

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23

# ESTIMATION OF INBREEDING DEPRESSION ON FEMALE FERTILITY IN THE FINNISH AYRSHIRE POPULATION

K. Martikainen<sup>1</sup>, A-M. Tyrisevä<sup>2</sup>, K. Matilainen<sup>2</sup>, J. Pösö<sup>3</sup>, P. Uimari<sup>1</sup>

<sup>1</sup>Department of Agricultural Sciences, University of Helsinki, Helsinki, Finland

<sup>2</sup>Natural Resources Institute Finland (Luke), Jokioinen, Finland

<sup>3</sup>Finnish Animal Breeding Association, Vantaa, Finland

## Correspondence

K. Martikainen, Department of Agricultural Sciences, P.O. Box 27, FI-00014 University of Helsinki,  
Finland

E-mail: [katja.martikainen@helsinki.fi](mailto:katja.martikainen@helsinki.fi)

Telephone: +358504489237

**Keywords:** cattle, fertility, inbreeding, inbreeding depression, runs of homozygosity, SNP

Running head: Inbreeding depression on fertility

## 24 Summary

25 SNP data enable the estimation of inbreeding at the genome level. In this study, we estimated  
26 inbreeding levels for 19 075 Finnish Ayrshire cows genotyped with a low-density SNP panel (8K).  
27 The genotypes were imputed to 50K density, and after quality control, 39 144 SNPs remained for  
28 the analysis. Inbreeding coefficients were estimated for each animal based on the percentage of  
29 homozygous SNPs ( $F_{PH}$ ), runs of homozygosity ( $F_{ROH}$ ), and pedigree ( $F_{PED}$ ). Phenotypic records  
30 were available for 13 712 animals including non-return rate (**NRR**), number of inseminations (**AIS**),  
31 and interval from first to last insemination (**IFL**) for heifers and up to three parities for cows, as well  
32 as interval from calving to first insemination (**ICF**) for cows. Average  $F_{PED}$  was 0.02,  $F_{ROH}$  0.06, and  
33  $F_{PH}$  0.63. A correlation of 0.71 was found between  $F_{PED}$  and  $F_{ROH}$ , 0.66 between  $F_{PED}$  and  $F_{PH}$ , and  
34 0.94 between  $F_{ROH}$  and  $F_{PH}$ . Pedigree-based inbreeding coefficients did not show inbreeding  
35 depression in any of the traits. However, when  $F_{ROH}$  or  $F_{PH}$  was used as a covariate, significant  
36 inbreeding depression was observed; a 10% increase in  $F_{ROH}$  was associated with 5 days longer  
37 IFL0 and IFL1, 2 weeks longer IFL3, and 3 days longer ICF2 compared to non-inbred cows.

## 39 Introduction

40 Mating of animals with common ancestors creates inbreeding. The inbreeding level, or inbreeding  
41 coefficient ( $F$ ), of an animal refers to the probability that two alleles at a locus are identical by  
42 descent (IBD; Falconer & Mackay, 1996). Inbreeding depression, in turn, is defined as the  
43 impairment of fertility or any other phenotypic value caused by inbreeding within a population  
44 (Falconer & Mackay, 1996). Multiple studies have reported reduced fertility due to inbreeding. For  
45 example, McParland et al. (2007) found that the calving interval of cows increased by 0.7 days  
46 and their survival to second lactation decreased by 0.3% for each 1% increase in the inbreeding  
47 coefficient. Bjelland et al. (2013) reported an increase in days open from 1.06 to 1.76 days per 1%  
48 increase in the inbreeding coefficient. Pryce et al. (2014), observed that a 1% increase in the  
49 inbreeding coefficient lengthened the calving interval by 0.18 days. Moreover, single lethal  
50 recessive alleles can cause embryo or fetus abortions at any stage of gestation, thus increasing

51 the time between parturitions of the dam. Impaired fertility also reduces profitability, because the  
52 lifetime milk production of the cow decreases and the costs related to inseminations and veterinary  
53 treatments increase. In the worst case, the cow must be involuntarily culled due to poor fertility,  
54 which incurs further costs through replacements.

55 The traditional way to estimate inbreeding coefficients is to use pedigree information  
56 ( $F_{PED}$ ). However, shallow or incomplete pedigree data may lead to an underestimation of  
57 inbreeding coefficients. An alternative method is to use single nucleotide polymorphism (SNP)  
58 marker data. The simplest estimate of genomic inbreeding is the percentage of homozygous  
59 alleles ( $F_{PH}$ ), but  $F_{PH}$  cannot distinguish between alleles that are identical by state (IBS) and those  
60 that are IBD. One way to overcome the problem is to look for continuous stretches of homozygous  
61 genotypes called runs of homozygosity (ROHs). Using ROHs increases the probability that the  
62 homozygosity is due to IBD, not IBS (Gibson, 2006). ROH length depends on the distance in  
63 generations to a common ancestor (Bjelland et al. 2013): a short ROH indicates that the common  
64 ancestor occurred several generations ago, whereas a long ROH reflects a more recent common  
65 ancestor (Purfield et al., 2012).

66 The objective of this study was to estimate inbreeding coefficients for Finnish Ayrshire  
67 cows from pedigree and genomic data, and to use this information to determine inbreeding  
68 depression of cow fertility traits in this cattle breed.

## 70 **Materials and methods**

71 The genotypes, pedigree data, breed proportions, raw phenotypes, solutions for fixed effects, and  
72 estimates of (co)variance components were obtained from NAV, Nordic Cattle Genetic Evaluation  
73 (Aarhus, Denmark) and from Faba, The Finnish Animal Breeding Association (Vantaa, Finland).

## 75 **Animals**

76 The present Red Dairy Cattle (RDC) population in Finland consists of the original Finnish Ayrshire  
77 (FAY) breed and Scandinavian and North American red breeds. Table 1 shows the number of

78 cows with genotypes and those with both genotypes and phenotypes, representing different  
79 proportions of the FAY breed. We used two sets of RDC cows in this study: one including all cows  
80 with both genotypes and phenotypes registered as Finnish RDC (RDCFIN; 13 712 cows), and the  
81 other including only those RDC cows with both genotypes and phenotypes and at least 50% of  
82 FAY based on the pedigree (FAY50; 7 547 cows). All cows were born between 2002 and 2014.

### 84 **Genomic data**

85 Genotyping was performed using the Illumina BovineLD v.2 BeadChip low-density panel (Illumina  
86 Inc., 2015), which contains 7 931 SNPs. To achieve 50K density, the genotypes were imputed by  
87 the Fimpute software (Sargolzaei et al., 2014) using the default values. The imputed genotypes  
88 were further pruned so that SNPs with minor allele frequency (MAF) of less than 0.05 or a P-value  
89 of the Chi-square test for Hardy-Weinberg equilibrium of less than 0.0001 were removed from the  
90 data. A total of 39 144 SNPs remained for the analysis.

### 92 **Phenotypic data**

93 We utilized Nordic fertility evaluation data for the RDC breed to select a sub-sample of Finnish  
94 RDC cows for this study. Phenotypes of female fertility were available for 1 805 454 animals.  
95 When combined with the available genomic data, the sub-sample comprised a total of 13 712  
96 animals with both genotypes and phenotypes. Fertility traits included non-return rate at 56 days  
97 after first insemination (NRR), number of inseminations (AIS), and intervals (in days) from calving  
98 to first insemination (ICF) and from first to last insemination (IFL). Fertility traits were considered  
99 separately for heifers (lactation 0) and for cows with one to three lactations. Descriptive statistics  
100 of unadjusted observations for each trait at each parity for both sets (RDCFIN and FAY50) of data  
101 are given in Table 2. The negative and non-integer values for observations in Table 2 are due to  
102 pre-corrections of the records for heterogeneous variance due to country, year of first calving, and  
103 parity (Fogh et al., 2003).

105 Since the data set used in this study was a sub-sample of the full Nordic fertility evaluation model  
106 for RDC, the raw phenotypic values were adjusted for systematic effects prior to the estimation of  
107 inbreeding depression using the solutions of the full evaluation model. The adjusted systematic  
108 effects included herd-birth year (for heifers) or herd-year of first calving (for cows), insemination  
109 year-month (for all traits except ICF), calving year-month (ICF), and heifer's age at first  
110 insemination. Figure 1 shows the variation of the adjusted phenotypic values.

## 112 **Estimation of inbreeding coefficients**

113 Pedigree-based inbreeding coefficients were estimated from the pedigree, with an average depth  
114 of 10 generations for the genotyped animals. Only animals with the pedigree completeness value  
115 (MacCluer et al. 1983) of 0.80 or greater based on five generations were included in the analysis.  
116 Genomic inbreeding coefficients were estimated based on either homozygous SNPs ( $F_{PH}$ ) or runs  
117 of homozygosity ( $F_{ROH}$ ). The first measure,  $F_{PH}$ , was determined for each animal as the proportion  
118 of homozygous genotypes of all genotypes. The other measures of genomic inbreeding were  
119 ROH-based. ROHs were detected with three different parameter settings using PLINK v1.07  
120 (Purcell et al., 2007). The first parameter setting (ROH\_1) was based on those used by Purfield et  
121 al. (2012). A minimum density of 1 SNP per 120 kb was set to prevent low SNP density from  
122 affecting ROH length. Short ROHs were eliminated by setting the minimum ROH length to 500 kb,  
123 but without limiting the number of SNPs per ROH (the corresponding PLINK parameters are --  
124 homozyg-density 120 --homozyg-kb 500 --homozyg-snp 0). For the second and third settings we  
125 reduced the size of the sliding window to 20 SNPs and the minimum ROH length to 10kb, and  
126 increased the minimum density to 1 SNP per 1000 kb, as in the article by Zhang et al. (2015). The  
127 difference between the second and third settings was the minimum number of SNPs per ROH: 30  
128 SNPs for ROH\_2 and 100 SNPs for ROH\_3 (the corresponding PLINK parameters are --homozyg-  
129 window-snp 20 --homozyg-density 1000 --homozyg-kb 10 --homozyg-snp 30 or 100). For all three  
130 settings we allowed one possible heterozygous genotype per window to account for potential  
131 errors in genotyping and imputation. Based on these three parameter settings ( $F_{ROH_1}$ ,  $F_{ROH_2}$  and

132  $F_{ROH_3}$ ), the inbreeding coefficient estimates were determined as the sum of SNPs in the ROHs  
133 divided by the total number of SNPs.

134

### 135 **Estimation of inbreeding depression**

136 Inbreeding depression was estimated by regressing the phenotypic values on the inbreeding  
137 coefficients. Inbreeding depression was estimated separately for each trait (NRR, AIS, ICF, and  
138 IFL) using the multi-lactation model, *i.e.* with heifer and cow traits jointly:

139

$$140 \mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e},$$

141

142 where vector  $\mathbf{y}$  contains the pre-adjusted phenotypes of a trait for each parity,  $\mathbf{b}$  is a vector of fixed  
143 effects including the mean  $\mu$  and the linear regression coefficient  $b$  for each parity,  $\mathbf{a}$  is a vector of  
144 random additive genetic effects, and  $\mathbf{e}$  is a vector of random residual effects. The matrix  $\mathbf{X}$   
145 includes the inbreeding coefficients, either  $F_{PED}$ ,  $F_{ROH_1}$  or  $F_{PH}$ , for each animal, and  $\mathbf{Z}$  is an  
146 incidence matrix that relates the appropriate effects to each observation. Furthermore, it was  
147 assumed that the random genetic effects were normally distributed with  $N(\mathbf{0}, \mathbf{A} \otimes \mathbf{G})$ , where  $\mathbf{A}$  is the  
148 pedigree-based additive relationship matrix and  $\mathbf{G}$  is the additive genetic variance-covariance  
149 matrix of the heifer and cow traits (*e.g.* between IFL0, IFL1, IFL2, and IFL3), and also that the  
150 random residual effects were normally distributed with  $N(\mathbf{0}, \mathbf{R})$ . Variances and covariances were  
151 the same as in the Nordic fertility evaluation (Muuttoranta et al., 2016). Genetic groups for animals  
152 with unknown parents were treated in the analysis as random effects. The statistical analyses  
153 were performed with the DMU program package (Madsen and Jensen, 2000).

154

## 155 **Results**

### 156 **Runs of homozygosity**

157 We used three settings in PLINK to detect ROHs in the population of 19 075 genotyped animals.

158 The first setting (ROH\_1) gave a total of 411 541 ROHs for all animals. The frequency distribution

of ROHs shorter than 50 Mb using the ROH\_1 setting is shown in Figure 2. There were 783 ROHs longer than 50 Mb, with a maximum length of 125.1 Mb. The number of SNPs in the ROHs varied from 6 to 1 897. With the ROH\_2 and ROH\_3 settings, the total numbers of ROHs found for all animals were 838 383 and 165 843, respectively. ROH lengths varied from 0.9 Mb to 138 Mb for ROH\_2, and from 3.3 Mb to 138 Mb for ROH\_3. The number of SNPs in the ROHs varied from 30 to 2 123 for ROH\_2 and from 100 to 2 123 for ROH\_3.

## **Inbreeding coefficients**

The average inbreeding coefficients of all RDCFIN animals were 0.02, 0.09, and 0.63 for  $F_{PED}$ ,  $F_{ROH}$ , and  $F_{PH}$ , respectively (Table 3). The corresponding averages for FAY50 animals were 0.03, 0.10, and 0.63 (Table 3). Homozygosity-based estimates ( $F_{PH}$ ) were on a different scale than the other estimates.  $F_{PH}$  could have been adjusted for the expected amount of homozygosity (Purcell et al., 2007) or calculated from the genomic relationship matrix (VanRaden et al., 2011) to be on a similar scale as the other inbreeding coefficient estimates. However, using scaled values instead of raw values would not have affected the obtained results of inbreeding depression. ROH-based inbreeding coefficients showed more variation than the pedigree- and homozygosity-based estimates (Tables 3 and 4, and Figures 3 and 4). Among the tested ROH methods, the lowest average estimates of inbreeding coefficients were obtained for ROH\_3 (0.04 for RDCFIN and 0.05 for FAY50) and the highest for ROH\_2 (0.09 for RDCFIN and 0.1 for FAY50). The average of  $F_{ROH_1}$  was 0.06 for RDCFIN and 0.07 for FAY50 (Table 3).

Correlations between the pedigree-based and genomic measures of inbreeding coefficients for RDCFIN were moderate: 0.66 between  $F_{PED}$  and  $F_{PH}$  and 0.71 between  $F_{PED}$  and  $F_{ROH_1}$  (Table 5). The corresponding correlations for FAY50 varied from 0.55 between  $F_{PED}$  and  $F_{PH}$  to 0.59 between  $F_{PED}$  and  $F_{ROH_1}$  (Table 5). Very strong correlations were detected between all genomic measures using all RDCFIN animals: from 0.90 between  $F_{ROH_3}$  and  $F_{PH}$  to 0.98 between  $F_{ROH_1}$  and  $F_{ROH_2}$ . The use of only FAY50 animals resulted in almost equal correlations, from 0.89 ( $F_{ROH_3}$  and  $F_{PH}$ ) to 0.98 ( $F_{ROH_1}$  and  $F_{ROH_2}$ ). Due to the very strong correlation between all ROH-

186 based estimates, only  $F_{ROH_1}$ , which had the strongest correlation with  $F_{PED}$ , was selected for  
187 subsequent analysis.

## 189 **Inbreeding depression**

190 There were large differences in the estimates of inbreeding depression between traits, parities,  
191 data sets, and measures of inbreeding (Tables 6–8). In general, no statistically significant results  
192 were obtained when  $F_{PED}$  was used as a covariate in the model. The only exception was NRR1 in  
193 the FAY50 data set ( $P < 0.1$ ), with a deteriorating effect of approximately 0.9% per 1% increase of  
194  $F_{PED}$  (Table 6).  $F_{PH}$  had a statistically significant effect on NRR1 in both data sets, deteriorating  
195 NRR1 by approximately 1.1% and 1.2% per 1% increase in  $F_{PH}$  in the RDCFIN and FAY50 data  
196 sets, respectively (Table 6).

197 A high genetic correlation (0.91) has been reported between IFL, which measures the  
198 service period in days, and AIS, which measures the number of inseminations in the same period  
199 (Berry et al., 2014). The results for both traits were congruent in our study, and thus only the  
200 results of IFL are presented here (Table 7). In the RDCFIN animals, an increase of 1% in  $F_{ROH_1}$   
201 lengthened IFL0 and IFL1 by approximately 0.4 and 0.5 days, respectively. Similarly, a 1%  
202 increase in  $F_{PH}$  was associated with an increase of 0.9 days in IFL0 and of 1.1 days in IFL1. Using  
203 the FAY50 data set, the corresponding estimates were 0.4 days (IFL0 and  $F_{ROH_1}$ ), 0.6 days (IFL1  
204 and  $F_{ROH_1}$ ), and 1.4 days (IFL1 and  $F_{PH}$ ). Moreover, an increase of 1% in  $F_{ROH_1}$  and  $F_{PH}$   
205 increased IFL3 by 1.5 days and 3 days, respectively (Table 7).

206 Among the three ICF traits, only ICF2 showed statistically significant inbreeding  
207 depression with  $F_{ROH_1}$  as a covariate; a 1% increase in  $F_{ROH_1}$  increased ICF2 by 0.3 and 0.4 days  
208 in the RDCFIN or FAY50 data sets, respectively.

## 210 **Discussion**

### 211 **Runs of homozygosity and inbreeding coefficient**



212 ROH determination depends on the selected parameters. Consequently, we compared three  
213 different parameter settings which differed in the minimum requirements for the number of SNPs in  
214 ROH, ROH length, and SNP density in ROH. The obtained results indicate that the number of  
215 ROH segments increases with a decrease in the number of SNPs required to determine ROH.  
216 Eventually, the ROH-based estimates of inbreeding coefficients converge to  $F_{PH}$  when ROH length  
217 diminishes to a single SNP. However, the possibility to detect ROH segments that are IBS but not  
218 IBD also increases along with a smaller number of required SNPs. The most stringent setting in  
219 our study was ROH\_3, for which the minimum number of SNPs in ROH was set to 100. The  
220 largest number of ROH segments was found using ROH\_2, which differed from ROH\_3 only in  
221 terms of the minimum number of SNPs in ROH (30 instead of 100 SNPs). As expected, the  
222 estimates of inbreeding coefficients depended on these settings. ROH\_3 gave the smallest  
223 average inbreeding coefficient, followed by ROH\_1 and ROH\_2.

224           The size of the sliding window may have had an effect on ROH lengths as well, since  
225 a SNP is included in a ROH only if 5% of the windows containing the SNP are completely  
226 homozygous (Howrigan et al., 2011). Additionally, Ferenčaković et al. (2013) showed that ROHs  
227 may depend on the density of the genotyping panel. In their study, a 50K panel gave a larger  
228 number of small (<4 Mb) ROH segments than a high-density panel. No differences between  
229 panels were obtained for ROHs longer than 4 Mb. The authors concluded that a 50K panel creates  
230 false positive findings of short ROHs, and therefore leads to an overestimation of  $F_{ROH}$ . In the  
231 present study, almost half (198 595 of 411 541) of ROHs (determined using the ROH\_1 setting)  
232 were shorter than 5 Mb (Figure 2). This indicates that ROH\_1, with average estimates of 0.06 for  
233 RDCFIN and 0.07 for FAY50, may overestimate the level of inbreeding. The ROH\_2 analysis  
234 revealed twice as many ROHs as ROH\_1, and presumably resulted in even higher overestimation  
235 of inbreeding (0.09 for RDCFIN and 0.10 for FAY50). In contrast, ROH\_3 only detected ROHs  
236 longer than 3.3 Mb, which may have led to an underestimation of  $F_{ROH}$ , with average values of 0.04  
237 for RDCFIN and 0.05 for FAY50. With the ideal criteria for detecting ROHs, the average  
238 inbreeding coefficient estimate would probably have been somewhere between the  $F_{ROH_3}$  and

239  $F_{ROH\_1}$  values. We also tested if pruning of SNPs based on linkage disequilibrium (LD) has an  
240 effect on detection of ROHs and estimates of inbreeding depression. For this we repeated the  
241 ROH\_1 analysis using the LD-pruned RDCFIN dataset (PLINK: --indep-pairwise 50 5 0.5, resulting  
242 in 29 390 SNPs). As a result, the correlation between  $F_{ROH\_1}$  values from the LD-pruned and  
243 unpruned data was 0.98. Also the pruning had a very minor effect on the estimates of inbreeding  
244 depression e.g. for IFL0 the LD-pruned data gave 43.4 (SE=13.8) compared the estimate of 43.2  
245 (SE=13.0) from the unpruned data. Despite the effect of panel density and the settings of ROH  
246 calling, many studies have concluded that  $F_{ROH}$  provides the most effective and consistent  
247 measure of the inbreeding coefficient compared to other methods (e.g. Keller et al., 2011; Bjelland  
248 et al., 2013).

249 All inbreeding coefficients ( $F_{PED}$ ,  $F_{PH}$ ,  $F_{ROH}$ ) calculated by the three methods were  
250 correlated, but the correlations were higher between genomic estimates ( $r = 0.89-0.98$ ) than  
251 between pedigree and genomic estimates ( $r = 0.55-0.71$ ). Similar high correlations have been  
252 reported in other studies as well. Bjelland et al. (2013) observed a correlation of 0.81 between  
253 genomic inbreeding measures ( $F_{PH}$  and  $F_{ROH}$ ), and Pryce et al. (2014) reported a corresponding  
254 correlation of 0.9. Keller et al. (2011) found that inbreeding coefficients calculated using ROH  
255 correlated strongly (0.6) with the homozygous mutation load, whereas the correlation between the  
256 homozygous mutation load and pedigree-based inbreeding coefficients was weak (0.25). Pryce et  
257 al. (2014) detected a correlation of 0.53 between  $F_{PED}$  and  $F_{ROH}$  and of 0.45 between  $F_{PED}$  and  
258  $F_{PH}$ , while Purfield et al. (2012) reported an even stronger positive correlation (0.73) between  $F_{PED}$   
259 and  $F_{ROH}$ . The correlations found in the present study between the genomic measures of  
260 inbreeding coefficients were consistent with previous studies, and correlations between estimates  
261 obtained by pedigree and genomic methods were almost as strong as those reported by Purfield  
262 et al. (2012).

## 263

## 264 **Inbreeding depression**

265 We observed virtually no inbreeding depression associated with  $F_{PED}$  in the present  
266 study. However, the genomic measures of inbreeding coefficients revealed inbreeding depression  
267 on NRR1 ( $F_{PH}$ ), ICF2 ( $F_{ROH_1}$ ), IFL0, IFL1, and IFL3 (both  $F_{PH}$  and  $F_{ROH_1}$ ). The phenotypic data  
268 used in the present study were from field records. Therefore, the number of highly inbred cows  
269 was small and the analyses were based primarily on cows with low levels of inbreeding. As  
270 Cassell et al. (2003) noted, in such conditions, non-significant results are common. In general,  
271 power of regression coefficient depends on data size and dispersion of the dependent variable. In  
272 our case, the number of animals was bigger for heifers than for cows. This may explain the  
273 differences between the parities. Smaller dispersion of  $F_{PED}$  than  $F_{ROH_1}$  may have resulted in  
274 reduced statistical power of pedigree based inbreeding depression estimation compared to  
275 genomic measures of inbreeding depression. Also, culling for poor fertility can create bias in  
276 estimates of inbreeding depression if the cause of poor fertility is inbreeding. This may be an  
277 additional explanation for differences between parities. Moreover, selection can create beneficial  
278 homozygosity at certain loci thus having an opposite effect on a trait compared to inbreeding  
279 based homozygosity. However, even though fertility traits are part of the current Nordic breeding  
280 goal, the effect of selection on our results is expected to be minimal.

281 Also Pryce et al. (2014) found differences between pedigree- and genome-based  
282 estimates of inbreeding depression, implying that the use of pedigree information probably  
283 underestimates inbreeding depression on female fertility. On the contrary, Ferenčaković et al.  
284 (2013) reported that the widely used Illumina BovineSNP50 BeadChip (Illumina Inc. 2016)  
285 overestimates the number of short ROHs and, consequently, the inbreeding coefficient. In  
286 addition, Pryce et al. (2014) reported that only ROHs longer than 60 SNPs or 3.5 Mb were  
287 associated with a decrease in milk yield independent of the overall level of homozygosity. They  
288 suggested that if inbreeding is due to an ancient common ancestor, selection has had an  
289 opportunity to purge deleterious mutations and therefore they did not find association between  
290 short ROHs and decrease in milk yield.

291 Among the studied fertility traits in the present study, NRR had a bimodal distribution  
292 even after adjustment for heterogeneous variance and systematic effects. The analysis with  
293 unadjusted data and logistic regression model might have resulted in more reliable estimates of  
294 inbreeding depression on NRR. However, the structure of the data (number of observations in the  
295 different classes of the systematic effects) did not allow this approach to be applied.

296 Thanks to its simplicity, the calving interval is among the most widely used fertility  
297 traits. Many studies have reported that inbreeding causes a lengthening of the calving interval  
298 (e.g. Smith et al., 1998; Wall et al., 2005; McParland et al., 2007; Pryce et al., 2014). The calving  
299 interval is defined as the interval from previous calving to conception (days open) and gestation  
300 length. Pereira et al. (2016) showed that increased inbreeding only affects days open, but not  
301 gestation length. Days open also comprises two periods: the interval from calving to first  
302 insemination and from first to last insemination, both of which were examined in the present study  
303 and revealed inbreeding depression. Bjelland et al. (2013) reported an increase of 1.76 and 1.72  
304 days in days open per 1% increase in  $F_{PH}$  and  $F_{ROH}$ , respectively. The corresponding results in the  
305 present study were approximately 0.3 days for ICF and approximately 0.5 days for IFL. Combining  
306 the estimates of inbreeding depression for the two intervals gave estimates from 0.59 (first parity  
307 and RDCFIN) to 0.94 (second parity and FAY50) with  $F_{ROH_1}$  and from 1.17 (first parity and  
308 RDCFIN) to 2.16 (second parity and FAY50) with  $F_{PH}$ . These results are in accordance with those  
309 of Bjelland et al. (2013).

310 In conclusion, we showed that genome-based estimates of inbreeding differ from  
311 pedigree-based estimates, and that genomic inbreeding estimates are associated with female  
312 fertility. Further examination of the effect of the density of the SNP panel and the length of ROHs  
313 on inbreeding depression would elucidate the role of inbreeding depression on fertility traits. It is  
314 possible that the sum of ROHs or homozygosity may not reveal all the harmful effects of  
315 inbreeding on fertility. We suggest that a more detailed intra-chromosomal approach (Kleinman-  
316 Ruiz et al., 2016) could reveal specific chromosomal regions that are strongly affected by  
317 inbreeding depression in Finnish Ayrshire cows.

318

## 319 **Acknowledgements**

320 The authors thank Suomen Naudanjalostussäätiö for funding this research. We are grateful to  
321 NAV (Nordic Cattle Genetic Evaluation) and Faba (The Finnish Animal Breeding Association) for  
322 providing the genotypic data, pedigree data, breed proportions, raw phenotypes, solutions for fixed  
323 effects, and estimates of (co)variance-components.

324

## 325 **References**

326 Berry, D. P., Wall, E., & Pryce, J. E. (2014). Genetics and genomics of reproductive performance  
327 in dairy and beef cattle. *Animal*, 2014, 8 (SUPPL. 1), 105–121.

328

329 Bjelland, D. W., Weigel, K. A., Vukasinovic, N., & Nkrumah, J. D. (2013). Evaluation of inbreeding  
330 depression in Holstein cattle using whole-genome SNP markers and alternative measures of  
331 genomic inbreeding. *J. Dairy Sci.*, 96, 4697–4706.

332

333 Cassell B. G., Adamec V., Pearson R. E. (2003) Maternal and fetal inbreeding depression for 70-  
334 day nonreturn and calving rate in Holsteins and Jerseys. *J. Dairy Sci.*, 86, 2977-2983.

335

336 Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. Longman,  
337 Essex, U.K., 4th edition.

338

339 Ferencakovic, M., Sölkner, J., & Curik, I. (2013). Estimating autozygosity from high-throughput  
340 information: Effects of SNP density and genotyping errors. *Genet. Sel. Evol.*, 45 (1), 42.

341

342 Fogh, A., Roth, A., Maagaard Pedersen, O., Eriksson, J.-Å., Juga, J., Toivonen, M., Ranberg,  
343 I.M.A., Steine, T., Sander Nielsen, U., & Pedersen Aamand, G. (2003). A joint Nordic model for  
344 fertility traits. *Interbull Bull.*, 31, 52–55.

345

346 Gibson, J. (2006). Extended tracts of homozygosity in outbred human populations. *Hum. Mol.*  
347 *Genet.*, 15, 789.

348

349 Howrigan, D. P., Simonson, M. A., & Keller, M. C. (2011). Detecting autozygosity through runs of  
350 homozygosity: A comparison of three autozygosity detection algorithms. *BMC Genomics*, 12.

351

352 Illumina Inc. (2015). Illumina Data Sheet: Agrigenomics. BovineLD v2.0 Genotyping BeadChip.  
353 Accessed Oct. 19, 2016.  
354 [http://www.illumina.com/Documents/products/datasheets/datasheet\\_bovineLD.pdf](http://www.illumina.com/Documents/products/datasheets/datasheet_bovineLD.pdf).

355

356 Illumina Inc. (2016). Illumina Data Sheet: DNA Analysis. BovineSNP50 Genotyping BeadChip.  
357 Accessed Oct. 20, 2016.  
358 [http://www.illumina.com/Documents/products/datasheets/datasheet\\_bovine\\_snp50.pdf](http://www.illumina.com/Documents/products/datasheets/datasheet_bovine_snp50.pdf)

359

360 Keller, M. C. (2011). Quantification of inbreeding due to distant ancestors and its detection using  
361 dense single nucleotide polymorphism data. *Genetics*, 189, 237.

362

363 Kleinman-Ruiz, D., Villanueva, B., Fernández, J., Toro, M. A., García-Cortés, L. A., & Rodríguez-  
364 Ramilo, S. T. (2016). Intra-chromosomal estimates of inbreeding and coancestry in the Spanish  
365 Holstein cattle population. *Livest. Sci.*, 185, 34–42.

366

367 Madsen, P., & Jensen, J. (2008). *A User's Guide to DMU – A package for analysing multivariate*  
368 *mixed models*. Version 6, release 4.7. University of Aarhus, Faculty of Agricultural Science, Tjele,  
369 Denmark.

370

371 Maccluer J. W., Boyce A. J., Dyke B., Weitkamp L. R., Pfenning D. W., Parsons C. J. (1983)  
372 Inbreeding and pedigree structure in standardbred horses. *J. Hered.*, 74, 394-399.

373

374 Mc Parland, S., Kearney, J. F., Rath, M., & Berry, D. P. (2007). Inbreeding effects on milk  
375 production, calving performance, fertility, and conformation in Irish Holstein-Friesians. *J. Dairy*  
376 *Sci.*, 90, 4411–4419.

377

378 Muuttoranta, K., Tyrisevä, A.-M., Mäntysaari, E. A., Pösö, J., Aamand, G. P., Eriksson, J.-Å.,  
379 Nielsen, U.S., & Lidauer, M.H. (2016). Female fertility in Nordic dairy cattle. Manuscript in  
380 preparation.

381

382 Pereira, R. J., Santana, M. L. Jr, Ayres, D. R., Bignardi, A. B., Menezes, G. R., Silva, L. O.,  
383 Machado, C. H., Josahkian, L. A., & Albuquerque, L. G.. (2016). Inbreeding depression in Zebu  
384 cattle traits. *J. Anim. Breed. Genet.*, May 19.

385

386 Pryce, J. E., Haile-Mariam, M., Goddard, M. E., & Hayes, B. J. (2014). Identification of genomic  
387 regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genet. Sel.*  
388 *Evol.*, 46, 71.

389

390 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar,  
391 P., De Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome  
392 association and population-based linkage analyses. *Am. J. Hum. Genet.*, 81, 559–575.

393

394 Purfield, D. C., Berry, D. P., McParland, S., & Bradley, D. G. (2012). Runs of homozygosity and  
395 population history in cattle. *BMC Genet.*, 13, 70.  
396

397 Sargolzaei, M., Chesnais, J. P., & Schenkel, F. S. (2014). A new approach for efficient genotype  
398 imputation using information from relatives. *BMC Genomics*, 15, 478.  
399

400 Smith, L. A., Cassell, B. G., & Pearson, R.E. (1998). The effects of inbreeding on the lifetime  
401 performance of dairy cattle. *J. Dairy Sci.*, 81, 2729–2737.  
402

403 VanRaden, P. M., Olson, K. M., Wiggans, G. R., Cole, J. B., & Tooker, M. E. (2011) Genomic  
404 inbreeding and relationships among Holsteins, Jerseys, and Brown Swiss. *J. Dairy Sci.*, 94, 5673–  
405 5682.  
406

407 Wall, E., Brotherstone, S., Kearney, J. F., Woolliams, J. A., & Coffey, M. P. (2005). Impact of  
408 nonadditive genetic effects in the estimation of breeding values for fertility and correlated traits. *J.*  
409 *Dairy Sci.*, 88, 376–385.  
410

411 Zhang, Q., Calus, M. P. L., Guldbbrandtsen, B., Lund, M. S., & Sahana, G. (2015). Estimation of  
412 inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds.  
413 *BMC Genet.*, 16, 88.  
414  
415  
416



417 Figure legends

418

419 Figure 1. Box-plots of adjusted phenotypic values

420

421 Figure 2. Frequency distribution of ROH length using the first parameter setting. Only ROHs  
422 shorter than 50 Mb are presented.

423

424 Figure 3. Box-plots of inbreeding coefficients estimated with different methods for RDCFIN and  
425 FAY50 data sets

426

427 Figure 4. Scatter density plot of  $F_{ROH_1}$  and  $F_{PED}$ . Each point is colored by the frequency of  
428 observations.

429

430

431 **Table 1** Number of cows with different proportions of the Finnish Ayrshire (FAY) breed

|                          | Animals with genotypes | Animals with genotypes and phenotypes |
|--------------------------|------------------------|---------------------------------------|
| RDCFIN <sup>1</sup>      | 19 075                 | 13 712                                |
| At least 25% FAY         | 18 393                 | 13 199                                |
| At least 50% FAY (FAY50) | 10 199                 | 7 547                                 |
| At least 75% FAY         | 420                    | 372                                   |

432 <sup>1</sup>Red Dairy Cattle registered in Finland

433 **Table 2** Descriptive statistics of unadjusted fertility traits of genotyped RDCFIN / FAY50 cows  
434

435 (values for both data sets are presented if different)

|            |   | Number of animals | Mean        | SD          | Min         | Max   |
|------------|---|-------------------|-------------|-------------|-------------|-------|
| NRR (%)    | 0 | 13 368 / 7 358    | 62.1 / 61.5 | 47.4 / 47.6 | -1.3        | 100.5 |
|            | 1 | 9 474 / 5 230     | 54.3 / 53.3 | 49.3        | 0.2         | 99.8  |
|            | 2 | 5 043 / 2 949     | 53.0 / 52.2 | 49.2 / 49.3 | 0.3         | 99.8  |
|            | 3 | 1 540 / 1 011     | 54.5        | 49.1        | 0.4         | 99.8  |
| AIS (n)    | 0 | 13 261 / 7 301    | 1.8         | 1.1         | 1.0         | 5.2   |
|            | 1 | 9 323 / 5 155     | 2.0 / 2.1   | 1.2         | 1.0         | 5.0   |
|            | 2 | 4 918 / 2 891     | 2.1         | 1.2 / 1.3   | 1.0         | 5.0   |
|            | 3 | 1 453 / 957       | 2.0         | 1.2         | 1.0         | 5.1   |
| ICF (days) | 1 | 9 453 / 5 220     | 85.9 / 86.5 | 28.4 / 28.6 | 19.1        | 189.4 |
|            | 2 | 5 067 / 2 968     | 87.4 / 88.0 | 29.2 / 29.3 | 15.4 / 24.9 | 188.7 |
|            | 3 | 1 509 / 991       | 86.3 / 87.8 | 28.0 / 28.3 | 23.1 / 24.2 | 188.5 |
| IFL (days) | 0 | 12 878 / 7 080    | 26.9 / 27.8 | 40.8 / 41.5 | -2.3        | 242.1 |
|            | 1 | 9 546 / 5 267     | 40.1 / 41.7 | 56.2 / 56.9 | -5.1        | 251.6 |
|            | 2 | 5 131 / 2 991     | 42.7 / 44.3 | 55.9 / 56.8 | -5.2        | 251.8 |
|            | 3 | 1 567 / 1 028     | 41.0 / 41.8 | 54.1 / 55.0 | -6.5        | 258.0 |

453

454 **Table 3** Descriptive statistics of inbreeding coefficients estimated with different methods for

455 RDCFIN / FAY50 data sets

|                    | Mean        | SD          | Min           | Max         |
|--------------------|-------------|-------------|---------------|-------------|
| F <sub>PED</sub>   | 0.02 / 0.03 | 0.01 / 0.01 | 0.0 / 0.0     | 0.29 / 0.29 |
| F <sub>ROH_1</sub> | 0.06 / 0.07 | 0.03 / 0.02 | 0.001 / 0.01  | 0.28 / 0.28 |
| F <sub>ROH_2</sub> | 0.09 / 0.10 | 0.03 / 0.02 | 0.008 / 0.03  | 0.30 / 0.30 |
| F <sub>ROH_3</sub> | 0.04 / 0.05 | 0.02 / 0.02 | 0.003 / 0.003 | 0.27 / 0.27 |
| F <sub>PH</sub>    | 0.63 / 0.63 | 0.01 / 0.01 | 0.60 / 0.60   | 0.71 / 0.71 |

456

457 **Table 4** Frequency distribution of F<sub>PED</sub> (N<sub>F<sub>PED</sub></sub>) and F<sub>ROH\_1</sub> (N<sub>F<sub>ROH\_1</sub></sub>)

| Inbreeding coefficient class | N <sub>F<sub>PED</sub></sub> | N <sub>F<sub>ROH_1</sub></sub> |
|------------------------------|------------------------------|--------------------------------|
| 0.00                         | 21                           | 2                              |
| 0.00-0.01                    | 2661                         | 123                            |
| 0.01-0.02                    | 3503                         | 379                            |
| 0.02-0.03                    | 3713                         | 917                            |
| 0.03-0.04                    | 2204                         | 1475                           |
| 0.04-0.05                    | 934                          | 1893                           |
| 0.05-0.06                    | 676 <sup>1</sup>             | 2006                           |
| 0.06-0.07                    |                              | 2057                           |
| 0.07-0.08                    |                              | 1743                           |
| 0.08-0.09                    |                              | 1235                           |
| 0.09-0.10                    |                              | 784                            |
| >0.10                        |                              | 1098                           |

458 <sup>1</sup>F<sub>PED</sub> > 0.05

459

460

461 **Table 5** Correlations between different estimates of inbreeding based on RDCFIN / FAY50 data  
 462 sets

|                    | F <sub>PED</sub> | F <sub>ROH_1</sub> | F <sub>ROH_2</sub> | F <sub>ROH_3</sub> | F <sub>PH</sub> |
|--------------------|------------------|--------------------|--------------------|--------------------|-----------------|
| F <sub>PED</sub>   | 1                |                    |                    |                    |                 |
| F <sub>ROH_1</sub> | 0.71 / 0.59      | 1                  |                    |                    |                 |
| F <sub>ROH_2</sub> | 0.70 / 0.58      | 0.98 / 0.98        | 1                  |                    |                 |
| F <sub>ROH_3</sub> | 0.69 / 0.57      | 0.96 / 0.95        | 0.92 / 0.92        | 1                  |                 |
| F <sub>PH</sub>    | 0.66 / 0.55      | 0.94 / 0.93        | 0.95 / 0.95        | 0.90 / 0.89        | 1               |

463

464

465 **Table 6** Estimates of inbreeding depression (SE in brackets) for non-return rate (NRR) in parities  
 466 0–3 for RDCFIN and FAY50 data sets

|      |        | F <sub>PED</sub> | F <sub>ROH_1</sub> | F <sub>PH</sub> |
|------|--------|------------------|--------------------|-----------------|
| NRR0 | RDCFIN | 18.0 (40.0)      | -9.4 (17.1)        | -17.1 (40.7)    |
|      | FAY50  | 27.4 (39.9)      | 17.1 (22.4)        | 64.3 (53.3)     |
| NRR1 | RDCFIN | -60.9 (40.2)     | -35.0 (22.4)       | -106.5** (53.2) |
|      | FAY50  | -93.6* (52.5)    | -40.2 (29.7)       | -118.9* (70.4)  |
| NRR2 | RDCFIN | 7.8 (53.2)       | -13.1 (29.8)       | -43.7 (71.2)    |
|      | FAY50  | 83.5 (68.9)      | -14.1 (39.2)       | -45.4 (92.9)    |
| NRR3 | RDCFIN | 38.4 (87.8)      | 17.0 (51.1)        | 21.6 (126.7)    |
|      | FAY50  | -7.6 (106.7)     | -0.6 (64.3)        | 10.9 (154.8)    |

467 \*P-values < 0.1, \*\*P-value < 0.05, \*\*\*P-value < 0.01

468

469 **Table 7** Estimates of inbreeding depression (SE in brackets) for interval from first to last  
 470 insemination (IFL) in parities 0–3 for RDCFIN and FAY50 data sets

|      |        | F <sub>PED</sub> | F <sub>ROH_1</sub> | F <sub>PH</sub> |
|------|--------|------------------|--------------------|-----------------|
| IFL0 | RDCFIN | 28.6 (23.4)      | 43.2*** (13.0)     | 89.1*** (30.9)  |
|      | FAY50  | 8.7 (30.1)       | 38.1** (17.0)      | 62.2 (40.4)     |
| IFL1 | RDCFIN | 41.2 (47.8)      | 54.8** (26.6)      | 110.0* (63.1)   |
|      | FAY50  | 53.3 (62.6)      | 64.8* (35.1)       | 139.6* (83.3)   |
| IFL2 | RDCFIN | 2.7 (62.5)       | 37.8 (34.8)        | 111.3 (83.3)    |
|      | FAY50  | -32.6 (81.6)     | 54.6 (46.0)        | 148.5 (109.1)   |
| IFL3 | RDCFIN | 27.7 (104.2)     | 79.8 (60.6)        | 217.6 (149.9)   |
|      | FAY50  | 148.5 (126.1)    | 145.9* (75.8)      | 326.1* (182.5)  |

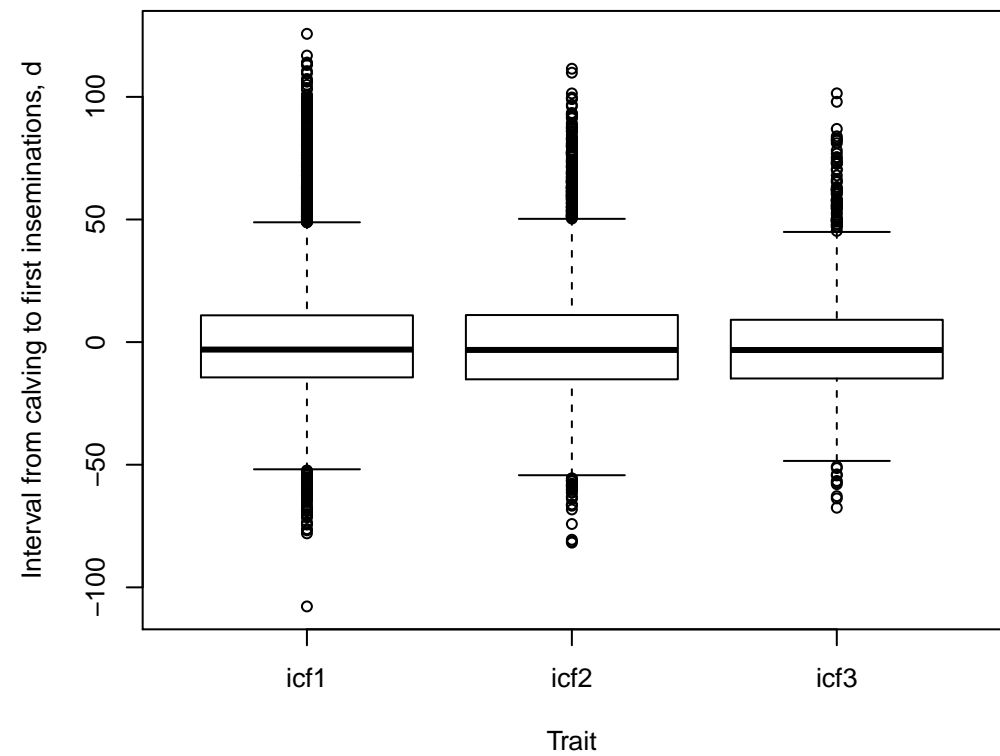
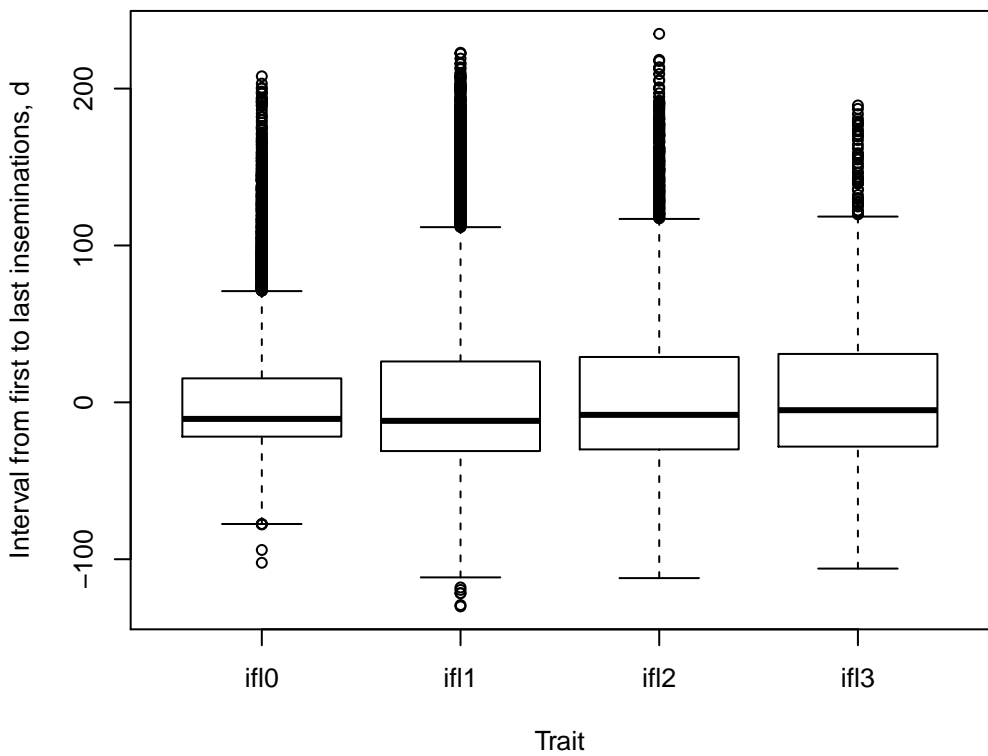
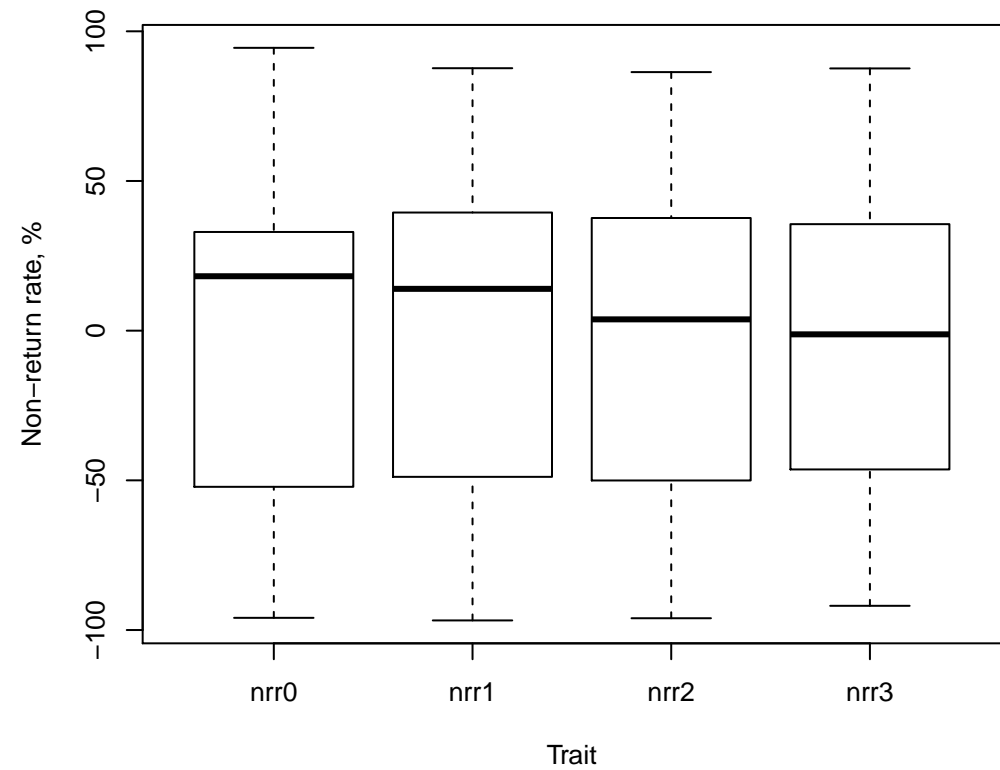
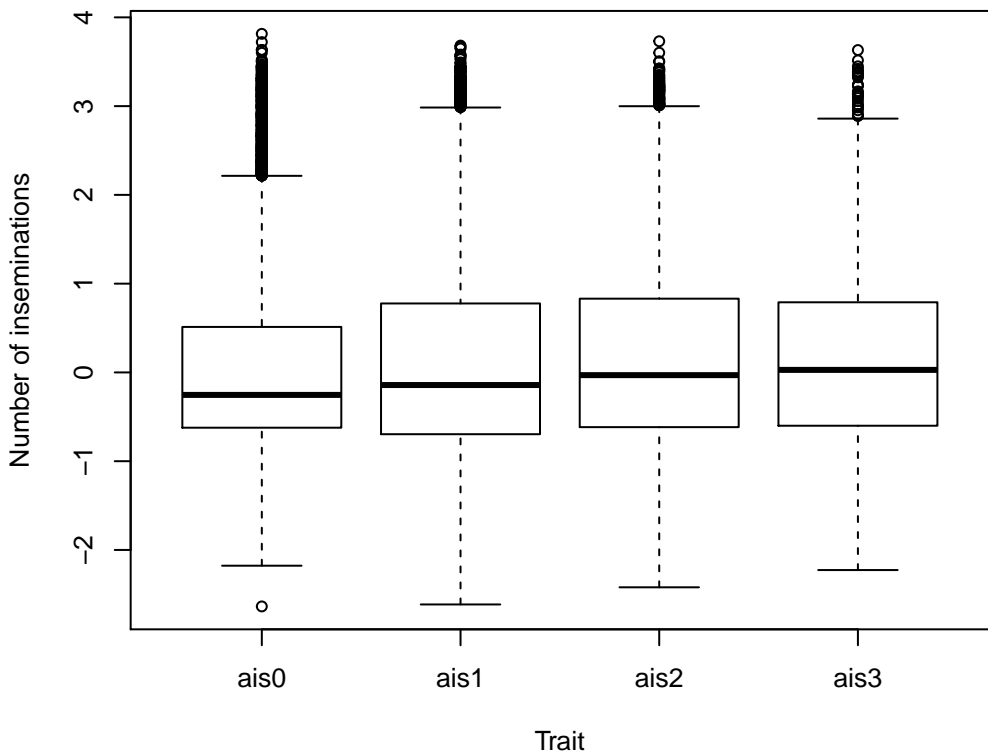
471 \*P-values < 0.1, \*\*P-value < 0.05, \*\*\*P-value < 0.01

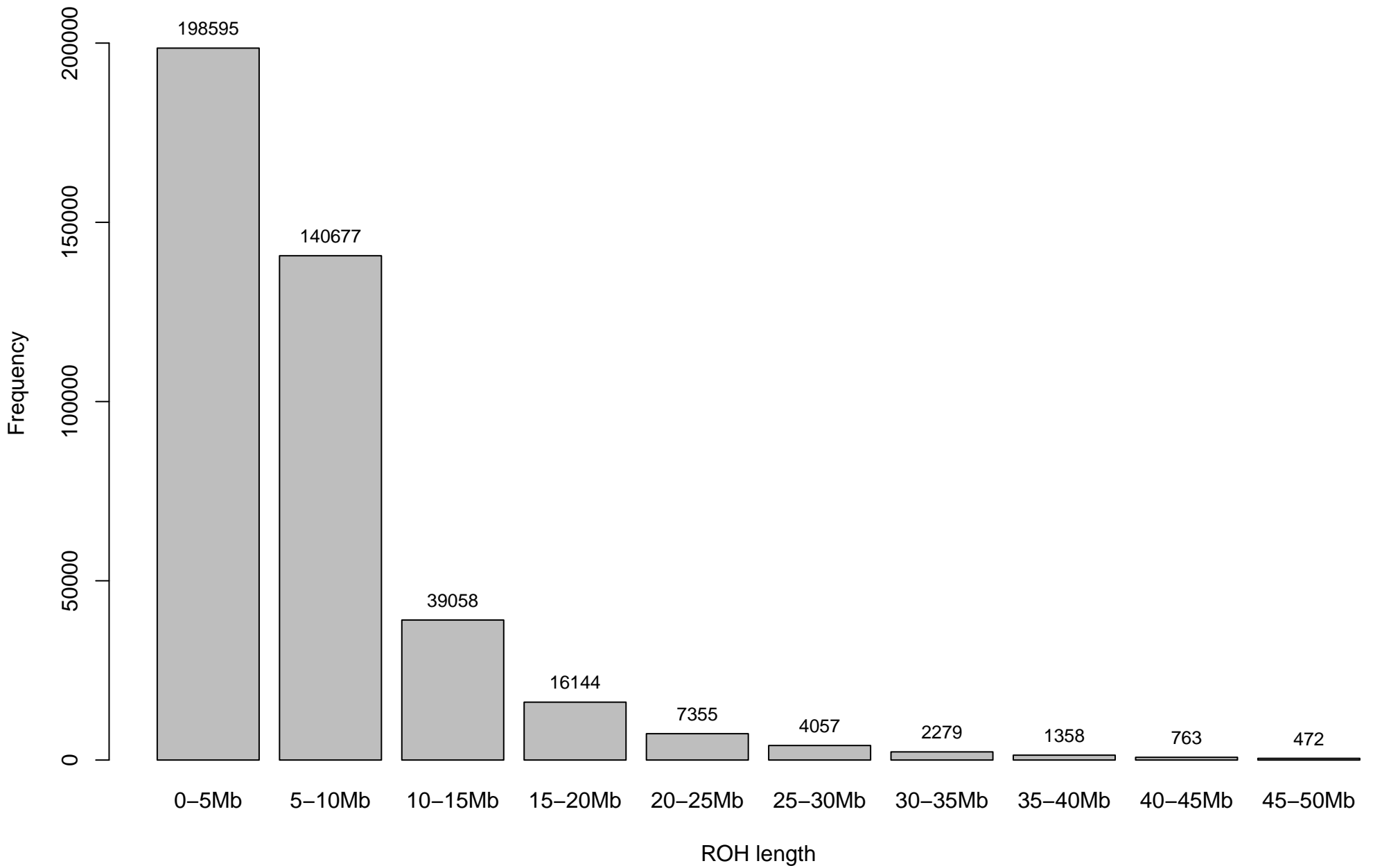
472

473 **Table 8** Estimates of inbreeding depression (standard error in brackets) for interval from calving to  
 474 first insemination (ICF) in parities 0–3 for RDCFIN and FAY50 data sets.

|      |        | F <sub>PED</sub> | F <sub>ROH_1</sub> | F <sub>PH</sub> |
|------|--------|------------------|--------------------|-----------------|
| ICF1 | RDCFIN | 15.1 (22.7)      | 4.0 (12.6)         | 7.1 (29.8)      |
|      | FAY50  | 4.8 (29.6)       | 13.0 (16.6)        | 15.9 (39.2)     |
| ICF2 | RDCFIN | 20.4 (28.9)      | 28.0*(16.1)        | 44.4 (38.4)     |
|      | FAY50  | 8.6 (37.4)       | 39.3*(21.1)        | 65.7 (49.9)     |
| ICF3 | RDCFIN | 15.6 (49.2)      | -3.6 (28.3)        | -8.3 (70.0)     |
|      | FAY50  | -18.3 (58.8)     | -14.4 (35.0)       | -30.0 (84.4)    |

475 \*P-values < 0.1, \*\*P-value < 0.05, \*\*\*P-value < 0.01







Inbreeding coefficient

