

Article

Population Genomics of American Mink Using Whole Genome Sequencing Data

Karim Karimi ¹, Duy Ngoc Do ¹ , Mehdi Sargolzaei ^{2,3} and Younes Miar ^{1,*} 

¹ Department of Animal Science and Aquaculture, Dalhousie University, Truro, NS B2N 5E3, Canada; karim.karimi@dal.ca (K.K.); duy.do@dal.ca (D.N.D.)

² Department of Pathobiology, University of Guelph, Guelph, ON N1G 2W1, Canada; msargol@uoguelph.ca

³ Select Sires Inc., Plain City, OH 43064, USA

* Correspondence: miar@dal.ca

Abstract: Characterizing the genetic structure and population history can facilitate the development of genomic breeding strategies for the American mink. In this study, we used the whole genome sequences of 100 mink from the Canadian Centre for Fur Animal Research (CCFAR) at the Dalhousie Faculty of Agriculture (Truro, NS, Canada) and Millbank Fur Farm (Rockwood, ON, Canada) to investigate their population structure, genetic diversity and linkage disequilibrium (LD) patterns. Analysis of molecular variance (AMOVA) indicated that the variation among color-types was significant ($p < 0.001$) and accounted for 18% of the total variation. The admixture analysis revealed that assuming three ancestral populations ($K = 3$) provided the lowest cross-validation error (0.49). The effective population size (N_e) at five generations ago was estimated to be 99 and 50 for CCFAR and Millbank Fur Farm, respectively. The LD patterns revealed that the average r^2 reduced to <0.2 at genomic distances of >20 kb and >100 kb in CCFAR and Millbank Fur Farm suggesting that the density of 120,000 and 24,000 single nucleotide polymorphisms (SNP) would provide the adequate accuracy of genomic evaluation in these populations, respectively. These results indicated that accounting for admixture is critical for designing the SNP panels for genotype-phenotype association studies of American mink.

Keywords: American mink; linkage disequilibrium; population structure; whole genome sequencing



Citation: Karimi, K.; Ngoc Do, D.; Sargolzaei, M.; Miar, Y. Population Genomics of American Mink Using Whole Genome Sequencing Data. *Genes* **2021**, *12*, 258. <https://doi.org/10.3390/genes12020258>

Academic Editor: Martien Groenen

Received: 12 December 2020

Accepted: 5 February 2021

Published: 11 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Characterizing the genetic structure is crucial to reveal the genetic diversity, domestication history, and genetic relationship of populations that eventually facilitate the development of efficient breeding strategies in domestic animals [1]. Advances in DNA sequencing technologies have provided the opportunity to use a large number of single nucleotide polymorphisms (SNPs) for investigation of genetic structure and diversity of livestock species, e.g., cattle [2,3], sheep [4], goat [5], and pig [6]. In addition, these markers are efficient tools to measure the linkage disequilibrium (LD) across the genome [7]. The LD refers to the non-random association of alleles at two or more loci and is indicative of the correlation between two nearby loci [8]. The LD might be caused by physical proximity between markers located on the same chromosome (linkage) whereas it is eroded by recombination events [9]. In addition, LD could be generated between unlinked markers due to several factors including genetic drift, selection, epistatic combinations and population structure [7]. For instance, admixture between distinct populations can arise LD between two unlinked loci, particularly those with different allele frequencies [10,11].

The LD levels are used as a measure to determine the required density of markers for genome-wide mapping studies [12]. Moreover, the magnitude and decay of LD over genomic intervals can be used to estimate the recent and past effective population sizes (N_e). Since recombination takes less time to break down the LD between loosely linked loci, LD at longer distances reflects N_e at recent generations whereas LD from physically

linked loci is an indicator of N_e at past generations [13,14]. The N_e is the key parameter to monitor the evolutionary history, inbreeding risk and conservation priority in animal populations [15–17]. Furthermore, assessment of relationship among individuals using molecular markers provides more accurate measures of inbreeding levels to manage ongoing mating [18].

American mink (*Neovison vison*) is a semiaquatic species of Mustelid family, which is known as one of the most popular resources of fur worldwide [19]. There is no much information available on the domestication history in American mink. However, it was stated that the American mink was initially bred in captivity in 1866 to produce fur [20]. Whereas some farmers tend to raise a particular color-type of mink, most farms are mixed of different color-types. Market demands may affect the color combination of mink farms. Unfortunately, most farms have no regular mating system and animals are mostly selected based on the phenotypic performances meaning that incidental gene flow could happen among populations [21,22]. Pelt size, fur quality, reproductive performance, and disease resistance are the main traits of interest for mink breeders [23–25]. Implementation of highly efficient breeding strategies is critical for the mink industry to adopt sustainable production practices, e.g., increasing resiliency against emerging diseases, resolving the ethical issues of fur production and reducing the cost of production [25]. Despite the successful utilization of genomic selection approaches in improvement of genetic merits in other livestock species, e.g., dairy cattle [26] and pig [27], this breeding strategy has not been applied in the mink industry. Whole genome sequence data provide a large number of SNP makers that can be used for genome-wide studies in American mink and accelerate the development of modern genomic selection in this industry [28,29].

Analysis of genetic structure and genome-wide LD pattern of mink populations can provide fundamental information for developing tools required for genomic selection in American mink. The genetic structure of farm and feral American mink was previously studied using 11 microsatellite markers in Japanese mink [30], 154 SNPs generated from restriction-site associated DNA sequencing (RAD-seq) in Polish and Danish mink populations [31], and mitochondrial DNA sequences along with 11 microsatellite markers in Chilean mink [32]. Microsatellite DNA markers were also used to estimate the N_e in American mink based on the LD information [33,34]. Moreover, 13,321 SNPs extracted by genotyping-by-sequencing (GBS) was used to investigate the LD patterns and estimate the N_e of black American mink [35]. However, there is no study on the LD patterns and N_e of different color-types of American mink using the whole genome sequencing data. In addition, admixture patterns raised by the presence of different color-types on mink farms would affect the extent of LD, which is not well-characterized in American mink. Therefore, the main ideas of this study were (1) to determine the genetic structure and admixture pattern of various color-types of American mink in Canada, (2) to investigate the LD pattern across the genome for estimation of the required marker density for genome-wide studies in American mink, and (3) to estimate the genomic inbreeding coefficients and N_e in two mink farms in Canada.

2. Materials and Methods

2.1. Animals and Sampling

All protocols were approved by the Dalhousie University Animal Care and Use Committee (certification number: 2018–009), and mink used in this study were cared for according to the Code of Practice for the Care and Handling of Farmed Mink guidelines [36]. Mink were housed under standard farming conditions and humanly euthanized in December 2018 [24]. Tongue samples of 100 individuals were collected from two farms including the Canadian Center for Fur Animal Research (CCFAR) at Dalhousie Faculty of Agriculture (Truro, NS, Canada) and Millbank Fur Farm (Rockwood, ON, Canada). Animals from CCFAR consisted of various color-types including Demi ($n = 32$), Pastel ($n = 10$), Mahogany ($n = 20$), Stardust ($n = 7$) and Black (Black-NS, $n = 16$). All animals from Millbank Fur Farm ($n = 15$) were of Black color-type (Black-ON). The relationship between animals was checked based on the pedigree information

and only animals with less degree of kinship were selected for further analyses (the median = 0.015; 1st–3rd quantile of relatedness = 0.008–0.039).

2.2. Whole Genome Sequencing and Variant Calling

Genomic DNA was extracted from tongue tissue using DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) following the manufacture protocol. DNA samples were then sequenced (100 bp pair-end reads) using the BGISEQ-500 platform at Beijing Genomics Institute (BGI, Guangdong, China). After sequencing, SOAPnuke software [37] was used to filter out sequencing adapters and low-quality reads. The clean reads were mapped to the American mink reference genome [38] using the mem-algorithm of the Burrows Wheeler Aligner (BWA) software [39]. The aligned files were converted from sequence alignment map (SAM) to binary alignment map (BAM) format and sorted using SAMtools package version 1.11 [40]. Potential PCR duplications were removed using ‘MarkDuplicates’ of Picard [41]. The BAM files were then indexed by SAMtools. Finally, variant calling was performed with the ‘mpileup’ module of SAMtools and BCFtools. All variants were filtered using Variant Call Format (VCF) tools based on the following parameters: minor allele frequency (MAF) >0.05; maximum missing rate <0.1; minor quality >30; and only bi-allelic variants were kept.

2.3. Inbreeding Coefficients and Genetic Distances

The average MAF, heterozygosity levels and inbreeding rate based on excess of homozygosity (F_{HOM}) were estimated for each color-type using SNP1101 version 1.0 [42]. Runs of homozygosity (ROH) were detected for each individual using SNP1101 software. The following parameters were set to detect ROH: minimum window size of 20 SNPs, genotyping error rate of 0.01 and sliding window step size of one SNP. The minimum length of ROH was set to be 500 kb, 1 Mb and 2 Mb in different runs. The inbreeding coefficients based on ROH (F_{ROH}) were calculated using Equation (1) as follows [43]:

$$F_{\text{ROH}} = \frac{\sum_K(\text{Length}(\text{ROH}_K))}{L}, \quad (1)$$

where $\sum_K(\text{Length}(\text{ROH}_K))$ indicates the total length of ROHs above a threshold length and L is the total length of genomic regions in which SNPs could be called, which was 802 Mb in this study.

Pairwise genetic distances were computed according to Nei [44] and Weir and Cockerham [45] procedures using ‘StAMPP’ package of R [46]. In addition, analysis of molecular variance (AMOVA) was performed using StAMPP package to describe the genetic variation among samples at different hierarchical levels. The AMOVA was computed based on the Nei’s genetic distance matrix using the `stamppAmova()` function following the procedure described by Excoffier et al. [47].

2.4. Genetic Structure and Admixture Patterns

Discriminant analysis of principal components (DAPC) was used to infer the genetic structure of the studied color-types. The DAPC was performed based on the procedure implemented in the ‘adegenet’ package of R [48]. Adegnet implements a prior principal component analysis (PCA) to transform the data and subsequently performs the discriminant analysis. The pre-defined number of clusters was determined using k-means algorithm, which found a given number (K) of clusters maximizing the variation between the clusters. The number of clusters between one and ten was explored and the Bayesian Information Criterion (BIC) was used to choose the optimum number of clusters following the elbow method. The number of retained principal components (PCs) was selected by the a -score [48,49].

Population structure and potential admixture among the populations (color-types) was further assessed with an unsupervised analysis in ADMIXTURE version 1.3.0 [50]. The ADMIXTURE was run for $K = 1$ to $K = 6$ with ten iterations per K . A cross-validation

procedure was applied to infer the best number of ancestral populations (K). The K value with the lowest standard error of cross-validation error was assumed as the best indicator of admixture pattern among populations.

2.5. Linkage Disequilibrium (LD) and Effective Population Size

The LD analysis and N_e estimation were performed based on the data of 100,000 SNPs randomly selected across scaffolds >10 Mb. The r^2 statistic was computed to measure the LD level between all possible pairs of SNPs located on the same scaffold using SNP1101 software. Equation (2) was used to estimate the r^2 values [51]:

$$r^2 = \frac{D^2}{(P_A P_B P_a P_b)}, \quad (2)$$

where: $D = P_{AB} - P_A P_B$ and P_A , P_a , P_B , and P_b are the frequencies of alleles A, a, B, and b, respectively; and P_{AB} presents the frequency of haplotype AB. Since haplotypes were not reconstructed, an unbiased estimator of D was computed using Equation (3) [52]:

$$D = \frac{N}{N-1} \left[\frac{4N_{AABB} + 2(N_{AABb} + N_{AaBB}) + N_{AaBb} - P_A P_B}{2N} \right], \quad (3)$$

where: N is the total number of samples and N_{AABB} , N_{AABb} , N_{AaBB} , and N_{AaBb} are the number of genotypes AABB, AABb, AaBb and AaBb, respectively. In addition, sample size correction for unphased data was made to all computed r^2 using Equation (4) as [53,54]:

$$r_{corrected}^2 = \frac{r_{computed}^2 - \frac{1}{n}}{1 - \frac{1}{n}}, \quad (4)$$

where n is the number of sampled haplotypes. The LD decay was assessed in three distance sets of ≤ 100 kb, ≤ 1000 kb, and ≤ 10 Mb, and SNP pairs were binned in each distance set using bin sizes of 10 kb, 100 kb, and 1000 kb, respectively. The average r^2 of each bin was plotted against the median size of the bin to depict the LD decay across the genome distances.

In addition, Equation (5) was used to estimate the effective population size based on the expectations for r^2 over different genomic intervals [55]:

$$r^2 = \frac{D^2}{(P_A P_B P_a P_b)}, \quad (5)$$

where: c shows the median of the recombination distance between SNPs in Morgans. The age of N_e at a given distance was estimated by $1/(2c)$ as suggested by Hayes et al. [56]. The recombination rate was assumed to be 1 centiMorgan per million base pairs. The N_e was assessed at 200, 150, 100, 50, 10, and 5 generations ago to monitor the changes in population size.

3. Results

3.1. Data Quality and Population Genetic Parameters

In total, more than 100 billion reads were generated with a range of 817,199,893 to 1,050,757,227 reads per sample. The genome coverage was in the range of $34\times$ to $44\times$ with an average of $40\times$ per sample. On average, 98.24% of reads were mapped to mink reference genome, with the range of 97.93 to 98.55%. In total, 22,990,329 variants were called by SAMtools and BCFtools and after quality control, 8,150,569 bi-allelic variants from 100 individuals were remained for further analysis.

Table 1 presents the average MAF, observed heterozygosity, F_{HOM} , and pedigree based inbreeding rate (F_{PED}) in different color-types. The average MAF ranged from 0.190 (Stardust) to 0.214 (Demi). The lowest level of heterozygosity (28.03%) was observed in Black-ON whereas the Black-NS had the highest percentage of heterozygosity (31.32%). The F_{HOM} ranged from -0.189 in Demi to -0.112 in Black-ON. The highest average of F_{PED}

was found in Pastel (0.026) and Demi had the lowest value of F_{PED} (0.006). The correlation between F_{HOM} and F_{PED} was estimated to be 0.59.

Table 1. Average minor allele frequency (MAF), observed heterozygosity, inbreeding rates based on excess of homozygosity (F_{HOM}) and pedigree (F_{PED}) for five color-types of American mink using whole genome sequence data.

Color-Types	Number of Animals	Average MAF	Observed Heterozygosity (%)	F_{HOM}	F_{PED}
Demi	32	0.214	31.12	−0.189	0.006
Pastel	10	0.195	29.42	−0.146	0.026
Mahogany	20	0.211	30.93	−0.184	0.009
Stardust	7	0.190	30.64	−0.153	0.016
Black-NS ¹	16	0.197	31.32	−0.166	0.011
CCFAR ²	85	0.216	30.87	−0.171	0.010
Black-ON ³	15	0.192	28.03	−0.112	NA ⁴
Total	100	0.216	30.45	−0.166	-

¹ Black color-type from Canadian Center for Fur Animal Research (CCFAR) at Dalhousie Faculty of Agriculture (Truro, NS, Canada); ² All samples collected at Canadian Center for Fur Animal Research (CCFAR); ³ Black color-type from Millbank Fur Farm (Rockwood, ON, Canada); ⁴ Not available.

Table 2 shows the average (\pm SD) F_{ROH} and the number of ROH segments detected in different studied groups based on the minimum ROH lengths of 500 kb, 1 Mb and 2 Mb. The highest average number of homozygous segments (120 ± 21) and F_{ROH} (0.138 ± 0.031) based on the minimum length of 500 kb were observed in Black-ON. The lowest average of $F_{ROH} > 500$ kb (0.072 ± 0.022) was found in Demi with an average (\pm SD) of 61 ± 17 segments per individual. Similar trends were also observed for higher thresholds of ROH length (Table 2). The short segments (500 kb to 1 Mb) included the most proportion of ROH detected among samples (69%). The highest number of identified ROH (length > 500 kb) per individual was 180, which was detected in one Black-ON mink. The correlations of F_{ROH} (length > 500 kb) with F_{HOM} and F_{PED} were 0.85 and 0.47, respectively.

Table 2. The average (\pm SD) number of runs of homozygosity (ROH) and inbreeding rates based on ROH (F_{ROH}) in different color-types of American mink.

Color-Types	Minimum Length of ROH					
	500 kb		1 Mb		2 Mb	
	Number	F_{ROH}	Number	F_{ROH}	Number	F_{ROH}
Demi	61 ± 17	0.072 ± 0.022	19 ± 7	0.036 ± 0.013	3 ± 2	0.009 ± 0.006
Pastel	90 ± 21	0.110 ± 0.027	30 ± 8	0.057 ± 0.016	5 ± 2	0.016 ± 0.008
Mahogany	68 ± 24	0.078 ± 0.030	21 ± 10	0.039 ± 0.018	3 ± 1	0.009 ± 0.005
Stardust	85 ± 12	0.100 ± 0.014	29 ± 4	0.056 ± 0.009	3 ± 2	0.011 ± 0.007
Black-NS ¹	82 ± 27	0.097 ± 0.034	26 ± 11	0.048 ± 0.021	3 ± 2	0.009 ± 0.007
CCFAR ²	72 ± 23	0.085 ± 0.029	23 ± 10	0.043 ± 0.018	3 ± 2	0.010 ± 0.006
Black-ON ³	120 ± 21	0.138 ± 0.031	39 ± 11	0.071 ± 0.019	4 ± 2	0.012 ± 0.007

¹ Black color-type from Canadian Center for Fur Animal Research (CCFAR) at Dalhousie Faculty of Agriculture (Truro, NS, Canada); ² All samples collected at Canadian Center for Fur Animal Research (CCFAR); ³ Black color-type from Millbank Fur Farm (Rockwood, ON, Canada).

3.2. Genetic Distances

Table 3 presents the Weir and Cockerham's F_{st} and Nei's genetic distance among different color-types. The highest F_{st} (0.124) and Nei's genetic distance (0.065) were found between Pastel and Stardust color-types. Both measures indicated that Demi and Mahogany color-types had the minimum genetic distances.

Table 3. Estimation of Nei's genetic distance (upper diagonal) and Weir and Cockerham's F_{st} (lower diagonal) between different color-types of American mink.

Color-Types	Demi	Pastel	Mahogany	Stardust	Black-NS	Black-ON
Demi	0	0.025	0.013	0.036	0.029	0.030
Pastel	0.036	0	0.036	0.065	0.055	0.058
Mahogany	0.015	0.062	0	0.034	0.021	0.028
Stardust	0.055	0.124	0.046	0	0.039	0.044
Black-NS ¹	0.057	0.114	0.033	0.066	0	0.037
Black-ON ²	0.059	0.122	0.048	0.077	0.076	0

¹ Black color-type from Canadian Center for Fur Animal Research (CCFAR) at Dalhousie Faculty of Agriculture (Truro, NS, Canada); ² Black color-type from Millbank Fur Farm (Rockwood, ON, Canada).

Table 4 shows the AMOVA for the studied color-types. The higher proportion of variation was explained by the differentiation within color-types (73%). The variation among color-types was also significant ($p < 0.001$) and represented 18% of the total molecular variation in the populations. On the other hand, the variation among farms was not significant and accounted for 9% of the total variation.

Table 4. Analysis of molecular variance (AMOVA) in different color-types of farmed mink.

Source of Variation	Degree of Freedom	Sum of Squares	Mean Squared Deviations	Variance Components	Percentage of Variation
Among farms	1	0.123	0.123	0.002	9
Among color-types	4	0.301	0.075	0.004 ***	18
Within color-types	94	1.560	0.016	0.016 ***	73
Total	99	1.984	0.020	0.022	100

*** $p < 0.001$.

3.3. Genetic Structure and Admixture Patterns

The lowest value of BIC was observed for $K = 3$ indicating the most likely number of genetic clusters in the data set. The DAPC was performed to reveal the population structure assuming three main clusters. Figure 1a presents the scatter plot of the first two linear discriminants (LD) for all samples. Most of Black-NS individuals were placed in a separate cluster. Whereas Mahogany and Demi shared the overlapping clusters, Pastel, Stardust and Black-ON were assigned to certain clusters. In addition, the membership probabilities of each individual to different clusters were presented in Figure 1b. The Black-ON and Stardust shared the same cluster and most Pastel and Black-NS were assigned to other two unique clusters. However, Demi and Mahogany were largely admixed from other clusters.

We also used the admixture analysis to reveal the admixture patterns assuming two to six ancestral populations (Figure 2). The lowest cross-validation error (0.49) was found for $K = 3$, indicating the best fitting model of admixture among the studied individuals (Figure 2b). Ancestral proportions at $K = 3$ indicated that Black-ON individuals were mainly assigned to the same cluster (on average, 92.36%). In addition, Pastel (on average, 78.17%) and Black-NS (on average, 62.94%) constructed two other main genetic compositions whereas other color-types (Demi, Mahogany, and Stardust) were admixed of these three clusters (Figure 2).

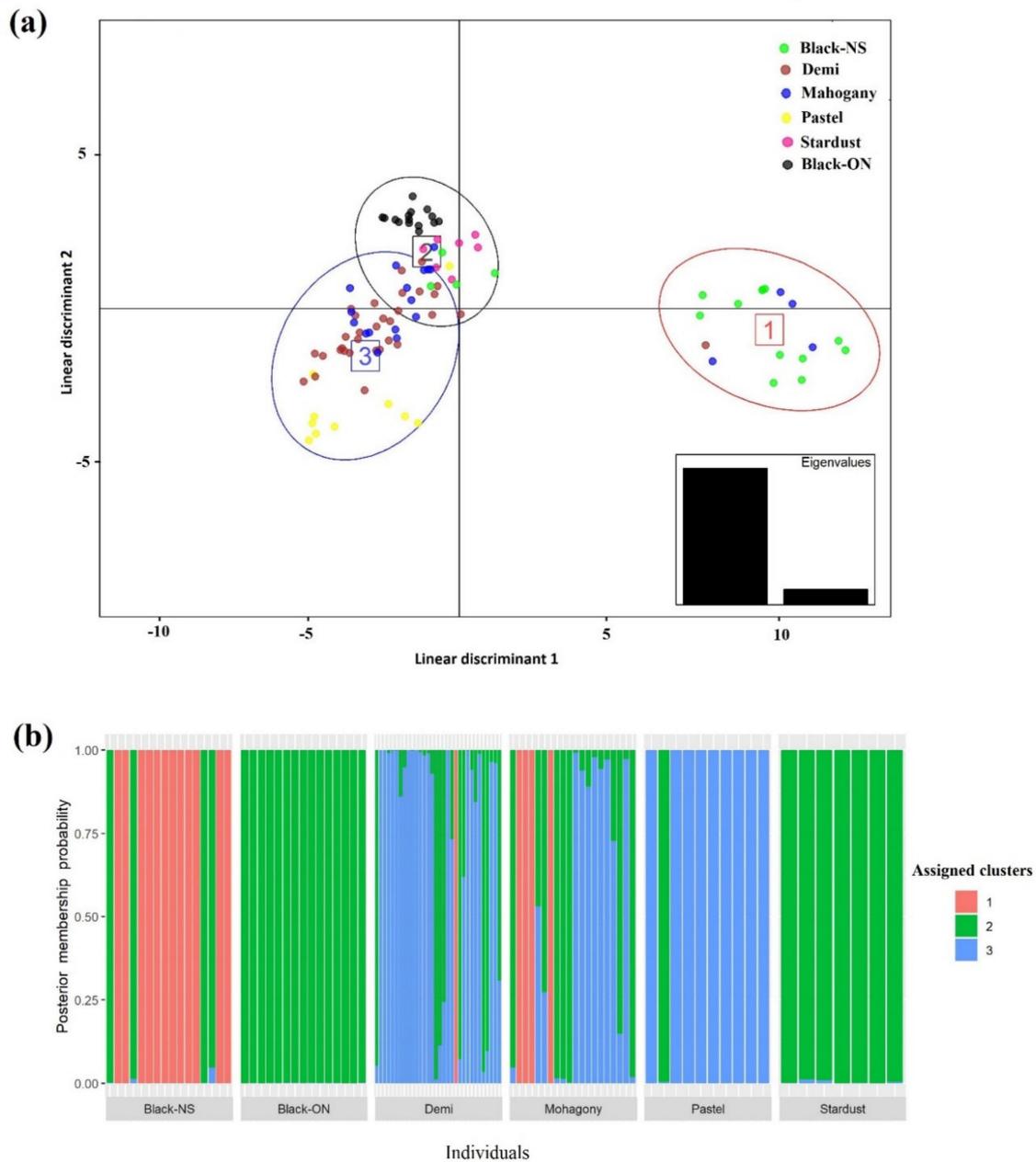


Figure 1. Discriminant analysis of principal components (DAPC) of five color-types of American mink: **(a)** The scatter plot of the first two linear discriminants for all samples. Color-types are presented by different colors and the three main clusters are shown by circles and numbers. The first two linear discriminants explained 89.68 and 10.32% of the variation, respectively; **(b)** The probability of membership of each sample in the three assigned clusters inferred by DAPC. Black-NS and Black-ON represent the black color-type samples collected from the Canadian Center for Fur Animal Research (Truro, NS, Canada) and Millbank Fur Farm (Rockwood, ON, Canada), respectively.

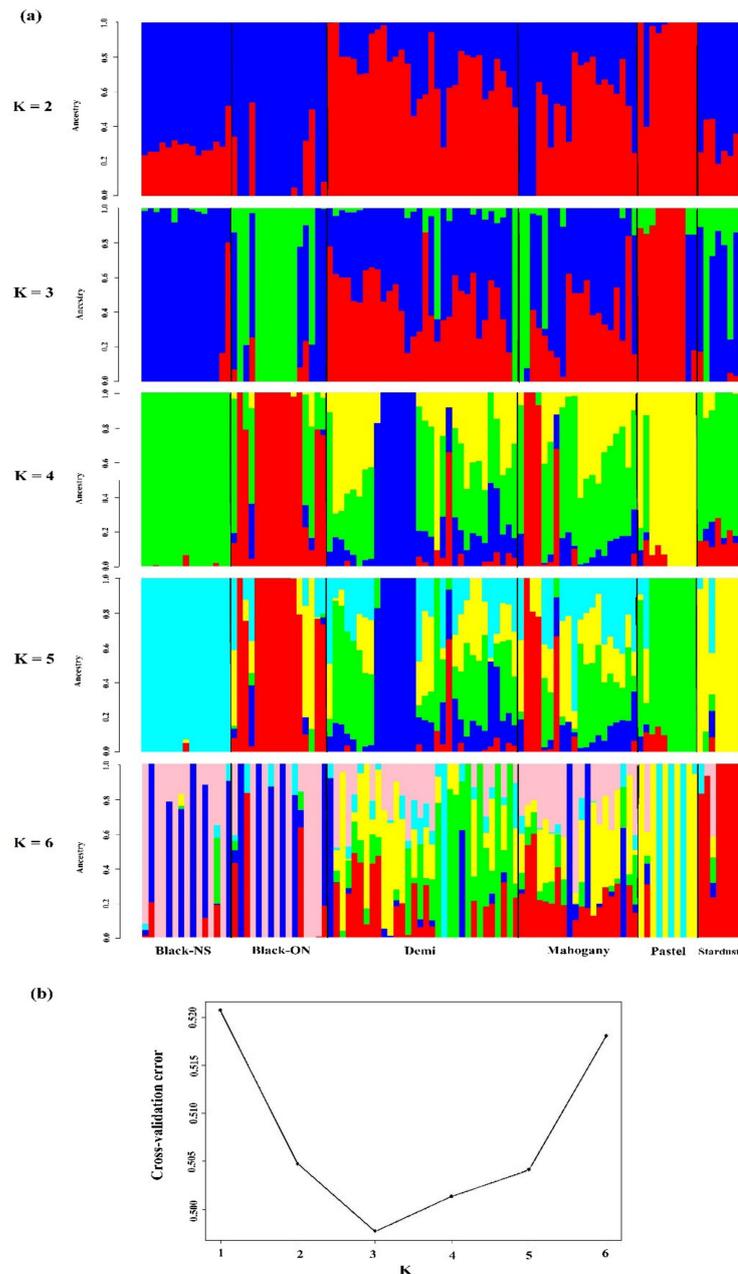


Figure 2. Admixture patterns of American mink color-types: (a) Admixture analysis at $K = 2$ to $K = 6$ in five color-types of American mink. Each individual was represented by a vertical bar and different colors were used to present color-types; (b) Cross-validation error as a function of K for admixture analysis of five color-types. The lowest cross-validation error was found at $K = 3$. Black-NS and Black-ON represent the black color-type samples collected from the Canadian Center for Fur Animal Research (Truro, NS, Canada) and Millbank Fur Farm (Rockwood, ON, Canada), respectively.

3.4. Linkage Disequilibrium and Effective Population Size

Table 5 shows the average $r^2 \pm SD$ between adjacent SNPs for different color-types. In addition, the average $r^2 \pm SD$ between adjacent SNPs over all studied scaffolds was presented in Table S1. The average distance between adjacent markers was 8.17 ± 2.84 kb. The highest and lowest average $r^2 \pm SD$ between adjacent markers were observed in Black-ON (0.366 ± 0.388) and Mahogany (0.280 ± 0.361) color-types, respectively. We also computed the pairwise r^2 between markers to describe the LD decay in different genomic distances. Figure 3 presents the LD decay up to 100 kb, 1000 kb, and 10 Mb using 10 kb,

100 kb, and 1 Mb bins, respectively. In addition, Table S2 presents the extension of LD (average $r^2 \pm SD$) at different intervals up to 1000 kb. The average $r^2 > 0.2$ extended to intervals up to 20 kb in Demi and Mahogany. On the other hand, higher levels of LD (average $r^2 > 0.2$) were observed for intervals up to 80 kb, 90 kb, and 100 kb in Black-NS, Pastel and Black-ON color-types, respectively. More rapid reduction was observed in LD decay of Black-ON over 200 kb genomic intervals (Figure 3).

Table 5. The average $r^2 \pm SD$ between adjacent SNPs in five color-types of American mink.

Number of Scaffolds	Length (Mb)	Number of SNPs	Average $r^2 \pm SD$					
			Demi	Mahogany	Pastel	Black-NS ¹	CCFAR ²	Black-ON ³
51	802.19	100,000	0.285 \pm 0.346	0.280 \pm 0.361	0.344 \pm 0.411	0.343 \pm 0.377	0.266 \pm 0.334	0.366 \pm 0.388

¹ Black color-type from the Canadian Center for Fur Animal Research (CCFAR) at Dalhousie Faculty of Agriculture (Truro, NS, Canada); ² All samples collected at the Canadian Center for Fur Animal Research (CCFAR); ³ Black color-type from Millbank Fur Farm (Rockwood, ON, Canada).

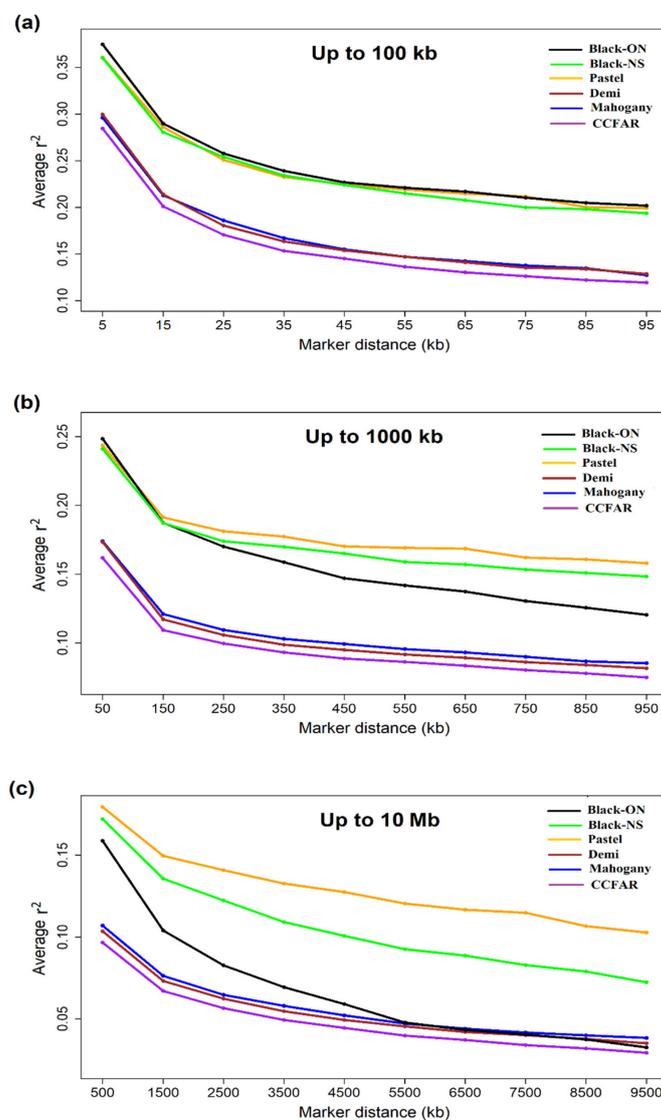


Figure 3. Linkage disequilibrium (LD) decays represented by the average r^2 for inter-marker distances of (a) up to 100 kb using consecutive 10 kb bins (b) up to 1000 kb using consecutive 100 kb bins and (c) up to 10 Mb using consecutive 1000 kb bins. Black-NS and Black-ON represent the black color-type samples collected from the Canadian Center for Fur Animal Research (Truro, NS, Canada) and Millbank Fur Farm (Rockwood, ON, Canada), respectively.

The N_e was also estimated based on the average LD obtained in different genomic intervals. Figure 4 represents the changes of estimated N_e over the last 200 generations. The recent N_e (five generations ago) was estimated to be 19, 47, 66, 28, and 50 in Pastel, Mahogany, Demi, Black-NS, and Black-ON, respectively. The recent N_e was 99 for all individuals from CCFAR farm. Furthermore, the maximum N_e was observed for Demi (846) at 200 generations ago whereas the lowest population size was for Pastel (414) at the same time.

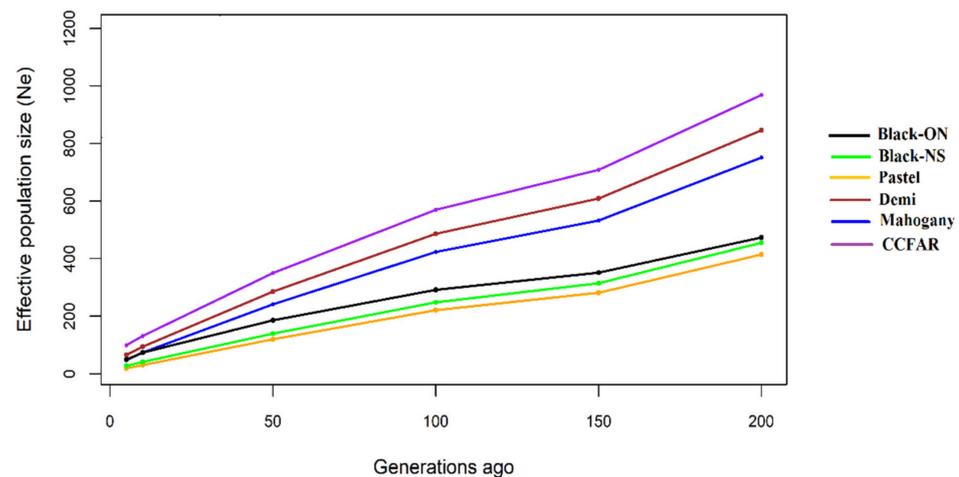


Figure 4. Estimated effective population sizes (N_e) for five color-types of American mink as a function of generations ago. Black-NS and Black-ON represent the black color-type samples collected from the Canadian Center for Fur Animal Research (Truro, NS, Canada) and Millbank Fur Farm (Rockwood, ON, Canada), respectively.

4. Discussion

This study used the whole genome sequence data to assess the genetic structure of 100 American mink from two Canadian mink farms. The results of this study confirmed that there is no wide genetic differentiation among studied color-types owing to high levels of gene flow. However, these findings were critical to determine the markers density required for genomic studies in the mink farms.

High percentage of reads mapped to the reference genome (on average, 98.24%) confirming the high efficiency of the whole genome sequence data in this study. The number of SNPs obtained in this study (8,150,569) was much higher than 52,714 and 34,816 SNPs that were extracted by GBS technique to study the LD pattern [35] and the genes associated with body size [57] in American mink, respectively. This is due to the fact that GBS is a restriction enzyme-based technique that can only sequence a reduced subset of genome [58]. These results are promising to design the SNP panel using whole genome data and subsequently develop the genome-enabled selection in American mink.

The F_{HOM} ranged between -0.189 and -0.112 in different color-types (Table 1), that were less than the average value of 0.150 obtained from 52,714 SNPs in black American mink [35]. This discrepancy can be due to the differences in the historical background, sampling approaches and the density of markers applied in these studies. In addition, the higher genomic inbreeding rate of 0.271 was reported using microsatellite markers for farm and wild American mink in Eastern Canada [59], which was attributed to line-breeding in the studied farms, and breeding between related individuals, as well as a limited movement of wild mink. On the other hand, inbreeding coefficients were in the range of -0.150 to 0.005 using a panel of 194 SNPs for farm and feral American mink in Denmark and Poland, which were in agreement with those (-0.189 to -0.112) observed in the current study. The lowest level of heterozygosity (28.03%) was observed for Black-ON, which was in agreement with the higher degree of inbreeding (-0.112) estimated for this

group. The highest level of inbreeding based on pedigree was found for Pastel (0.026) that was in accordance with the low N_e (19) estimated for this color-type. The high correlation of 0.59 was observed between F_{PED} and F_{HOM} across different color-types. This value is comparable with the correlations of 0.58, 0.48, and 0.51 reported in three cattle breeds [60], seven pig breeds [61] and Spanish Holstein population [62], respectively.

The inbreeding rates were also investigated using ROH analysis. The highest average F_{ROH} was observed for Black-ON across all the defined lengths, which was in accordance with the higher F_{HOM} (−0.112), lower heterozygosity (28.03%) and N_e (50) found in this group. The average $F_{ROH} > 500$ kb ranged from 0.072 to 0.138 among different groups in this study, which were lower than 0.186 reported for black American mink using 13,321 SNPs [35]. These differences might be attributed to the genetic history, sample size, coverage of genome and density of markers existed in these studies. The length of homozygous segments can provide evidence to infer the recent and ancient inbreeding events where the short autozygous tracts are related with ancient inbreeding and longer segments correlate with recent inbreeding events [63]. It was approximated that the ROHs >1 Mb date back about 50 generations ago [64]. The abundance of shorter ROH (500 kb to 1 Mb) indicated that the inbreeding events likely occurred over 50 generations ago in the studied populations, which falls into the approximated time of reduction in population sizes due to domestication [20]. Since there is no chromosome-scale genome assembly available for American mink [38], the analyses of this study were restricted to longer scaffolds (>10 Mb). Accordingly, it was not possible to assess the recent inbreeding events across the populations based on the detection of longer ROH. The availability of chromosome-scale genome assembly for American mink can open more opportunities to reveal the genetic diversity and develop the genomic studies of this species. High correlation of 0.85 was found between F_{HOM} and F_{ROH} , which was comparable with the values of 0.87, 0.89, and 0.83–0.95 reported in black American mink [35], four cattle breeds [65] and Italian local cattle [66], respectively. However, a lower correlation (0.47) was observed between F_{ROH} and F_{PED} . These results implied that genomic inbreeding coefficients can provide more accurate measure to control assortative mating on mink farms.

Although the AMOVA indicated that the genetic differentiation among color-types was significant, this explained a low proportion of total variance (18%) was explained by that. The multivariate DAPC could also differentiate color-types in certain groups; however, clusters were overlapped and the distances between groups were not considerable, which might be due to high gene flow and the admixture events occurred in recent generations among studied mink. The Black-ON samples were closely clustered in a separate group whereas the Black-NS individuals were mainly grouped in another cluster, showing the differences between two black color-types in this study (Figure 1). These results revealed that geographical distribution may play more important role in genetic differentiation of mink compared to color-types. It seems that extensive admixture events have shaped the genetic structure among mink color-types in CCFAR farm (Figures 1b and 2), which was expected due to raising different color-types on the farm without systematic mating. However, Pastel and Black-NS were recognized as two main genetic compositions constructing the genetic structure of the studied individuals from CCFAR. Assuming three genetic clusters, Black-ON was the phenotypic category with the least evidence of admixture as all individuals were inferred to be largely from one ancestry component (Figures 1b and 2a). The low genetic distances indicated that there was no extensive discrimination among the studied color-types (Table 3). Although the color-types existing in the CCFAR were not widely differentiated, the intensity of admixture and inbreeding rates were not similar across the color-types due to the differences in the group sizes, historical background and unmanaged mating in the past generations. In accordance with the result of this study, Thirstrup et al. [31] observed low genetic differentiation (F_{st}) among mink color-types (0.076) and mink from two different geographical origins (0.087). However, pairwise F_{st} ranged between 0.017 and 0.364 among mink from 12 locations in southern Chile [32], which were higher than those observed in this study (0.013 to 0.065). Besides the genetic

origins and sampling methods, the differences in the marker types and density can be possible reasons for the higher F_{st} values in the Chilean mink populations.

The estimations of N_e at five generations ago were 99 and 50 for CCFAR and Millbank Fur Farm, respectively. These estimations were less than the N_e of 116 observed in black American mink using SNP data extracted by GBS technique [35]. This difference might be related to the history of populations, sample sizes, and density of markers used in these studies. In addition, the range of 17.5 to 78.8 [34] and 7.2 to 34.8 [33] were estimated using microsatellite markers for mink population size in Swedish coasts and Spain, respectively. The low N_e estimated in these studies were attributed to the conservation programs conducted in these areas to control the mink population size. Furthermore, the discrepancy of estimations might be due to the differences in the density and type of DNA markers, and genetic background of populations. Despite the different estimations for N_e among the studied color-types, the similar trends were observed for changes in N_e from 200 to 5 generations ago (Figure 4). Our results indicated that the decline of population sizes was more rapid between 200 and 150 generations ago, which was in accordance with the domestication events initiated to produce fur in 1866 [20]. These trends were in agreement with those reported for black American mink using GBS data [35]. It is suggested to use the high-throughput genomic data to control the reducing population size and levels of inbreeding in mink farms.

The average r^2 between adjacent markers were in range of 0.280 to 0.366 for different color-types, which was in agreement with the value of 0.300 ± 0.350 reported by Karimi et al. [35] for American mink. Genomic selection and genome wide association studies (GWAS) rely on the assumption of the presence of LD between markers and the genes controlling the traits of interest [67]. The r^2 of > 0.2 was suggested as the required LD level between casual gene and marker to achieve an accuracy of > 0.85 in genomic selection programs [68]. Similarly, $r^2 > 0.3$ can be accounted as the threshold LD to obtain adequate power and precision in GWAS [7]. In this study, the LD patterns revealed that the average r^2 reduced to < 0.2 at inter-marker intervals of > 20 kb and > 100 kb in CCFAR and Millbank Fur Farm, indicating that the density of 120,000 (2.4 Gb/20 kb, where 2.4 Gb is the size of genome assembly) and 24,000 (2.4 Gb/100 kb) SNPs would provide the acceptable accuracy for genomic selection in these populations, respectively. Whereas the average $r^2 > 0.2$ was extended to marker distances of < 20 kb for both Demi and Mahogany, the higher extensions of 80 and 90 kb were observed for Black-NS and Pastel, respectively. In addition, the average $r^2 > 0.3$ at genomic distances of < 10 kb for both CCFAR and Millbank Fur Farm suggested that 240,000 informative SNPs can capture the LD levels necessary for GWAS in these populations. Based on the GBS data, Karimi et al. [35] suggested the density of 60,000 SNPs for genomic selection in black American mink, which is higher than the 24,000 suggested for Millbank Fur Farm, but is less than the 120,000 for CCFAR in this study. In addition, the 120,000 SNPs suggested by Karimi et al. [35] for GWAS is less than the density of 240,000 required for both farms in the current study.

Our results indicated that higher marker densities are required for admixed population (CCFAR) compared to non-mixed population (i.e., Millbank Fur Farm). Similar to our results, lower LD was observed for a group of admixed Danish Jersey cattle in comparison with the original Danish and United States Jersey dairy cattle whereas the LD persistence was lower at longer marker distances for the original groups [69]. In addition, lower LD levels were reported at short distances for a heterogeneous breed (Rangeland) in Australian and Canadian goat breeds, suggesting that higher dense panel is required for genomic selection in this breed [70]. Admixture can generate the LD between both linked and unlinked loci, particularly those with different allele frequencies in the founder populations [11]. Although admixture-induced LD (ALD) between unlinked loci disappears quickly owing to recombination events, ALD between linked loci will persist for several generations. Despite the increasing levels of long-range ALD, the short-range LD, which is originated from the founder populations, is simultaneously reduced due to the admixture events [69,71]. Although sparse markers can effectively capture the long-range ALD, only a limited proportion of genetic

variation is likely to be explained, restricting the accurate genomic predictions. The efficiency of genomic selection in the admixed population relies on the LD levels remained from founder populations, the long-range LD induced by admixture, and the degree of genetic differentiation between original populations [71]. The significance of these factors depends on the heritability, genetic architecture of the traits, density of markers, and the genomic selection approaches applied in the data [71,72]. In the mink farms, the traits of interest are usually measured in mixed populations with no certain genetic composition. This implies that not accounting for admixture pattern can bias the estimates of marker effects for genomic selection [72]. Our results confirmed that admixed population would require a higher density of SNPs for genomic selection and GWAS.

The results revealed that three main genetic compositions recognized by admixture analysis (i.e., Pastel, Black-NS and Black-ON) presented the higher levels of LD whereas the admixed color-types (Mahogany and Demi) had lower LD levels (Figure 3). The trend of LD decay and N_e suggested that Mahogany and Demi were under the similar evolutionary forces, which is also in accordance with their close genetic distance. High LD levels of Pastel and Black-NS might be due to the small number of founders initially brought to the CCFAR farm. In addition, the higher genetic differentiation of Pastel and Black-NS could lead to higher ALD for these color-types. The extension of $r^2 > 0.2$ to genomic distances of 100 kb for Black-ON might be indicative of strong founder effects in this population. Generally, observation of long-range LD could be due to population admixture [73,74], random genetic drift [75], structural variations of genome [76,77], epistatic selection [78], hitchhiking effect of positively selected mutations [79] and demographic changes [80]. For instance, recurrent bottlenecks were suggested as the contributing force to long-range LD observed in the European human populations [81,82]. Furthermore, false appearance of LD could originate from some types of errors in the data or reference genome [82]. In the current study, the pattern of N_e as well as the abundance of shorter ROH (500 kb to 1 Mb) indicated that the domestication of Black-ON mink was likely initiated with a smaller number of founders or earlier than other groups. There is evidence that the LD curve is high and flat after the end of the population collapse and as population size increases at the end of the bottleneck, the right portion of the LD curve descends more quickly than the left portion [83], which is in accordance with the LD pattern observed for Black-ON in our study. Although it seems that association mapping may require the lower density of SNP markers in Black-ON population, the likelihood of false positive associations between markers and causal genes would increase due to the high level of LD. In other words, the hitchhiking effects of high LD haplotype-blocks could lead to an excess of candidate SNPs and less precise mapping of causative variants [83].

Results of this study confirmed no wide genetic distinction among the studied mink indicating that the samples might not be ideal to reveal the population structure. For instance, the existence of wild groups and samples from a wider geographical distribution could enrich the analysis of population history. In addition, the small sample size can result in biased and less accurate estimation of LD measures and the bias might vary with the distances between markers [84]. Although the r^2 values in the current study were corrected for the sampling error, they might be subjected to bias and any comparison should be conservative. In fact, the LD estimations may vary depending on the sample size, MAF threshold, marker density and inter-marker distances applied in different studies. Despite all the restrictions, the main idea of this study was to provide the preliminary results for development of genomic evaluations in the mink industry. Given the fact that the mink industry has faced with economic issues in recent years, especially with the effect of SARS-Cov-2 on mink farms in Europe and North America, this industry will be required to increase the production efficiency. Genomic selection is suggested as an appropriate strategy to improve the economically important traits such as SARS-Cov-2 disease resistance and fur quality in mink [29]. The results of this study could provide the essential information to design the SNP panel for American mink and subsequently apply modern genomic strategies in the mink industry.

5. Conclusions

In this study, whole genome sequences were used to reveal the genetic structure and LD patterns of 100 American mink. The genetic analysis confirmed that there was no wide genetic differentiation among the studied individuals. However, three ancestral genetic compositions were recognized among the mink from two studied farms. Our analyses showed that the CCFAR is an admixed population whereas animals from Millbank Fur Farm were nearly pure. We observed a relationship between the admixture levels and LD extensions across the studied color-types where the three less admixed color-types (i.e., Pastel, Black-NS, and Black-ON) presented higher levels of LD and the highly admixed color-types (Mahogany and Demi) had lower LD levels. The extension of LD (r^2) to genomic distances of 20 kb and 100 kb for CCFAR and Millbank Fur Farm suggested that 120,000 and 24,000 SNPs would require for applying genomic selection programs in these farms, respectively. These results indicated that accounting for admixture is critical to design the SNP panels and achieve the required accuracy in genomic studies of American mink.

Supplementary Materials: The following are available online at <https://www.mdpi.com/2073-4425/12/2/258/s1>. Table S1: The average $r^2 \pm$ SD between adjacent SNPs over different scaffolds in five color-types of American mink. Table S2: The average $r^2 \pm$ SD over physical distances up to 1000 kb, pooled over all scaffolds, in five color-types of American mink.

Author Contributions: Conceptualization, K.K. and Y.M.; data analyses, K.K. and D.N.D.; software, M.S.; writing—original draft preparation, K.K.; writing—review and editing, Y.M. and D.N.D.; project administration and funding acquisition, Y.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada, Canada Mink Breeders Association, Nova Scotia Mink Breeders Association, and Mink Veterinary Consulting and Research Service Ltd.

Institutional Review Board Statement: The study was conducted according to the Code of Practice for the Care and Handling of Farmed Mink guidelines, and approved by the Dalhousie University Animal Care and Use Committee (certification number: 2018-009).

Informed Consent Statement: Not applicable.

Data Availability: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Acknowledgments: We greatly acknowledge Jingyi Wang for technical assistance and laboratory works. We would like to extend thanks to CCFAR (Dalhousie Agricultural Campus) and Millbank Fur Farm staff for collecting and providing the data.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Wellmann, R.; Bennewitz, J. Key Genetic Parameters for Population Management. *Front. Genet.* **2019**, *10*, 667. [[CrossRef](#)]
2. Karimi, K.; Strucken, E.M.; Moghaddar, N.; Ferdosi, M.H.; Esmailizadeh, A.; Gondro, C. Local and Global Patterns of Admixture and Population Structure in Iranian Native Cattle. *BMC Genet.* **2016**, *17*, 108. [[CrossRef](#)]
3. The Bovine HapMap Consortium; Gibbs, A.R.; Taylor, J.F.; Van Tassell, C.; Barendse, W.; Eversole, A.K.; Gill, A.C.; Green, R.D.; Hamernik, D.L.; Kappes, S.M.; et al. Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds. *Science* **2009**, *324*, 528–532. [[CrossRef](#)]
4. Kijas, J.W.; Lenstra, J.A.; Hayes, B.; Boitard, S.; Neto, L.R.P.; Cristobal, M.S.; Servin, B.; McCulloch, R.; Whan, V.; Gietzen, K.J.; et al. Genome-Wide Analysis of the World's Sheep Breeds Reveals High Levels of Historic Mixture and Strong Recent Selection. *PLoS Biol.* **2012**, *10*, e1001258. [[CrossRef](#)]
5. Colli, L.; The ADAPTMAP Consortium; Milanese, M.; Talenti, A.; Bertolini, F.; Chen, M.; Crisà, A.; Daly, K.G.; Del Corvo, M.; Guldbbrandtsen, B.; et al. Genome-Wide SNP Profiling of Worldwide Goat Populations Reveals Strong Partitioning of Diversity and Highlights Post-Domestication Migration Routes. *Genet. Sel. Evol.* **2018**, *50*, 58. [[CrossRef](#)]

6. Edea, Z.; Kim, S.-W.; Lee, K.-T.; Kim, T.H.; Kim, K.-S. Genetic Structure of and Evidence for Admixture between Western and Korean Native Pig Breeds Revealed by Single Nucleotide Polymorphisms. *Asian-Australas. J. Anim. Sci.* **2014**, *27*, 1263–1269. [[CrossRef](#)]
7. Ardlie, K.G.; Kruglyak, L.; Seielstad, M. Patterns of Linkage Disequilibrium in the Human Genome. *Nat. Rev. Genet.* **2002**, *3*, 299–309. [[CrossRef](#)] [[PubMed](#)]
8. Slatkin, M. Linkage Disequilibrium—Understanding the Evolutionary Past and Mapping the Medical Future. *Nat. Rev. Genet.* **2008**, *9*, 477–485. [[CrossRef](#)]
9. Reich, D.E.; Cargill, M.; Bolk, S.; Ireland, J.; Sabeti, P.C.; Richter, D.J.; Lavery, T.; Kouyoumjian, R.; Farhadian, S.F.; Ward, R.; et al. Linkage Disequilibrium in the Human Genome. *Nat. Cell Biol.* **2001**, *411*, 199–204. [[CrossRef](#)] [[PubMed](#)]
10. Qanbari, S. On the Extent of Linkage Disequilibrium in the Genome of Farm Animals. *Front. Genet.* **2020**, *10*, 1304. [[CrossRef](#)] [[PubMed](#)]
11. Pfaff, C.L.; Parra, E.J.; Bonilla, C.; Hiester, K.; McKeigue, P.M.; Kamboh, M.I.; Hutchinson, R.G.; Ferrell, R.E.; Boerwinkle, E.; Shriver, M.D. Population Structure in Admixed Populations: Effect of Admixture Dynamics on the Pattern of Linkage Disequilibrium. *Am. J. Hum. Genet.* **2001**, *68*, 198–207. [[CrossRef](#)]
12. Goddard, M.E.; Hayes, B.J. Mapping Genes for Complex Traits in Domestic Animals and Their Use in Breeding Programmes. *Nat. Rev. Genet.* **2009**, *10*, 381–391. [[CrossRef](#)]
13. Hollenbeck, C.M.; Portnoy, D.S.; Gold, J.R. A Method for Detecting Recent Changes in Contemporary Effective Population Size from Linkage Disequilibrium at Linked and Unlinked Loci. *Heredity* **2016**, *117*, 207–216. [[CrossRef](#)]
14. Hill, W.G. Estimation of Effective Population Size from Data on Linkage Disequilibrium. *Genet. Res.* **1981**, *38*, 209–216. [[CrossRef](#)]
15. Karimi, K.; Koshkoiyeh, A.E.; Asadi-Fozi, M.; Porto-Neto, L.R.; Gondro, C. Prioritization for Conservation of Iranian Native Cattle Breeds Based on Genome-Wide SNP Data. *Conserv. Genet.* **2015**, *17*, 77–89. [[CrossRef](#)]
16. Waples, R.K.; Larson, W.A. Estimating Contemporary Effective Population Size in Non-model Species Using Linkage Disequilibrium across Thousands of Loci. *Heredity* **2016**, *117*, 233–240. [[CrossRef](#)] [[PubMed](#)]
17. Theodorou, K.; Couvet, D. On the Expected Relationship between Inbreeding, Fitness, and Extinction. *Genet. Sel. Evol.* **2006**, *38*, 371–387. [[CrossRef](#)] [[PubMed](#)]
18. Howard, J.T.; Pryce, J.E.; Baes, C.; Maltecca, C. Invited Review: Inbreeding in the Genomics Era: Inbreeding, Inbreeding Depression, and Management of Genomic Variability. *J. Dairy Sci.* **2017**, *100*, 6009–6024. [[CrossRef](#)] [[PubMed](#)]
19. Abramov, A.V. A Taxonomic Review of the Genus *Mustela* (Mammalia Carnivora). *Zoosyst. Ross.* **2000**, *8*, 357–364.
20. Bowness, E.R. An Historical Perspective on the North American Mink Industry. In *Mink Biology, Health and Disease*; Hunter, D.B., Lemieux, N., Eds.; University of Guelph: Guelph, ON, Canada, 1996; pp. 1–9.
21. Joergensen, G. *Mink Production*; Scientifur: Hillerød, Denmark, 1985; p. 399.
22. Lagerkvist, G.; Johansson, K.; Lundeheim, N. Selection for Litter Size, Body Weight, and Pelt Quality in Mink (*Mustela vison*): Correlated Responses. *J. Anim. Sci.* **1994**, *72*, 1126–1137. [[CrossRef](#)]
23. Hu, G.; Do, D.N.; Gray, J.; Miar, Y. Selection for Favorable Health Traits: A Potential Approach to Cope with Diseases in Farm Animals. *Animals* **2020**, *10*, 1717. [[CrossRef](#)] [[PubMed](#)]
24. Do, D.N.; Miar, Y. Evaluation of Growth Curve Models for Body Weight in American Mink. *Animals* **2019**, *10*, 22. [[CrossRef](#)]
25. Karimi, K.; Sargolzaei, M.; Plastow, G.; Wang, Z.; Miar, Y. Genetic and Phenotypic Parameters for Litter Size, Survival Rate, Gestation Length, and Litter Weight Traits in American Mink1. *J. Anim. Sci.* **2018**, *96*, 2596–2606. [[CrossRef](#)]
26. Wiggans, G.; Cole, J.; Hubbard, S.M.; Sonstegard, T.S. Genomic Selection in Dairy Cattle: The USDA Experience. *Annu. Rev. Anim. Biosci.* **2017**, *5*, 309–327. [[CrossRef](#)] [[PubMed](#)]
27. Knol, E.F.; Nielsen, B.; Knap, P.W. Genomic Selection in Commercial Pig Breeding. *Anim. Front.* **2016**, *6*, 15–22. [[CrossRef](#)]
28. Davey, J.W.; Hohenlohe, P.A.; Etter, P.D.; Boone, J.Q.; Catchen, J.M.; Blaxter, M.L. Genome-Wide Genetic Marker Discovery and Genotyping Using Next-Generation Sequencing. *Nat. Rev. Genet.* **2011**, *12*, 499–510. [[CrossRef](#)]
29. Karimi, K.; Sargolzaei, M.; Plastow, G.S.; Wang, Z.; Miar, Y. Opportunities for Genomic Selection in American Mink: A Simulation Study. *PLoS ONE* **2019**, *14*, e0213873. [[CrossRef](#)]
30. Shimatani, Y.; Fukue, Y.; Kishimoto, R.; Masuda, R. Genetic Variation and Population Structure of the Feral American Mink (*Neovison vison*) in Nagano, Japan, Revealed by Microsatellite Analysis. *Mammal Study* **2010**, *35*, 1–7. [[CrossRef](#)]
31. Thirstrup, J.; Ruiz-González, A.; Pujolar, J.M.; Larsen, P.F.; Jensen, J.; Randi, E.; Zalewski, A.; Pertoldi, C. Population Genetic Structure in Farm and Feral American Mink (*Neovison vison*) Inferred from RAD Sequencing-Generated Single Nucleotide Polymorphisms1. *J. Anim. Sci.* **2015**, *93*, 3773–3782. [[CrossRef](#)]
32. Mora, M.; Medina-Vogel, G.; Sepúlveda, M.A.; Noll, D.; Álvarez-Varas, R.; Vianna, J.A. Genetic Structure of Introduced American Mink (*Neovison vison*) in Patagonia: Colonisation Insights and Implications for Control and Management Strategies. *Wildl. Res.* **2018**, *45*, 344. [[CrossRef](#)]
33. Lecis, R.; Ferrando, A.; Ruiz-Olmo, J.; Mañas, S.; Roura, D.X. Population Genetic Structure and Distribution of Introduced American Mink (*Mustela vison*) in Spain, Based on Microsatellite Variation. *Conserv. Genet.* **2007**, *9*, 1149–1161. [[CrossRef](#)]
34. Zalewski, A.; Zalewska, H.; Lunneryd, S.-G.; André, C.; Mikusiński, G. Reduced Genetic Diversity and Increased Structure in American Mink on the Swedish Coast following Invasive Species Control. *PLoS ONE* **2016**, *11*, e0157972. [[CrossRef](#)]
35. Karimi, K.; Farid, A.H.; Sargolzaei, M.; Myles, S.; Miar, Y. Linkage Disequilibrium, Effective Population Size and Genomic Inbreeding Rates in American Mink Using Genotyping-by-Sequencing Data. *Front. Genet.* **2020**, *11*, 223. [[CrossRef](#)]

36. Turner, P.; Buijs, S.; Rommers, J.M.; Tessier, M. *The Code of Practice for the Care and Handling of Farmed Mink*; The National Farm Animal Care Council: Rexdale, ON, Canada, 2013; p. 58.
37. Chen, Y.; Chen, Y.; Shi, C.; Huang, Z.; Zhang, Y.; Li, S.; Li, Y.; Ye, J.; Yuxin, C.; Li, Z.; et al. SOAPnuke: A Mapreduce Acceleration-Supported Software for Integrated Quality Control and Preprocessing of High-Throughput SE-Quencing Data. *GigaScience* **2018**, *7*, 1–6. [[CrossRef](#)] [[PubMed](#)]
38. Cai, Z.; Petersen, B.; Sahana, G.; Madsen, L.B.; Larsen, K.; Thomsen, B.; Bendixen, C.; Lund, M.S.; Guldbrandtsen, B.; Panitz, F. The First Draft Reference Genome of the American Mink (Neovison Vison). *Sci. Rep.* **2017**, *7*, 14564. [[CrossRef](#)]
39. Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
40. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
41. Broad Institute. Picard Toolkit, Broad Institute. GitHub Repository . Available online: <http://broadinstitute.github.io/picard/> (accessed on 3 February 2021).
42. Sargolzaei, M. *SNP1101 User's Guide*; Higgs Gene Solutions Inc.: Guelph, ON, Canada, 2014; Version 1.0.
43. McQuillan, R.; Leutenegger, A.-L.; Abdel-Rahman, R.; Franklin, C.S.; Pericic, M.; Barac-Lauc, L.; Smolej-Narancic, N.; Janicijevic, B.; Polasek, O.; Tenesa, A.; et al. Runs of Homozygosity in European Populations. *Am. J. Hum. Genet.* **2008**, *83*, 359–372. [[CrossRef](#)]
44. Nei, M. Genetic Distance between Populations. *Am. Nat.* **1972**, *106*, 283–292. [[CrossRef](#)]
45. Weir, B.S.; Cockerham, C.C. Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **1984**, *38*, 1358–1370. [[CrossRef](#)]
46. Pembleton, L.W.; Cogan, N.O.I.; Forster, J.W. St AMPP: An R Package for Calculation of Genetic Differentiation and Structure of Mixed-Ploidy Level Populations. *Mol. Ecol. Resour.* **2013**, *13*, 946–952. [[CrossRef](#)]
47. Excoffier, L.; Smouse, P.E.; Quattro, J.M. Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics* **1992**, *131*, 479–491. [[CrossRef](#)] [[PubMed](#)]
48. Jombart, T.; Devillard, S.; Balloux, F. Discriminant Analysis of Principal Components: A New Method for the Analysis of Genetically Structured Populations. *BMC Genet.* **2010**, *11*, 94. [[CrossRef](#)]
49. Jombart, T.; Ahmed, I. adegenet 1.3-1: New Tools for the Analysis of Genome-Wide Snp Data. *Bioinformatics* **2011**, *27*, 3070–3071. [[CrossRef](#)]
50. Alexander, D.H.; Novembre, J.; Lange, K. Fast Model-Based Estimation of Ancestry in Unrelated Individuals. *Genome Res.* **2009**, *19*, 1655–1664. [[CrossRef](#)] [[PubMed](#)]
51. Hill, W.G.; Robertson, A. Linkage Disequilibrium in Finite Populations. *Theor. Appl. Genet.* **1968**, *38*, 226–231. [[CrossRef](#)] [[PubMed](#)]
52. Lynch, M.; Walsh, B. *Genetics and Analysis of Quantitative Traits*; Sinauer Associates: Sunderland, MA, USA, 1998.
53. Hui, T.-Y.J.; Burt, A. Estimating Linkage Disequilibrium from Genotypes under Hardy-Weinberg Equilibrium. *BMC Genet.* **2020**, *21*, 21. [[CrossRef](#)] [[PubMed](#)]
54. Nilsen, H.; Hayes, B.; Berg, P.R.; Roseth, A.; Sundsaasen, K.K.; Nilsen, K.; Lien, S. Construction of a Dense SNP Map for Bovine Chromosome 6 to Assist the Assembly of the Bovine Genome Sequence. *Anim. Genet.* **2008**, *39*, 97–104. [[CrossRef](#)]
55. Sved, J. Linkage Disequilibrium and Homozygosity of Chromosome Segments in Finite Populations. *Theor. Popul. Biol.* **1971**, *2*, 125–141. [[CrossRef](#)]
56. Hayes, B.J.; Visscher, P.M.; McPartlan, H.C.; Goddard, M.E. Novel Multilocus Measure of Linkage Disequilibrium to Estimate Past Effective Population Size. *Genome Res.* **2003**, *13*, 635–643. [[CrossRef](#)]
57. Cai, Z.; Villumsen, T.M.; Asp, T.; Guldbrandtsen, B.; Sahana, G.; Lund, M.S. SNP Markers Associated with Body Size and Pelt Length in American Mink (Neovison Vison). *BMC Genet.* **2018**, *19*, 103. [[CrossRef](#)] [[PubMed](#)]
58. Gurgul, A.; Miksza-Cybulska, A.; Szmatola, T.; Jasielczuk, I.; Piestrzyńska-Kajtoch, A.; Fornal, A.; Semik-Gurgul, E.; Bugno-Poniewierska, M. Genotyping-by-Sequencing Performance in Selected Livestock Species. *Genomes* **2019**, *111*, 186–195. [[CrossRef](#)]
59. Belliveau, A.M.; Farid, A.; O'Connell, M.; Wright, J.M. Assessment of Genetic Variability in Captive and Wild American Mink (Mustela Vison) Using Microsatellite Markers. *Can. J. Anim. Sci.* **1999**, *79*, 7–16. [[CrossRef](#)]
60. Zhang, Q.; Calus, M.P.L.; Guldbrandtsen, B.; Lund, M.S.; Sahana, G. Estimation of Inbreeding Using Pedigree, 50k SNP Chip Genotypes and Full Sequence Data in Three Cattle Breeds. *BMC Genet.* **2015**, *16*, 1–11. [[CrossRef](#)] [[PubMed](#)]
61. Schiavo, G.; Bovo, S.; Bertolini, F.; Tinarelli, S.; Dall'Olio, S.; Costa, L.N.; Gallo, M.; Fontanesi, L. Comparative Evaluation of Genomic Inbreeding Parameters in Seven Commercial and Autochthonous Pig Breeds. *Animal* **2020**, *14*, 910–920. [[CrossRef](#)] [[PubMed](#)]
62. Rodríguez-Ramilo, S.T.; Fernández, J.; Toro, M.Á.; Hernández, D.; Villanueva, B. Genome-Wide Estimates of Coancestry, Inbreeding and Effective Population Size in the Spanish Holstein Population. *PLoS ONE* **2015**, *10*, e0124157. [[CrossRef](#)]
63. Pemberton, T.J.; Absher, D.; Feldman, M.W.; Myers, R.M.; Rosenberg, N.A.; Li, J.Z. Genomic Patterns of Homozygosity in Worldwide Human Populations. *Am. J. Hum. Genet.* **2012**, *91*, 275–292. [[CrossRef](#)] [[PubMed](#)]
64. Howrigan, D.; Simonson, M.A.; Keller, M.C. Detecting Autozygosity through Runs of Homozygosity: A Comparison of Three Autozygosity Detection Algorithms. *BMC Genom.* **2011**, *12*, 460. [[CrossRef](#)]

65. Ferenčaković, M.; Hamzić, E.; Gredler, B.; Solberg, T.; Klemetsdal, G.; Curik, I.; Sölkner, J. Estimates of Autozygosity Derived from Runs of Homozygosity: Empirical Evidence from Selected Cattle Populations. *J. Anim. Breed. Genet.* **2013**, *130*, 286–293. [[CrossRef](#)]
66. Mastrangelo, S.; Tolone, M.; Di Gerlando, R.; Fontanesi, L.; Sardina, M.T.; Portolano, B. Genomic Inbreeding Estimation in Small Populations: Evaluation of Runs of Homozygosity in Three Local Dairy Cattle Breeds. *Animal* **2016**, *10*, 746–754. [[CrossRef](#)]
67. Bush, W.S.; Moore, J.H. Chapter 11: Genome-Wide Association Studies. *PLoS Comput. Biol.* **2012**, *8*, e1002822. [[CrossRef](#)]
68. Hayes, B.J.; Goddard, M.E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* **2001**, *157*, 1819–1829.
69. Thomasen, J.R.; Sørensen, A.C.; Su, G.; Madsen, P.; Lund, M.S.; Guldbrandtsen, B. The Admixed Population Structure in Danish Jersey Dairy Cattle Challenges Accurate Genomic Predictions. *J. Anim. Sci.* **2013**, *91*, 3105–3112. [[CrossRef](#)]
70. Brito, L.F.; Jafarikia, M.; Grossi, D.D.A.; Kijas, J.W.; Porto-Neto, L.R.; Ventura, R.V.; Sargolzaei, M.; Schenkel, F.S. Characterization of Linkage Disequilibrium, Consistency of Gametic Phase and Admixture in Australian and Canadian Goats. *BMC Genet.* **2015**, *16*, 1–15. [[CrossRef](#)] [[PubMed](#)]
71. Ødegård, J.; Moen, T.; Santi, N.; Korsvoll, S.A.; Kjøglum, S.; Meuwissen, T.H.E. Genomic Prediction in an Admixed Population of Atlantic Salmon (*Salmo Salar*). *Front. Genet.* **2014**, *5*, 402. [[CrossRef](#)]
72. Toosi, A.; Fernando, R.L.; Dekkers, J.C.M. Genomic Selection in Admixed and Crossbred Populations. *J. Anim. Sci.* **2010**, *88*, 32–46. [[CrossRef](#)] [[PubMed](#)]
73. Nei, M.; Li, W.-H. Linkage Disequilibrium in Subdivided Populations. *Genetics* **1973**, *75*, 213–219. [[CrossRef](#)] [[PubMed](#)]
74. Wilson, J.F.; Goldstein, D.B. Consistent Long-Range Linkage Disequilibrium Generated by Admixture in a Bantu-Semitic Hybrid Population. *Am. J. Hum. Genet.* **2000**, *67*, 926–935. [[CrossRef](#)]
75. Ohta, T. Linkage Disequilibrium Due to Random Genetic Drift in Finite Subdivided Populations. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 1940–1944. [[CrossRef](#)] [[PubMed](#)]
76. Bansal, V.; Bashir, A.; Bafna, V. Evidence for Large Inversion Polymorphisms in the Human Genome from Hapmap Data. *Genome Res.* **2007**, *17*, 219–230. [[CrossRef](#)]
77. Cáceres, A.; Sindi, S.S.; Raphael, B.J.; Cáceres, M.; González, J.R. Identification of Polymorphic Inversions from Genotypes. *BMC Bioinform.* **2012**, *13*, 28. [[CrossRef](#)] [[PubMed](#)]
78. Lewontin, R.C.; Kojima, K.-I. The Evolutionary Dynamics of Complex Polymorphisms. *Evolution* **1960**, *14*, 458. [[CrossRef](#)]
79. Voight, B.F.; Kudravalli, S.; Wen, X.; Pritchard, J.K. A Map of Recent Positive Selection in the Human Genome. *PLoS Biol.* **2006**, *4*, e72. [[CrossRef](#)]
80. Slatkin, M. Linkage Disequilibrium in Growing and Stable Populations. *Genetics* **1994**, *137*, 331–336. [[CrossRef](#)] [[PubMed](#)]
81. Schmegner, C.; Hoegel, J.; Vogel, W.; Assum, G. Genetic Variability in a Genomic Region with Long-Range Linkage Disequilibrium Reveals Traces of a Bottleneck in the History of the European Population. *Qual. Life Res.* **2005**, *118*, 276–286. [[CrossRef](#)]
82. Koch, E.; Ristroph, M.; Kirkpatrick, M. Long Range Linkage Disequilibrium across the Human Genome. *PLoS ONE* **2013**, *8*, e80754. [[CrossRef](#)]
83. Franssen, S.U.; Nolte, V.; Tobler, R.; Schlötterer, C. Patterns of Linkage Disequilibrium and Long Range Hitchhiking in Evolving Experimental *Drosophila Melanogaster* Populations. *Mol. Biol. Evol.* **2015**, *32*, 495–509. [[CrossRef](#)]
84. Khatkar, M.S.; Nicholas, F.W.; Collins, A.R.; Zenger, K.R.; Cavanagh, J.A.L.; Barris, W.; Schnabel, R.; Taylor, J.F.; Raadsma, H.W. Extent of Genome-Wide Linkage Disequilibrium in Australian Holstein-Friesian Cattle Based on a High-Density SNP Panel. *BMC Genom.* **2008**, *9*, 187. [[CrossRef](#)] [[PubMed](#)]