Interpretive summary

Estimation of intra-chromosomal inbreeding depression on female fertility using runs of homozygosity in Finnish Ayrshire cattle, Martikainen et al.

Inbreeding depression impairs fertility and reduces the profitability of dairy cattle production. Homozygous stretches called runs of homozygosity, which occur due to inbreeding, can be used to detect the regions showing inbreeding depression. Multiple regions in the genome of Finnish Ayrshire showed association between increased inbreeding and impaired fertility. Hence, managing inbreeding is important for limiting the impact of inbreeding depression on fertility.

Estimation of intra-chromosomal inbreeding depression on female fertility using runs of homozygosity in Finnish Ayrshire cattle

K. Martikainen*, A. Sironen, † and P. Uimari*

*Department of Agricultural Sciences, University of Helsinki, Helsinki, Finland

†Natural Resources Institute Finland (Luke), Green Technology, Jokioinen, Finland.

Correspondence

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

Telephone: +358504489237

E-mail: katja.martikainen@helsinki.fi

ABSTRACT

Inbreeding increases homozygosity, which in turn increases the frequency of harmful recessive alleles, resulting in inbreeding depression. Inbreeding depression on fertility reduces the profitability of dairy farming by decreasing the lifetime milk production of cows and by increasing insemination and veterinary costs. Continuous homozygous segments called runs of homozygosity (ROH) are currently considered to provide an effective measure of genomic inbreeding. The aim of this study was to estimate the effect of increased intra-chromosomal homozygosity for female fertility in the Finnish Ayrshire population using ROH and haplotype analysis. Genotypes were obtained from 13,712 females with the Illumina BovineLD v.2 BeadChip low-density panel and imputed to 50K density. After quality control, 40,554 SNP (single nucleotide polymorphism) remained for the analysis. Phenotypic data consisted of records for non-return rate (NRR), intervals from first to last insemination (IFL), and intervals from calving to first insemination (ICF). The raw phenotypic values were pre-adjusted for systematic effects prior to statistical analyses. ROH-based inbreeding coefficients (F_{ROH}) were used as covariates in the mixed model equation to estimate the association between inbreeding and inbreeding depression on female fertility. First, we estimated the effect of increased chromosomal F_{ROH}. We detected significant inbreeding depression on IFL. Based on our results, a 10% increase in F_{ROH} on chromosomes 2, 18, and 22 were associated with IFL of heifers lengthening by 1.6, 0.9, and 0.7 days, respectively. Similarly, a 10% increase in F_{ROH} on chromosome 15 was associated with IFL of second parity cows increasing by 2.3 days. Next, we located the regions within the chromosomes showing inbreeding depression. Our analysis revealed regions near the beginning of chromosome 2 and towards the ends of chromosomes 15, 18, and 22 that were associated with inbreeding depression on IFL. Last, we performed a haplotype analysis for the detected regions. The most promising haplotypes of each region were associated with IFL of heifers increasing by 4.4, 3.2, and 4.1 days on chromosomes 2, 18, and 22, respectively. The haplotype on chromosome 15 associated with IFL of second parity cows increasing by 7.6 days. Overall, the breeding program requires inbreeding control, as increased genomic inbreeding in our study was associated with reduced reproductive ability in Finnish Ayrshire cattle.

Key Words: cattle, fertility, inbreeding, inbreeding depression, runs of homozygosity

INTRODUCTION

Intensive selection for improved milk production traits and the widespread use of a limited number of artificial insemination (**AI**) sires has resulted in reduced effective population size and increased levels of inbreeding in many dairy cattle populations. Inbreeding increases homozygosity throughout the genome and can lead to inbreeding depression in fitness traits, such as fertility, as a result of an increased frequency in recessive deleterious alleles (Falconer and Mackay, 1996). As fertility is an economically important trait in dairy cows, problems with fertility can lead to reduced profitability by decreasing lifetime milk production and by increasing the number of inseminations, veterinary treatments, and replacement heifers due to cullings.

Pedigree information has traditionally been used to estimate inbreeding coefficients (F). Pedigree-based estimates of F are commonly underestimated due to a limited number of known generations, making distant parental relationships undetectable. Estimation of pedigree-based F can also be distorted by inconsistencies and errors in ancestry recording. In addition, pedigree-based F is an expectation of the proportion of alleles that are identical by descent from a common ancestor, and it ignores the variation caused by recombination (Keller et al., 2011). The availability of **SNP** (single nucleotide polymorphism) panels has enabled the estimation of F from genomic data without pedigree records. Genomic data also enables the estimation of F on certain regions of the genome such as those harboring QTL (Saura et al., 2015).

Continuous homozygous segments called runs of homozygosity (**ROH**) are currently considered the best genomic estimates of inbreeding (Curik et al., 2013). Runs of homozygosity additionally have an increased probability of carrying deleterious recessive alleles, and can therefore be connected to inbreeding depression (Szpiech et al., 2013; Curik et al., 2014; Bosse et al., 2015). Multiple studies show an association between increased ROH-based inbreeding ($\mathbf{F_{ROH}}$) and impaired fertility. For example, Bjelland et al. (2013) reported an increase of 1.72 days in days open per a 1% increase in F_{ROH} and Martikainen et al. (2017) estimated that a 10% increase in F_{ROH} was associated with a five-day increase between the first and last insemination in heifers. Furthermore, as inbreeding patterns vary between different regions of the genome (Kleinman-Ruiz et

al., 2016), so does the association between ROH and fertility. Pryce et al. (2014) reported an unfavorable association between ROH and calving interval on chromosomes 2, 5, 8, 9, 15, and 24 for the Holstein breed. A study of US Jersey cattle by Kim et al. (2015) found a negative association between ROH and daughter pregnancy rate on chromosomes 3, 7, 8, and 12.

The aim of our study was to estimate associations between ROH and female fertility traits between and within chromosomes in the Finnish Ayrshire population. We further examined the chromosomal regions showing significant inbreeding depression on fertility using varying lengths of haplotypes.

MATERIALS AND METHODS

Genomic Data

Genomic data used in this study were the same as in Martikainen et al. (2017), and were obtained from the Nordic Cattle Genetic Evaluation (**NAV**) (Aarhus, Denmark). The original SNP genotypes of 19,075 Finnish Ayrshire females were generated using the Illumina BovineLD v.2 BeadChip low-density panel (Illumina Inc., 2015), which contains 7,931 SNP. The genotypes were further imputed into 50K density using the Fimpute software (Sargolzaei et al., 2014) with the default parameters and a reference population of Nordic Red AI-bulls with 50K genotypes. SNP positions were based on Illumina's assembly "Native platform". The imputed genotypes were pruned so that SNP with minor allele frequency (**MAF**) less than 0.05 were removed from the data. Deviation from the Hardy-Weinberg equilibrium (**HWE**) was not used as exclusion criteria because inbreeding is one possible reason for the deviation. However, removing SNP with significant ($P < 10^{-4}$) deviation from HWE (1,476 SNP) had no effect on the results (data not shown). After pruning, a total of 40,554 SNP remained for the analysis.

Phenotypic Data

Phenotypic data included 13,712 animals with both genotypes and phenotypes available. The same data were also used in Martikainen et al. (2017). Phenotypic data consisted of records for the non-return rate at 56 days after first insemination (**NRR**) and the intervals (in days) from calving to first insemination (**ICF**) and

from first to last insemination (**IFL**). Traits were considered separately for heifers (parity 0), for first parity cows, for second parity cows, and for third parity cows. Prior to statistical analyses the raw phenotypic values were pre-adjusted for the main systematic effects, e.g. heifer's age at first insemination, herd-birth year or herd-year of first calving (for heifers and cows, respectively), insemination year-month (NRR and IFL), and calving year-month (ICF). The phenotypic data and solutions for the systematic effects were obtained from NAV and The Finnish Animal Breeding Association (**Faba**) (Vantaa, Finland). When combined with the available genomic data, the sub-sample consisted of a total 13,712 animals, with both genotypes and phenotypes included.

Estimation of Inbreeding Coefficients

Runs of homozygosity were determined prior to the estimation of intra-chromosomal inbreeding coefficients (F_{ROH}). The software PLINK v1.07 (Purcell et al., 2007) was used to detect ROH with the following criteria (originally used by Purfield et al., 2012.): a minimum density of 1 SNP per 120 kb, a minimum ROH length of 500 kb, no limit for the number of SNP per ROH, and one possible heterozygous genotype allowed per window (the corresponding PLINK parameters are --homozyg-density 120 --homozyg-kb 500 --homozyg-snp 0 --homozyg-window-het 1). These settings were selected based on a comparison of three different parameter settings from the previous study by Martikainen et al. (2017).

After ROH determination, chromosomal F_{ROH} was estimated for each autosomal chromosome K as:

$$F_{ROH\_K} = \frac{ROH\_SNP_K}{N\_SNP_K},$$

where ROH_SNP_K is the number of SNP in ROH on chromosome K and N_SNP_K is the total number of SNP on that chromosome. Based on chromosomal analysis, only chromosomes associated with inbreeding depression were selected for the intra-chromosomal study. Within chromosomes, the inbreeding estimation was performed regionally using a sliding window approach. The window moved along the chromosome in steps that were half of the window size until the end of the chromosome was reached. For each selected

chromosome, the first window contained half of the chromosome's SNP. Window size was then reduced by halving it for each round and the analysis was re-run over each selected chromosome. For each window, F_{ROH} was estimated in the same way as chromosomal F_{ROH}: the number of SNP_is in ROH divided by the total number of SNP in the window.

Estimation of Inbreeding Depression

Inbreeding depression on the chromosomal level and within significant chromosomes were estimated separately for each trait (NRR, ICF, and IFL) using a linear mixed model $y_i = \mu + bF_{ROH\_i} + u_i + e_i$, where y_i is a pre-adjusted phenotype of animal i, b is the estimate of inbreeding depression, F_{ROH_i} is the ROH-based inbreeding coefficient of animal i related either on given chromosome or window, u_i is an additive genetic effect and e_i is a residual effect of animal i, respectively. Within a trait (NRR, ICF, and IFL), parities 0 (heifers for NRR and IFL), 1, 2 and 3, were analyzed jointly with a multi-trait model. Thus, additive genetic effects were normally distributed with $N(\mathbf{0}, \mathbf{A} \otimes \mathbf{G})$, where $\mathbf{A}$ is the pedigree-based additive relationship matrix and $\mathbf{G}$ is the additive genetic variance-covariance matrix between the parities, and residual effects were normally distributed with $N(\mathbf{0}, \mathbf{R})$. Variances and covariances were the same as in the Nordic fertility evaluation (K. Muuttoranta, Natural Resources Institute Finland (Luke), Jokioinen, Finland, personal communication). Genetic groups for animals with unknown parents were treated as random effects in the analysis. Estimation of inbreeding depression was conducted using the DMU program package (Madsen and Jensen, 2013).

Haplotype Analysis

The chromosomal regions indicating inbreeding depression were further investigated for haplotypes that show impaired fertility in a homozygous state. An association of a certain haplotype j on a fertility trait was first tested with a linear model without considering relatedness between animals (no additive genetic effect in the model): $y_i = \mu + g_{ij} + e_i$ (g_{ij} is given a value of 1 if the cow i is homozygous for the tested haplotype and 0 otherwise, *i.e.* a recessive mode of inheritance was assumed) to save computing time. All haplotypes with a length of 1 to 10 SNP were tested and the haplotypes were formed using a sliding approach (in 1 SNP steps) along the selected chromosomal regions. To test whether haplotypes longer than 10 SNP gave a stronger

association than the best haplotype, the length of the haplotype was extended up to 30 SNP around the best haplotype. For each region, the haplotype with the smallest P-value (only haplotypes with a homozygous frequency greater than 5% were considered) was validated using a full model $y_i = \mu + g_{ij} + u_i + e_i$. In the validation step, an additive mode of inheritance (g_{ij} equals to number of copies of the haplotype j a cow i has, either 0, 1, or 2) was also tested. Haplotype analysis with a different length of haplotypes and sliding windows creates a large number of tests with correlated P-values. Bonferroni correction to define a threshold for a significant P-value would have been too stringent and a permutation-based test was too laborious. Thus, instead of claiming significance we report the best haplotypes of all studied chromosomal regions with nominal unadjusted P-values.

RESULTS

Runs of Homozygosity and Inbreeding

We found 426,753 ROH in the population of 19,075 genotyped animals. As expected, the number of detected ROH increased (P-value 1.01e-11) with the increase in chromosome length (measured as number of SNP in chromosome). The average number of ROH per chromosome was 14,716, varying from 30,435 (chromosome 1) to 6,866 (chromosome 27) (Figure 1A).

Median of the number of SNP in ROH across all autosomal chromosomes was 82, and it varied from 67 (chromosome 29) to 95 (chromosome 8) (Figure 1B). Median length of ROH across all autosomal chromosomes was 5.0 Mb, and it varied from 3.89 Mb (chromosome 29) to 5.77 Mb (chromosome 5) (Figure 1C). Figures 1B and 1C show that the number of long ROH (number of SNP in ROH or length of ROH) decrease as the chromosome length decreases. In addition, the distributions of number of SNP in ROH or length of ROH were skewed to the right. This indicates that the majority of detected ROH fell at the shorter end of the range of values, with a tail on the right consisting of very long ROH.

Median F_{ROH} across all autosomal chromosomes was 0.027 and it varied from 0.00 (chromosomes 12, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, and 29) to 0.05 (chromosomes 1, 6, 13, 14, and 22) (Figure 1D). Also the distribution of F_{ROH} was skewed to the right, as the estimated F_{ROH} was low for the majority of the individuals

with some individuals having high estimates for F_{ROH}. However, no association was observed between chromosome size and the estimates of F_{ROH}.

Chromosomes Associated with Inbreeding Depression

An increase in chromosomal F_{ROH} was associated with inbreeding depression on twelve chromosomes: 1, 2, 6, 8, 9, 13, 14, 15, 18, 21, 22, and 27 (P-value < 0.05, Table 1). When adjusting the P-value for the number of tests (here 29 chromosomes and 11 traits), only chromosome 2 and IFL0 remained statistically significant (P-value < 0.05/(29*11)) at a 5% significance level. However, Bonferroni correction assumes independent testing and the fertility traits are correlated between parities and between IFL, NRR, and IFC. We therefore also selected chromosomes with P-values smaller than 0.01 for more detailed intra-chromosomal analysis. The estimated effect of inbreeding on IFL0 was 1.6 days for every 10% increase in F_{ROH} on chromosome 2. Similarly, a 10% increase in F_{ROH} on chromosome 15 was associated with the lengthening of IFL2 by 2.3 days, F_{ROH} on chromosome 18 with the lengthening of IFL0 by 0.9 days, and F_{ROH} on chromosome 22 with the lengthening of IFL0 by 0.7 days.

Chromosomal Regions Associated with Inbreeding Depression

Regions within chromosomes associated with IFL lengthening were analyzed on chromosomes 2, 15, 18, and 22. Results for chromosome 2 are presented in Table 2. The first screening was performed with a window size of 1,076 SNP. The windows overlapped by 538 SNP. Based on the results, a 10% increase in F_{ROH} on the first window (window 1^1: rs109402527 - rs41576795) was associated with the lengthening of IFL0 by 1.6 days (P-value = 4.52E-07). When window size was reduced to 538 SNP, the first, second, and third windows (window 2^1: rs109402527 - rs109949037, window 2^2: rs41630342 - rs109375535, and window 2^3: rs29012101 - rs41576795) showed inbreeding depression on IFL0 (P < 0.01). A 10% increase in F_{ROH} in these windows was associated with the lengthening of IFL0 of approximately one day in each of the aforementioned intervals. These windows overlap the first window of the first screening. As window size was further reduced to 270 SNP, the first six windows were associated with the lengthening of IFL0 (P < 0.01). The estimated effect of a

10% increase in F_{ROH} in these windows varied from 0.5 days (window 3^3: rs41660074 - rs110616599) to 0.8 days (window 3^6: rs29019574 - rs42320820).

Table 3 presents the regional results of chromosome 15. The first screening was performed with a window size of 661 SNP. Based on our results, a 10% increase in F_{ROH} on the last window (window 1^3: rs109462292 - rs109571290) was associated with the lengthening of IFL by 2.2 days for second parity cows ($P < 0.01$). When window size was reduced to 330 SNP, the last two windows (window 2^6: rs41582685 - rs41568421 and window 2^7: rs110890928 - rs29012401) were associated with the lengthening of IFL2 ($P < 0.01$). The effect of these windows were 1.4 days and 1.7 days with a 10% increase in F_{ROH}, respectively.

The first screening on chromosome 18 with a window size of 532 SNP did not reveal any significant regions (Table 4). When the window size was reduced to 266 SNP, the last window (window 2^7: rs41610988 - rs29014989) showed inbreeding depression on IFL0 with a P-value <0.01. The estimated effect of a 10% increase in F_{ROH} in the last window was 0.5 days.

Table 5 presents the regional results of chromosome 22. Based on results from a window sized 490 SNP, a 10% increase in F_{ROH} on the first and second windows (window 1^1: rs42940539 - rs43709794 and window 1^2: rs41999532 - rs41603517) was associated with the lengthening of IFL0 by approximately 0.6 days ($P < 0.01$). When window size was reduced to 245 SNP, the first two windows and fifth window (window 2^1: rs42940539, - rs41998685, window 2^2: rs29023397 - rs109480467, and window 2^5: rs109028040 - rs42013343) were associated with the lengthening of IFL0 by approximately 0.5 days ($P < 0.01$).

Haplotype Results

The rapid screening of haplotypes (ignoring the relatedness of the cows) revealed several haplotypes that were associated with either IFL0 (chromosomes 2, 18, and 22) or IFL2 (chromosome 15). The best haplotypes were validated using the full model by taking into account the relatedness of the cows. The selected haplotypes were six (haplotype 1, P-value 8.03e-06), eight (haplotype 2, P-value 1.44e-06), four SNP (haplotype 3, P-value 3.64e-05), and five SNP (haplotype 4, P-value 9.4e-04) in length on chromosomes 2, 15, 18, and 22, respectively. Using the full model and a recessive mode of inheritance confirmed the results from the initial screening (Table 6). We estimated heifers that were homozygous for haplotype 1 on chromosome 2, haplotype

3 on chromosome 18, or haplotype 4 on chromosome 22 to have IFLs of 4.4, 3.2, or 4.1 days longer compared to heifers that were not homozygous for the haplotype, respectively. Similarly, haplotype 2 on chromosome 15 was estimated to increase IFL of the second parity cows by 7.6 days. Table 7 shows the results from the additive model. The obtained results support the recessive mode of inheritance in all these haplotypes. Given that the associated haplotypes were on different chromosomes, the joint effect of these haplotypes was close to additive, e.g. cows with homozygous haplotypes 1, 3, and 4 had IFL0s of 11.7 days longer than the average cows in our dataset.

DISCUSSION

Runs of Homozygosity and Inbreeding Coefficients

We investigated ROH patterns between chromosomes. The number of ROH detected increased as the size of the chromosome increased, which was expected. The size of ROH (number of SNP in ROH or length of ROH) varied between chromosomes, but showed a trend where longer chromosomes had larger-sized ROH. The parameters used in ROH detection can affect the number and length of ROH (Howrigan et al., 2011). McQuillan et al. (2008) noted that accepting all ROH longer than 0.5 Mb might generate some ROH resulting from linkage disequilibrium (**LD**) rather than from parental relatedness. Ferenčaković et al. (2013) also reported that the 50K panel overestimated the number of shorter ROH (<5 Mb). In our present study, half of all ROH (213,626 out of 426,753) were shorter than 5 Mb. The parameters used in our study therefore possibly overestimated the total number of ROH, as some of the shorter ROH we identified might be false-positives resulting from LD or panel density. False-positive short ROH would also have resulted in the underestimation of the mean size of ROH. In addition, ROH overestimation would have led to overestimation of the inbreeding coefficients.

Despite the possible bias in ROH detection, we expect the variation of F_{ROH} between chromosomes to reflect the variation in genetic diversity along the genome. Regions with higher estimates of F_{ROH} could indicate selection advantage, as the selection of favorable alleles will increase the frequency of these alleles, but also the frequency of genetically linked neutral sites (“hitch-hiking”, Maynard-Smith and Haigh, 1974).

Therefore, higher levels of inbreeding will be detected on sites under selection. Correspondingly, regions with lower F_{ROH} levels can contain selection disadvantage. However, our results concerning the patterns of genetic diversity are suggestive, and more careful consideration of the ROH parameters should be performed if the purpose is to accurately estimate the variation in genetic diversity along the genome.

Inbreeding Depression

The data used in our study consisted of genotypes and phenotypes from 13,712 animals, which is expected to result in acceptable power of the regression coefficient. Using ROH in the estimation of inbreeding depression revealed four chromosomes (2, 15, 18, and 22) in the Finnish Ayrshire population, where increased F_{ROH} was associated with the lengthening of IFL ($P < 0.01$). The estimated effects of a 10% increase in F_{ROH} on chromosomes 2, 18, and 22 on IFL0 were 1.64 days, 0.85 days, and 0.69 days, respectively. The sum of these effects is 3.18 days, whereas the estimated effect of a 10% increase in F_{ROH} over the genome was 4.32 days ($P < 0.01$) in a previous study by Martikainen et al. (2017) using the same dataset. When adding the effects of chromosomes 8 and 21 (0.71 days and 0.74 days, $P < 0.05$), the sum of the effects is 4.63 days, which is close to the estimate over the entire genome. Even though these two estimates are similar, it is not expected that only five chromosomes explain all the inbreeding depression on IFL0. Also, the previous study (Martikainen et al. 2017) showed no significant effect of F_{ROH} on IFL2, whereas in our study a 10% increase in F_{ROH} in chromosome 15 was associated with a 2.3-day lengthening of IFL2.

We estimated inbreeding depression within chromosomes applying a sliding window approach. The intra-chromosomal estimates of inbreeding depression on IFL varied from 0.5 days (IFL0 on chromosome 22 with a window size of 245 SNP) to 2.2 days (IFL2 on chromosome 15 with a window size of 661 SNP). The estimated effect decreased in all chromosomes as the window size decreased. This was expected, as inbreeding depression is mostly caused by cumulative effects of many mildly deleterious mutations that are not easily purged, unlike large-effect mutations (Charlesworth and Willis, 2009).

The sliding window method we used to investigate inbreeding depression locally using ROH has certain limitations. First, in each window, only SNP belonging to that particular window were considered, and ROH reaching over the border of the window were broken down. Some long ROH may also have been longer than

the window size. Therefore, the analyzed homozygosity could have consisted of full and partial ROHs especially in the smaller windows. Pryce et al. (2014) reported that only ROH longer than 60 SNP or 3.5 Mb were associated with inbreeding depression. They suggested that the lack of association between short ROH and inbreeding depression is explained by the purging of deleterious mutations through natural selection. In addition, they noted that ROH with a significant association with inbreeding depression formed clusters at overlapping positions on the genome, indicating that these ROH were in LD with the same QTL. Although we used window overlapping to take into account the breaking of ROH, the window possibly separated the ROH clusters and broke down long ROHs, particularly with the smaller window size, thus affecting the analysis accuracy. Second, Gomez-Raya et al. (2015) noted an additional problem when using ROH in F estimation: the same F estimate can be derived from various proportions of short and long ROH. In these cases, the effects of long ROH containing deleterious mutations would be undetectable. In our present study, the same estimates of F within the window could have been due to either a long ROH or multiple short ROH that would also have complicated the estimation of inbreeding depression. Third, Howard et al. (2017) pointed out that the unfavorable effect within a region of interest is likely due to a single ROH genotype. If a region contains some ROH with unfavorable effect together with a large number of ROH with no unfavorable effect, the regression coefficient will be pushed towards zero thus hiding the effect of the unfavorable ROH. In our present study we did not estimate the effect of each ROH separately which might have further complicated the estimation of inbreeding depression. Despite these limitations, the sliding-window approach proved an effective way of locating chromosomal regions with increased inbreeding depression.

Haplotypes Associated with Fertility

To investigate further whether the detected inbreeding depression was due to certain common haplotypes and recessive QTL within these haplotypes, we tested haplotypes of length 1–10, 20, and 30 SNP within the significant chromosomal regions. We found that the length of the most promising haplotypes varied from four to eight SNP between regions. This is consistent with previous studies concerning optimal haplotype length. For example, Grapes et al. (2006) studied haplotypes of length 1, 2, 4, 6, and 10 SNP, and discovered that the greatest QTL detection accuracy was obtained with haplotypes of 4 to 6 SNP in length. In addition, Pryce et

al. (2010) investigated haplotypes of length 3, 5, 7, and 9 SNP, and showed that the precision of mapping QTL affecting fertility was greatest when haplotype length was 7 SNP. They observed that optimal haplotype length is affected by the patterns of LD and the frequency of the QTL alleles.

We found several haplotypes that showed impaired fertility at the recessive mode of inheritance. However, the proportion of homozygous haplotypes belonging to any detected ROH varied. Haplotype 1 on chromosome 2 was a part of ROH on 26.5% of the individuals. Correspondingly, for haplotype 2 on chromosome 15, haplotype 3 on chromosome 18, and haplotype 4 on chromosome 22 the proportions were 51.4%, 18.5%, and 31.9%, respectively. Homozygosity in a short region could be due to inbreeding or due to random haplotype pairing. In addition, the effect of the haplotype can be due to one or possibly several harmful mutations of various origins. Thus we do not expect full concordance with ROH and haplotype-based approaches. A recessive haplotype with a strong negative effect on fertility could indicated inbreeding depression despite only part of the observed homozygous haplotypes being in ROH. Given that the detection of ROH depends on SNP density, the proportion of harmful recessive haplotypes belonging to ROH could be greater if more SNP or full sequence data were available for ROH detection. However, if the inbreeding depression is due to several mildly deleterious mutations within a chromosomal region, haplotype analysis may not discover any haplotypes showing a large enough effect on fertility despite the ROH-based approach indicating inbreeding depression.

Candidate Genes within Fertility-associated Haplotypes

The effect on inbreeding depression comes from the functional changes in genes within fertility-associated regions. Thus, genomic alterations in candidate genes with a role in fertility can explain inbreeding depression. The haplotype regions in chromosomes 2, 15, and 22 contained only a few genes, but the region in chromosome 18 included several genes or uncharacterized proteins. The Serine/Threonine Kinase 2 (*Akt2)* gene adjacent to the fertility-associated region on chromosome 18 (49,904,012–49,950,072 bp) has been shown to affect fertility in mice (Restuccia et al., 2012). Akt2 is a serine/threonine protein kinase, the absence of which results in increased testosterone levels and development of ovarian cysts. Furthermore, decrease in male fertility has been reported in *Akt2* knockout mice due to decreased glucose transporter levels and

increased apoptosis during sperm development (Kim et al., 2012). The correct glucose transporter levels are also crucial for mouse blastocyst development (Carayannopoulos et al., 2000) and Akt2 may therefore have a role in the female reproduction system in addition to its effect on testosterone levels and male fertility.

Interestingly, a galactosyltransferase gene beta-1,3-galactosyltransferase 1 *(B3GALT1)* is localized close to the haplotype on chromosome 2 (28,400,755–28,401,735 bp). *B3GALT1* is a member of the beta-1,3-galactosyltransferase (*beta3GalT*) gene family, which encodes type II membrane-bound glycoproteins. Absence of the core 1 β1,3-galactosyltransferase 1 enzyme (C1β3galt1*)* causes modified follicular development leading to the maturation and ovulation of more follicles and increased fertility (Williams and Stanley, 2008). Oocyte-specific double deletion of Alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase (*Mgat1)* and *C1galt1* results in female subfertility in mice at six weeks and infertility by nine weeks (Grasa et al., 2016). In *C1galt1* mutant mice, modified oocyte glycoproteins alter the ratio of Growth/differentiation factor 9 (GDF9): Bone Morphogenetic Protein 15 (BMP15) expression. The altered expression ratio modifies follicle development resulting in the generation of more follicles and an extended estrous cycle (Grasa et al., 2015). These results underline important roles for glycoprotein regulators of follicular integrity and fertility.

The haplotype region in chromosome 15 has been associated with milk yield during first lactation (Saowaphak et al., 2017). This region includes the Leucine-rich Repeat Containing 4C (*LRRC4C)* gene (15: 71,243,845–71,245,767 bp), which encodes a transmembrane protein belonging to the netrin-G family of cell adhesion molecules. A copy number variation (CNV) covering the *LRRC4C* gene region in humans has been identified in infertile meiosis arrest male patients (Eggers et al., 2015). LRRC4C was localized in meiotic cells, which may indicate a role in female meiosis in addition to male fertility. The main role of LRRC4C has been implicated in promoting the presynaptic differentiation and growth of axons in the brain (Lin et al., 2003; Song et al., 2013). The identification of exact biological reasons for the inbreeding depression would require functional studies of the genome regions and associated genes such as *B3GALT1* in chromosome 2. As a first step towards this goal our study identified the ROH levels across chromosomes and fertility-associated haplotypes with candidate genes in Finnish Ayrshire cattle.

CONCLUSIONS

In our study, we showed that inbreeding patterns vary across the genome in Finnish Ayrshire cattle. Using ROH, we were able to locate regions in the genome showing inbreeding depression on female fertility. On these regions, the increased ROH-based inbreeding coefficient was associated with the lengthening of the female fertility trait interval from first to last inseminations (IFL). We also investigated the effect of increased homozygosity by performing a haplotype analysis, which revealed common homozygous haplotypes that were associated with the lengthening of IFL. Overall, balancing between genetic gain and levels of inbreeding in the breeding program is needed, as the increased genomic inbreeding observed in our study was associated with reduced reproductive ability in dairy cows.

ACKNOWLEDGEMENTS

The authors thank the funders of this research, Suomen Naudanjalostussäätiö and the Ministry of Agriculture and Forestry. We are grateful to NAV (Nordic Cattle Genetic Evaluation) and Faba (The Finnish Animal Breeding Association) for providing the genotypes, pedigree data, raw phenotypes, solutions for the fixed effects, and estimates of (co)variance-components used in our study.

REFERENCES

Bjelland, D. W., K. A. Weigel, N. Vukasinovic and J. D. Nkrumah. 2013. Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. J. Dairy Sci. 96:4697-4706. https://doi.org/10.3168/jds.2012-6435.

Bosse, M., H. Megens, O. Madsen, R. P. M. A. Crooijmans, O. A. Ryder, F. Austerlitz, M. A. M. Groenen and M. A. de Cara R. 2015. Using genome-wide measures of coancestry to maintain diversity and fitness

in endangered and domestic pig populations. Genome Res. 25:970-981. http://doi.org/10.1101/gr.187039.114.

Carayannopoulos, M. O., M. M. -. Chi, Y. Cui, J. M. Pingsterhaus, R. A. McKnight, M. Mueckler, S. U. Devaskar and K. H. Moley. 2000. GLUT8 is a glucose transporter responsible for insulin-stimulated glucose uptake in the blastocyst. Proc. Natl. Acad. Sci. USA. 97:7313-7318. https://doi.org/10.1073/pnas.97.13.7313.

Charlesworth, D. and J. H. Willis. 2009. The genetics of inbreeding depression. Nat. Rev. Genet. 10:783-796.

Curik, I., M. Ferenčaković and J. Sölkner. 2014. Inbreeding and runs of homozygosity: A possible solution to an old problem. Livest. Sci. 166:26-34. https://doi.org/10.1016/j.livsci.2014.05.034.

Eggers, S., K. D. DeBoer, J. van den Bergen, L. Gordon, S. J. White, D. Jamsai, R. I. McLachlan, A. H. Sinclair and M. K. O'Bryan. 2015. Copy number variation associated with meiotic arrest in idiopathic male infertility. Fertil. Steril. 103:214-219. https://doi.org/10.1016/j.fertnstert.2014.09.030.

Falconer, D. S. and T. F. C. Mackay, T. F. C. 1996. Introduction to quantitative genetics, 4th edition. Essex, UK: Longman.

Ferencakovic, M., J. Sölkner and I. Curik. 2013. Estimating autozygosity from high-throughput information: Effects of SNP density and genotyping errors. Genet. Sel. Evol. 45:42. http://doi.org/10.1186/1297-9686-45-42.

Gomez-Raya, L., C. Rodríguez, C. Barragán and L. Silió. 2015. Genomic inbreeding coefficients based on the distribution of the length of runs of homozygosity in a closed line of Iberian pigs. Genet. Sel. Evol. 47:81. https://doi.org/10.1186/s12711-015-0153-1.

Grapes, L., M. Z. Firat, J. C. M. Dekkers, M. F. Rothschild and R. L. Fernando. 2006. Optimal haplotype structure for linkage disequilibrium-based fine mapping of quantitative trait loci using identity by descent. Genetics. 172:1955-1965. http://doi.org/10.1534/genetics.105.048686.

Grasa, P., P. Ploutarchou and S. A. Williams. 2015. Oocytes lacking O-glycans alter follicle development and increase fertility by increasing follicle FSH sensitivity, decreasing apoptosis, and modifying GDF9:BMP15 expression. FASEB J. 29:525-539. https://doi.org/10.1096/fj.14-253757.

Grasa, P., S. Sheikh, N. Krzys, K. Millar, S. Janjua, P. Nawaggi and S. A. Williams. 2016. Dysregulation of follicle development in a mouse model of premature ovarian insufficiency. Reproduction. 152:591-601. http://doi.org/10.1530/REP-16-0091.

Howard, J. T., F. Tiezzi, Y. Huang, K. A. Gray, C. Maltecca. 2017. A heuristic method to identify runs of homozygosity associated with reduced performance in livestock. J. Anim. Sci. 95: 4318–4332. https://doi.org/10.2527/jas2017.1664

Howrigan, D. P., M. A. Simonson and M. C. Keller. 2011. Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. BMC Genomics. 12:460-460. http://doi.org/10.1186/1471-2164-12-460.

Illumina Inc. 2015. Illumina Data Sheet: Agrigenomics. BovineLD v2.0 Genotyping BeadChip. Accessed Mar 8, 2018. http://www.illumina.com/Documents/products/datasheets/datasheet_bovineLD.pdf.

Keller, M. C., P. M. Visscher and M. E. Goddard. 2011. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics. 189:237-249. http://doi.org/10.1534/genetics.111.130922.

Kim, S. T., K. Omurtag and K. H. Moley. 2012. Decreased Spermatogenesis, Fertility, and Altered Slc2A Expression in Akt1−/− and Akt2−/− Testes and Sperm. Reprod. Sci. 19:31-42. http://doi.org/10.1177/1933719111424449.

Kim, E., T. S. Sonstegard, C. P. Van Tassell, G. Wiggans and M. F. Rothschild. 2015. The relationship between runs of homozygosity and inbreeding in Jersey cattle under selection. Plos One. 10:e0129967. http://doi.org/10.1371/journal.pone.0129967.

Kleinman-Ruiz, D., B. Villanueva, J. Fernández, M. A. Toro, L. A. García-Cortés and S. T. Rodríguez-Ramilo. 2016. Intra-chromosomal estimates of inbreeding and coancestry in the Spanish Holstein cattle population. Livest. Sci. 185:34-42. https://doi.org/10.1016/j.livsci.2016.01.002.

Lin, J. C., W. Ho, A. Gurney and A. Rosenthal. 2003. The netrin-G1 ligand NGL-1 promotes the outgrowth of thalamocortical axons. Nat. Neurosci. 6:1270-1276. http://dx.doi.org/10.1038/nn1148.

Madsen P. and J. Jensen. 2013. A user's guide to DMU – A Package for Analysis Multivariate Mixed Models. Version 6, release 5.2, University of Aarhus. http://dmu.agrsci.dk/DMU/Doc/Current/dmuv6_guide.5.2.pdf.

Martikainen, K., A. Tyrisevä M., K. Matilainen, J. Pösö and P. Uimari. 2017. Estimation of inbreeding depression on female fertility in the Finnish Ayrshire population. J. Anim. Breed. Genet. 134:383–392. http://dx.doi.org/10.1111/jbg.12285.

Maynard-Smith, J., and J. Haigh. 1974. The hitch-hiking effect of a favourable gene. Genet. Res. 23:23-35.

McQuillan, R., A. Leutenegger, R. Abdel-Rahman, C. S. Franklin, M. Pericic, L. Barac-Lauc, N. Smolej-Narancic, B. Janicijevic, O. Polasek, A. Tenesa, A. K. MacLeod, S. M. Farrington, P. Rudan, C. Hayward, V. Vitart, I. Rudan, S. H. Wild, M. G. Dunlop, A. F. Wright, H. Campbell and J. F. Wilson. 2008. Runs of homozygosity in European populations. Am. J. Hum. Genet. 83:359-372. http://doi.org/10.1016/j.ajhg.2008.08.007.

Pryce, J. E., S. Bolormaa, A. J. Chamberlain, P. J. Bowman, K. Savin, M. E. Goddard and B. J. Hayes. 2010. A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using variable length haplotypes. J. Dairy Sci. 93:3331-3345. https://doi.org/10.3168/jds.2009-2893.

Pryce, J. E., M. Haile-Mariam, M. E. Goddard and B. J. Hayes. 2014. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. Genet. Sel. Evol. 46:71. http://doi.org/10.1186/s12711-014-0071-7.

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

Purfield, D. C., D. P. Berry, S. McParland and D. G. Bradley. 2012. Runs of homozygosity and population history in cattle. BMC Genet. 13:70. https://doi.org/10.1186/1471-2156-13-70.

Restuccia, D. F., D. Hynx and B. A. Hemmings. 2012. Loss of PKBβ/Akt2 predisposes mice to ovarian cyst formation and increases the severity of polycystic ovary formation in vivo. Dis. Model. Mech. 5:403-411. http://doi.org/10.1242/dmm.008136.

Saowaphak P., M. Duangjinda, S. Plaengkaeo, R. Suwannasing and W. Boonkum. 2017. Genetic correlation and genome-wide association study (GWAS) of the length of productive life, days open, and 305-days milk

yield in crossbred Holstein dairy cattle. Genet. Mol. Res. 16 (2): gmr16029091. http://dx.doi.org/10.4238/gmr16029091.

Sargolzaei, M., J. P. Chesnais and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics. 15:478. https://doi.org/10.1186/1471-2164-15-478.

Saura, M., A. Fernández, L. Varona, A. I. Fernández, M. Á. R. de Cara, C. Barragán and B. Villanueva. 2015. Detecting inbreeding depression for reproductive traits in Iberian pigs using genome-wide data. Genet. Sel. Evol. 47:1. https://doi.org/10.1186/s12711-014-0081-5.

Song, Y. S., H. Lee, P. Prosselkov, S. Itohara and E. Kim. 2013. Trans-induced cis interaction in the tripartite NGL-1, netrin-G1 and LAR adhesion complex promotes development of excitatory synapses. J. Cell. Sci. 126:4926-4938.

Szpiech, Z., J. Xu, T. Pemberton, W. Peng, S. Zöllner, N. Rosenberg and J. Li. 2013. Long runs of homozygosity are enriched for deleterious variation. Am. J. Hum. Genet. 93:90-102. http://doi.org/10.1016/j.ajhg.2013.05.003.

Williams, S. A. and P. Stanley. 2008. Mouse fertility is enhanced by oocyte-specific loss of core 1-derived O-glycans. FASEB J. 22:2273-2284. http://doi.org/10.1096/fj.07-101709.

TABLES

Table 1. Chromosomes associated with inbreeding depression (P-value < 0.05)

Chromosome	Trait[1]	b^2 in days (SE)	P-value
1	IFL1	1.59 (0.71)	0.025
2	IFL0	1.64 (0.38)	1.81E-05*
2	NRR0	-1.07 (0.50)	0.032
6	ICF2	0.92 (0.42)	0.028
8	IFL0	0.71 (0.35)	0.042
9	NRR1	-1.29 (0.59)	0.030
13	NRR0	0.89 (0.41)	0.031
14	IFL1	1.25 (0.61)	0.041
15	IFL2	2.32 (0.87)	7.50E-03*
15	ICF1	0.67 (0.30)	0.027
18	IFL0	0.85 (0.31)	5.36E-03*
18	NRR0	-0.95 (0.40)	0.019
21	IFL0	0.74 (0.32)	0.020
22	IFL0	0.69 (0.26)	7.82E-03*
27	NRR1	-1.11 (0.54)	0.038

[1] IFL: interval from calving to first insemination; NRR: non-return rate at 56 days after insemination; ICF: interval from calving to first insemination. Heifers (value 0) and cows with first three parities (values 1–3) were considered separately.

[2] Estimate of inbreeding depression for a 10% increase in inbreeding coefficient

* Selected for intra-chromosomal analysis (P-value < 0.01)

Table 2. Regions on chromosome 2 associated with inbreeding depression in interval from calving to first insemination on heifers (IFL0) using three different window sizes

1st round[1]				2nd round[2]				3rd round[3]			
Window[4]	First and last SNP	b[5] (SE)	P-value	Window[4]	First and last SNP	b[5] (SE)	P-value	Window[4]	First and last SNP	b[5] (SE)	P-value
1[1]	rs109402527, rs41576795	1.58 (0.31)	4.52E-07	2[1]	rs109402527, rs109949037	0.92 (0.23)	5.31E-05	3[1]	rs109402527, rs41630342	0.67 (0.19)	3.92E-04
1[2]	rs29012101, rs110289171	0.88 (0.27)	0.001	2[2]	rs41630342, rs109375535	0.99 (0.24)	3.78E-05	3[2]	rs43043879, rs29021257	0.67 (0.19)	3.02E-04
1[3]	rs41622659, rs110731179	0.47 (0.27)	NS	2[3]	rs29012101, rs41576795	0.90 (0.25)	2.97E-04	3[3]	rs41660074, rs110616599	0.53 (0.18)	3.07E-03
				2[4]	rs43706860, rs109413238	0.30 (0.19)	NS	3[4]	rs109193035, rs109596311	0.55 (0.20)	6.37E-03
								3[5]	rs109622109, rs110708198	0.73 (0.20)	3.14E-04
								3[6]	rs29019574, rs42320820	0.78 (0.21)	2.56E-04
								3[7]	rs110693558, rs109992151	0.40 (0.19)	0.033
								3[8]	rs110139290, rs109018046	0.08 (0.16)	NS

[1]Size of the window 1,076 SNP

[2] Size of the window 538 SNP

[3] Size of the window 270 SNP

[4] X^Y, where X is the round and Y is the number of windows, numbered from the beginning of the chromosome

[5] Estimate of inbreeding depression for a 10% increase in inbreeding coefficient

546 **Table 3.** Regions within chromosome 15 associated with inbreeding depression on interval from calving to

547 first insemination on second parity cows (IFL2) using two different window sizes

1st round[1]				2nd round[2]			
Window[3]	First and last SNP	b[4] (SE)	P-value	Window[3]	First and last SNP	b[4] (SE)	P-value
1^1	rs43056590, rs109462292	0.33 (0.70)	NS	2^3	rs41753591, rs41631168	-0.06 (0.54)	NS
1^2	rs41753591, rs110890928	0.71 (0.62)	NS	2^4	rs41763024, rs110675698	0.77 (0.51)	NS
1^3	rs109462292, rs109571290	2.17 (0.63)	5.78E-04	2^5	rs109462292, rs109173346	0.83 (0.46)	NS
				2^6	rs41582685, rs41568421	1.37 (0.46)	0.003
				2^7	rs110890928, rs29012401	1.70 (0.49)	5.17E-04

548 [1]Size of the window 661 SNP

549 [2] Size of the window 330 SNP

550 [3] X^Y, where X is the round and Y is the number of the windows, numbered from the beginning of the

551 chromosome

552 [4] Estimate of inbreeding depression for a 10% increase in inbreeding coefficient

553

554

555

556

557

558

559

560

561

562

563

564

Table 4. Regions within chromosome 18 associated with inbreeding depression on interval from calving to first insemination on heifers (IFL0) using two different window sizes

1st round[1]				2nd round[2]			
Window[3]	First and last SNP	b[4] (SE)	P-value	Window[3]	First and last SNP	b[4] (SE)	P-value
1^1	rs109662639, rs110835265	0.45 (0.24)	NS	2^3	rs41636244, rs110835265	0.27 (0.18)	NS
1^2	rs41636244, rs41604526	0.38 (0.22)	NS	2^4	rs110672596, rs109456364	0.37 (0.18)	0.037
1^3	rs41637283, rs29014989	0.54 (0.23)	0.018	2^5	rs41637283, rs41604526	0.22 (0.18)	NS
				2^6	rs43706886, rs41897446	0.36 (0.19)	NS
				2^7	rs41610988, rs29014989	0.50 (0.19)	0.009

[1]Size of the window 532 SNP

[2] Size of the window 266 SNP

[3] X^Y, where X is the round and Y is the number of the windows, numbered from the beginning of the chromosome

[4] Estimate of inbreeding depression for a 10% increase in inbreeding coefficient

584 **Table 5.** Regions within chromosome 22 associated with inbreeding depression on interval from calving to

585 first insemination on heifers (IFL0) using two different window sizes

1st round[1]				2nd round[2]			
Window[3]	First and last SNP	b[4] (SE)	P-value	Window[3]	First and last SNP	b[4] (SE)	P-value
1^1	rs42940539, rs43709794	0.60 (0.22)	0.007	2^1	rs42940539, rs41998685	0.54 (0.19)	0.006
1^2	rs41999532, rs41603517	0.61 (0.19)	0.002	2^2	rs29023397, rs109480467	0.49 (0.17)	0.004
1^3	rs110298230, rs43502224	0.27 (0.18)	NS	2^3	rs41998685, rs109028040	0.26 (0.16)	NS
				2^4	rs109480467, rs109936697	0.38 (0.16)	0.02
				2^5	rs109028040, rs42013343	0.47 (0.14)	0.001

586 [1]Size of the window 490 SNP

587 [2] Size of the window 245 SNP

588 [3] X^Y, where X is the round and Y is the number of the windows, numbered from the beginning of the

589 chromosome

590 [4] Estimate of inbreeding depression for a 10% increase in inbreeding coefficient

591

592

593

594

595

596

597

598

599

600

601

602

603 **Table 6.** Effects of haplotypes on interval from calving to first insemination of heifers (IFL0) or second parity

604 cows (IFL2) estimated using the recessive model

Haplotype	CHR	Trait	First SNP	Position (Mb)	Freq. homozygotes	Number of homozygotes	b[1] (SE)	P-value
1	2	IFL0	rs109350674	28.6–28.8	0.10	1,323	4.38 (1.01)	1.55E-05
2	15	IFL2	rs41655010	70.4–70.9	0.18	2,496	7.57 (2.16)	4.73E-04
3	18	IFL0	rs41257394	49.7–49.9	0.15	2,068	3.23 (0.83)	1.01E-04
4	22	IFL0	rs42233434	42.0–42.1	0.05	728	4.13 (1.29)	0.001

605 [1] Difference between individuals homozygous / not homozygous for the haplotype

606 **Table 7.** Effects of haplotypes 1, 2, 3 and 4 on interval from calving to first insemination of heifers (IFL0) or

607 second parity cows (IFL2) estimated using the additive model

Haplotype	CHR	Trait	Number of haplotype copies	Number of cows	b[1] (SE)
1	2	IFL0	0	6,236	-1.33 (2.66)
			1	6,153	-1.55 (2.68)
			2	1,323	2.89 (2.83)
2	15	IFL2	0	4,457	-0.42 (6.95)
			1	6,759	-2.95 (6.98)
			2	2,496	5.27 (7.27)
3	18	IFL0	0	5,088	-1.47 (2.67)
			1	6,556	-1.73 (2.67)
			2	2,068	1.59 (2.75)
4	22	IFL0	0	7963	-1.29 (2.66)
			1	5021	-1.55 (2.68)
			2	728	2.71 (2.92)

608 [1]Estimates effect for individuals with 0, 1 or 2 copies of the haplotype

609

610

611

612

613

614

615

616

617

618

FIGURES

Figure 1: Number of detected ROH (A), number of SNP in ROH (B), length of ROH (C), and F_{ROH} (D) per each autosomal chromosome.

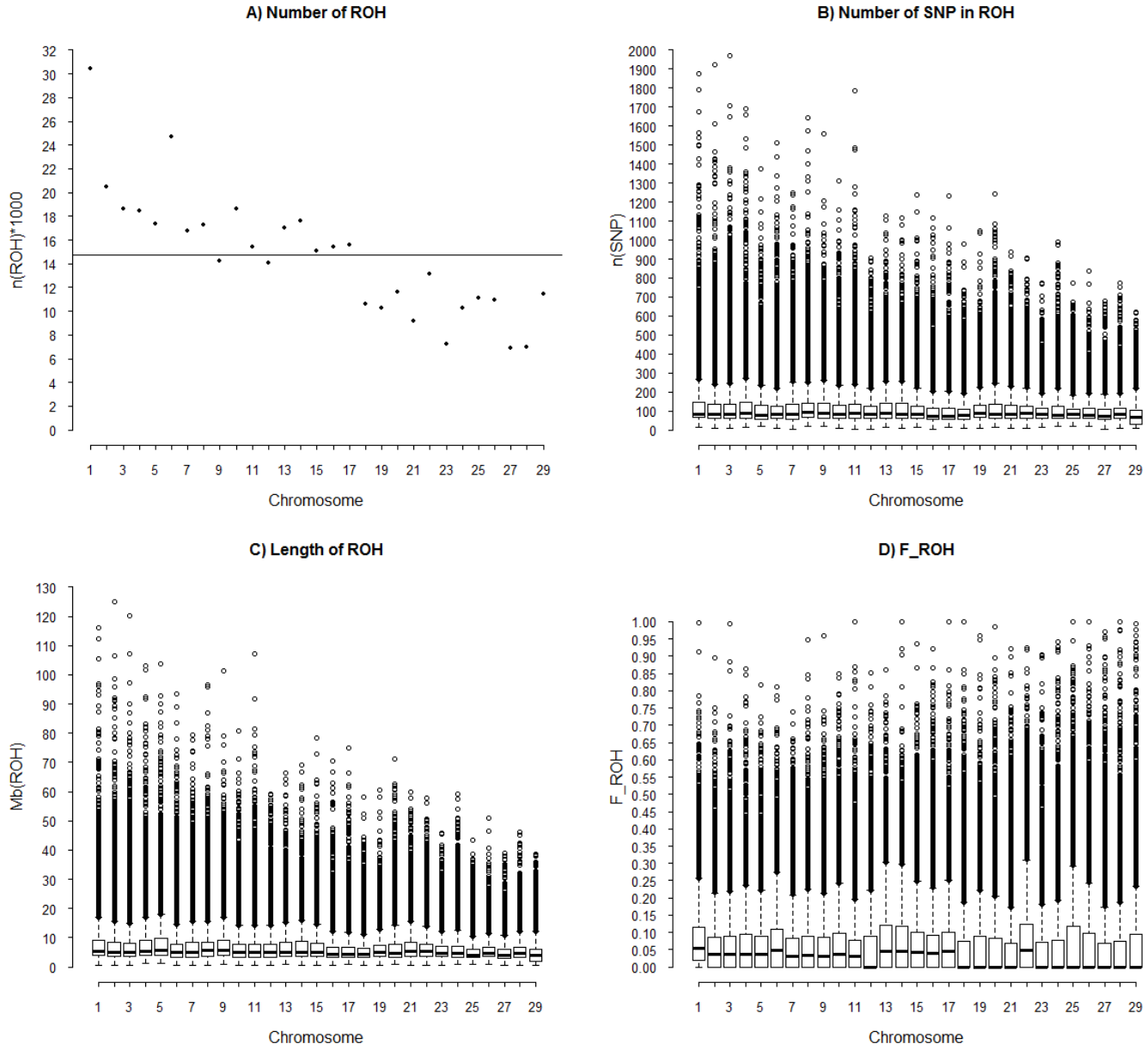

Figure 1. Number of detected ROH (A), number of SNP in ROH (B), length of ROH (C), and F_{ROH} (D) per each autosomal chromosome.

JDS-18-14805, Figure 1