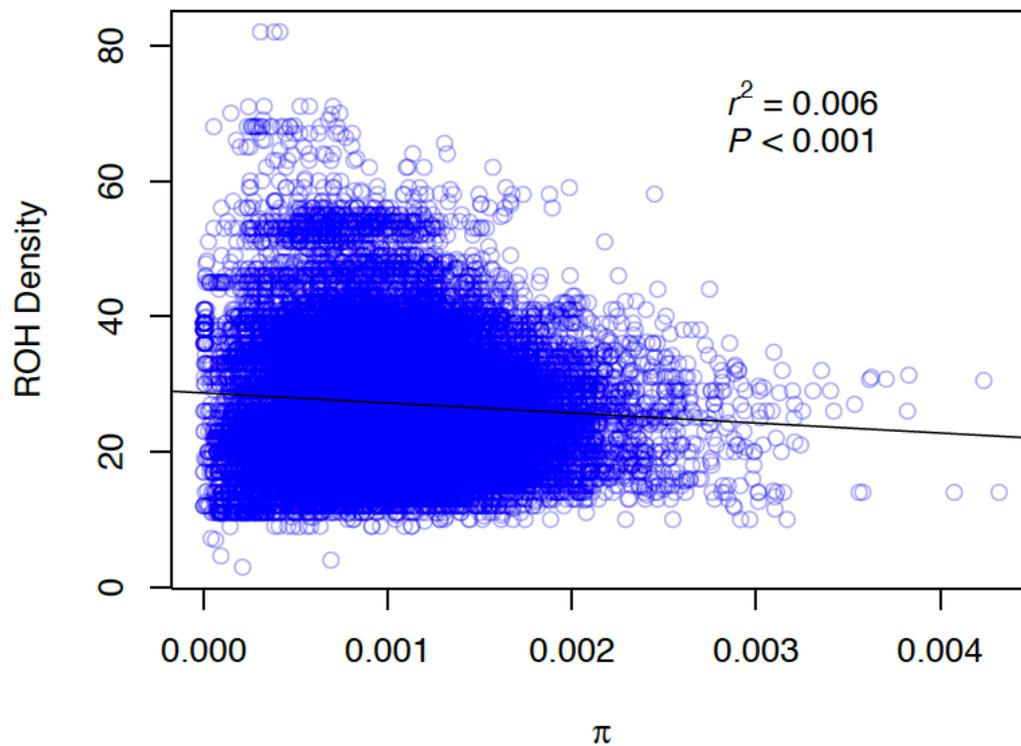
919
920 **Figure S10.** ROH density versus π measured in 100 kb windows of the 97 wolf
921 genomes. The *P*-value, r^2, and solid black fitted line are from a linear regression
922 model.
923
924
925

926
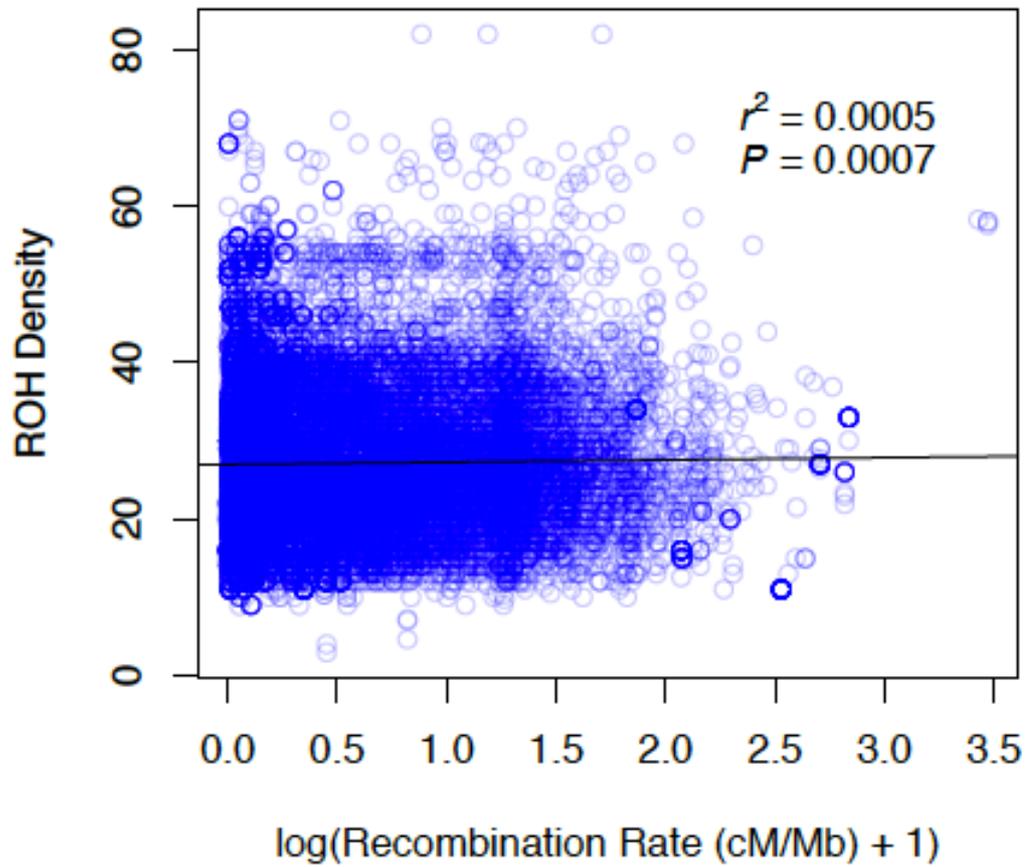
927
928
929 **Figure S11.** Scatterplot of ROH density versus the log-transformed
930 recombination rate (cM/Mb) in 100 kb windows of the 97 wolf genomes. The *P*-
931 value, r^2, and solid black fitted line are from a linear regression model.
932
933
934
935
936

937