Detection of SNP markers and practical applications in mink (*Neovison vison*)

P.F. Larsen[1,4], D. Demontis[2], V.H. Nielsen[3], J.P. Thirstrup[3] and C. Pertoldi[4]
[1]Kopenhagen Research, Agro Food Park 15, 8200 Aarhus N, Denmark; [2]Institute of Biomedicine, Aarhus University, Wilhelm Meyrs Allé 4, 8000 Århus C, Denmark; [3]Department of Genetics and Biotechnology, Aarhus University, Blichers Allé, 8830 Tjele, Denmark; [4]Department of Biological Sciences, Aarhus University, Ny Munkegade, 8000 Århus C, Denmark; pfl@kopenhagenfur.com

Abstract

Single nucleotide polymorphisms (SNPs) are widespread sources of sequence variation in genomes and are currently applied in fields such as forensic and animal breeding genetics. Here we present how it is possible to quickly identify informative SNPs in mink based on a SNP marker panel created for the domestic dog which is phylogenetic far away from the mink. Applying the commercial available Canine SNP Genotyping BeadChip (Illumina®) to genotype 12 American mink from 4 different color strains we identified 111 SNPs in mink. Based on the 111 SNPs genetic population structure analysis showed similar levels of relatedness and genetic differentiation than previously observed using microsatellite markers in mink. We suggest strategies for applying SNP markers in practical farm management, by genotyping individuals with a few very informative genetic markers, considering genetic information that would increase the heterozygosity and have positive effects on the breeding result in mink production. Alternative applications of SNP markers for parentage analysis, individual identification and forensic identification in mink are discussed. Finally, identifying markers associated with important production traits and combined with neutral markers can combined be applied to improve genetic screening on farm levels and for developing new breeding strategies on Danish mink farms.

Keywords: genetic distance, relatedness, inbreeding, mink farm

Introduction

In recent years a large number of microsatellite markers have been published for mink (*Neovison vison*) (e.g. Anistoroaei *et al.,* 2006, 2007). However, microsatellite markers are expensive in genotyping and difficult to genotype in a large scale approach, furthermore standardization among different laboratories is a big problem. Therefore new markers are desired and here SNPs represent the most widespread source of sequence variation within genomes (Brumfield *et al.,* 2003) and they have emerged as valuable genetic markers also in animal and conservation genetics. SNPs are biallelic markers, and are inherently less informative if compared to the multiallelic microsatellites, when used for individual identification, parentage analysis and population genetics. However, SNPs compared to microsatellites has several potential advantages like a better amplification power, lower genotyping errors and unambiguous allele comparisons. However, their simpler mutational dynamics strongly reduce risks of homoplasy (Chen and Sullivan, 2003). Furthermore, fast and inexpensive methods are now being developed to screen hundreds or thousands of SNPs per sample per population (Wang *et al.,* 2009).

From the few genetic studies conducted on mink so far, it has been shown that there are large genetic differences between different strains (color types) of mink (e.g. Michalska-Parda *et al.,* 2009; Larsen *et al.,* 2009). Demontis *et al.* (2011) have recently shown a very strong and highly significant negative correlation between the level of relatedness (R) and fecundity ($r = 0.536$, $P<0.001$). The high correlation between the level of R and breeding result, indicate that inbreeding decreases fecundity in farmed mink, demonstrating therefore, the importance of avoiding severe inbreeding in order to keep the fecundity high. This information could potentially be applied in practical farm management. By genotyping individuals with a few very informative markers, the degree of genetic differentiation (F_{ST}) estimates between farms within strains, and the degree of R within farms and strains etc., can be estimated and used to evaluate which animals would be the best to purchase in practice in order to keep the fecundity as high as possible (Larsen *et al.,* 2011).

The main aims of this study, were to genotype American mink on the Illumina Canine BeadChip in order: (1) to identify Single Nucleotide Polymorphisms (SNPs) in the mink; (2) to use these SNPs to estimate the level of R within strains as a measure of the inbreeding level (as inbreeding is produced by mating among relatives); (3) to elucidate the genetic divergence between strain, and finally (4) to estimate the probability of identity (PI) for an increasing number of loci, i.e. the probability that different individuals by chance share an identical genotype (potential useful for individual analysis and forensic genetics in mink).

Material and methods

DNA extraction and SNPs genotyping

12 American mink representing four different strains were included in this study; three Mahogany (consisting of a trio: female, male and their male offspring, included to follow segregation of alleles from parent to offspring and ensure SNPs are 'real SNPs'), three Sapphire, three Black and three Brown/Glow, from different Danish mink farms. Except for the mink in the Mahogany strain, all collected mink were unrelated. The unrelated mink did not have the same mother or father in the breeding season 2010. Blood or liver samples were kept in a freezer at -18 °C until analyses were performed. DNA was extracted the using the QIAGEN DNeasy Blood and Tissue Extraction Kit (QIAGEN Ltd). Genotyping of SNPs (single nucleotide polymorphisms) was performed using Canine SNP BeadChip (Illumina®) according to the manufacturer's protocol (Infinium II Multi-Sample), followed by analysis using the software GenomeStudio (Illumina®).

Statistical analyses

Variability analysis: Unbiased Expected Heterozygosity (UH_E), percent of polymorphic loci (P%), was performed using the software GENALEX v. 6.1 (Peakall and Smouse, 2006). GENALEX was also used for estimating the level of relatedness (R) within each strain (Ritland, 1996), and for estimating the probability of identity (PI) for an increasing number of loci, i.e. the probability that different individuals by chance share an identical genotype (the male Mahogany offspring was excluded from this test).

The genetic divergence between strains and within strains among farms was estimated by calculating the pair-wise F_{ST} (Weir and Cockerham, 1984) values using the software GENALEX and the significance of the F_{ST}s and linkage disequilibrium between loci (within strain) were estimated using the software GENEPOP 3.4 (Raymond and Rousset, 1995), (the male Mahogany offspring was excluded from this test).

Results

The CanineHD genotyping BeadChip SNP chip contains more than 170,000 markers placed on the CanFam2.0 reference sequence. Very stringent quality control was performed in order to identify SNPs with reliable calls. Only SNPs with a call rate above 91% (called in all but one individual) having a minor allele frequency above 0.25 and below 0.45 were included. Only 440 SNPs passed these criteria. After manually checking all clusters 111 SNPs remained for further analysis all demonstrated nice clusters with a call rate of 100%.

Among the known mother-father-offspring trio, no errors were detected for the 111 SNPs. Although the power to detect errors at individual loci was low, the fact that no errors were detected among the loci investigated suggests that SNPs can be typed reliably and that the true error rate may be lower than the 1% error rate assumed in the paternity simulations. If the error rate was assumed to be zero, a slightly smaller panel of SNPs would be sufficient for successful paternity analysis (data not shown).

The unbiased expected heterozygosity (UH_E) varied between 0.110 in Mahogany to 0.570 in Brown/Glow and the percent of polymorphic loci (P%) varied between 19.82% in Mahogany and 96.40% in Black and Brown/Glow (see Table 1). No linkage disequilibrium was detected between the 111 loci investigated, which were therefore all included in the statistic analyses.

Table 1. Id-number (N) of individuals genotyped from each farm and breeding line. The unbiased expected heterozygosity (UH_E), Standard Error (SE), % of polymorphic loci (P %) for each strain.

N	Strain	UH_E	SE	P %
1	SAPH	0.491	0.013	93.69%
2	SAPH			
3	SAPH			
4	BLACK	0.475	0.012	96.40%
5	BLACK			
6	BLACK			
7	WILD	0.570	0.012	96.40%
8	WILD			
9	WILD			
10 (son)	MAH	0.110	0.021	19.82%
11 (female)	MAH			
12 (male)	MAH			

When comparing the Black, Sapphire, Brown/Glow and Mahogany mink we found the highest mean R for the Mahogany strain (0.478) (bringing however in mind that the Mahogany strain is composed by a trio) followed by the Sapphire strain (0.238) and the lowest R in the Brown/Glow strain (0.067) and lastly, Black (-0.099) (see Table 2).

The genetic divergence (F_{ST}) between strains ranged from 0.074 (between the Black and Brown/Glow strains) to 0.449 (between the Sapphire and Mahogany strains) and were all highly significant ($P<0.0001$) (Table 3).

Probability of identity based on the 111 SNPs was estimated to be between 9.4E-09 for the Mahogany strain and 5.4E-45 for the Brown/Glow (see Figure 1).

Discussion

The rank of the R-values is mostly in agreements with the rank obtained by Demontis *et al.* (2011). It was not surprising that the highest degree of R was found in the Mahogany as the individuals representing this strain was a trio. However the Mahogany strain has been maintained isolated at very low population size for five generations and that is also reflected in the extremely

Table 2. Mean relatedness R (Ritland 1996) Upper (Uv error) and lower (Lv error) error bars bound the 95% confidence interval about the mean values as determined by bootstrap resampling (99 bootstrap) for SAPH, BLACK and BROWN/GLOW (B/G) without MAHOGANY (MAHO) (a) and with MAHO (b) (TRIO, mother, father and son).

		SAPH	BLACK	B/G	MAHO
a. Mean Bootstrap Value		0.163	-0.073	0.032	
	Uv error	0.058	0.08	0.015	
	Lv error	0.1	0.137	0.023	
b. Mean Bootstrap Value		0.238	-0.099	0.067	0.478
	Uv error	0.078	0.084	0.023	0.032
	Lv error	0.138	0.119	0.017	0.02

Table 3. Pairwise F_{ST} between strains. All the F_{ST} values were highly significant (P<0.0001).

	Sapphire	Black	Brown/glow
Sapphire			
Black	0.161		
Brown/glow	0.041	0.074	
Mahogany	0.449	0.169	0.301

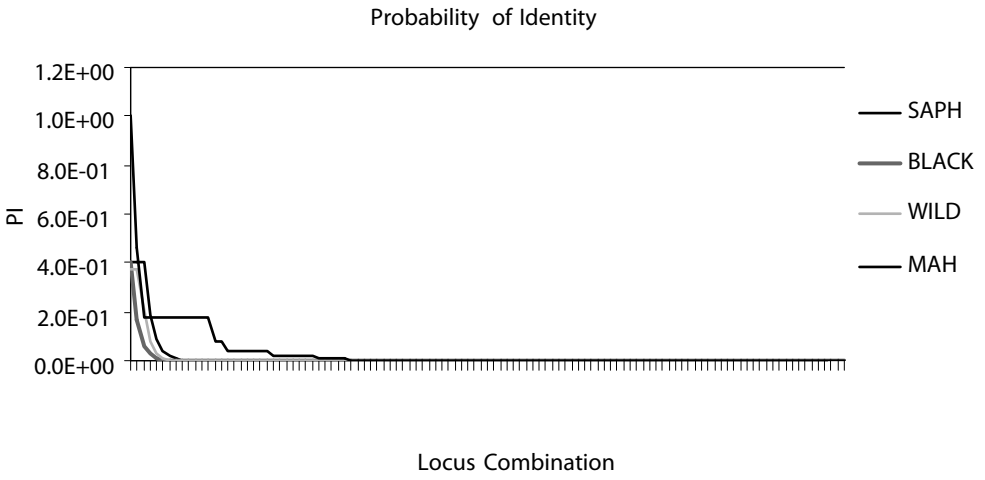

Figure 1. Probability of identity with increasing locus combination (excluding the son in Mahogany).

low level of polymorphism observed in this strain compared to the others. The relatively low level of R found in the Black strain in this study is the result of a large production group of Black mink at the Research Farm in Foulum. The observed degree of genetic differentiation between strains was also in agreement with the results shown by Demontis *et al.* (2011).

The use of SNPs as a genetic marker has proven advantageous although some disadvantages are experienced. Thus SNP identification is a challenge due to major laboratory efforts and cost associated with the findings of SNPs (Ryynänen *et al.,* 2007). Identification of polymorphic and informative SNPs in small or differentiated populations is difficult, even when polymorphic sequences for closely related species are available. Furthermore, not all SNPs detected in the sequences are in specific positions suitable for primer and probe design. Thus in this study we have shown that it is possible to use a chip developed for another species and detect useful SNP, this allow for detecting and applying SNPs in species where genomic resources are still limited. To improve genotyping success and reliability, it is preferable to amplify DNA fragments as short as possible. The use of single nucleotide polymorphisms (SNPs) requires only the amplification of very short fragments and this makes them particularly suitable for non-invasive genetic monitoring projects (Seddon *et al.,* 2005).

In mink SNPs could be used for selection of breeding animals or by farmers when buying in new males to ensure high fecundity in their breeding stock. Knowing the genetic distance between pure lines will make it possible to suggest better breeding strategies. Thus it is expected to obtain heterosis by mating genetically distant strains/lines. Qualified choices of breeding animals can be made when knowing both the 'genetic profile' of the animals on the farm as well as the profile of the potential breeding animals on top farms. Thus heterosis may be obtained for several traits such as fecundity, body size, skin length, weight at birth within color types. Knowing, the genetic distance among breeding lines from different farms within color types, it may be possible to observe heterosis within color types, which is often more attractive in practical farm

management. QTL studies or association mapping of quantitative traits would help to identify markers associated with important production traits and these important markers combined with neutral markers can combined be applied to improve genetic screening for optimal animals for breeding on Danish mink farms.

A subset of very informative SNPs (48 or 96) could be genotyped using relatively cheap technologies like, for example, VeraCode or Fluidigm EP1 System. The costs of SNP microchip beads application is gradually decreasing, making it available to a wider range of users. These new techniques for high-throughput SNP genotyping could be used in non-invasive DNA monitoring projects that usually require a huge number of samples to be analyzed.

Based on the present study, a panel of 50-60 SNP-based genetic markers that are highly informative in mink could be designed and would be sufficient to conduct identity analyses in mink (Figure 1). This SNP panel would be conveniently genotyped using VeraCode (Illumina Inc.) or SNPlex (ABI) technology at a much reduced cost relative to typical microsatellite genotyping. Usage of an SNP panel as proposed here would allow not only reliable parentage and identity analysis but also would provide a standardized panel for exchange of genotype data between laboratories.

Conclusions

In mink, genotyping of individuals could help overcome problems with reduced fecundity, by mating genetically divergent breeding animals to optimize fitness or to obtain heterosis effects among strains within color types. Genotyping could also be employed to check identity of samples for diagnostic purposes and to detect identity of recovered stolen skins. We therefore see a huge advantage in applying molecular genetic information in practical mink farming in the future.

References

Anistoroaei R., Farid A., Benkel B., Cirera S. and Christensen K., 2006. Isolation and characterization of 79 microsatellite markers from the American mink (*Mustela vison*). Animal Genetics 37: 185-188.

Anistoroaei R., Menzorov A., Serov O., Farid A. and Christensen K., 2007. The first linkage map of the American mink (*Mustela vison*). Animal Genetics 38: 384-388.

Bonin A., Bellemain E., Eidesen P.B., Pompanon F., Brochmann C. and Taberlet P., 2004. How to track and assess genotyping errors in population genetics studies. Mol Ecol, 13: 3261-3173.

Broquet T. and Petit E., 2004. Quantifying genotyping errors in noninvasive population genetics. Mol Eco 13: 3601-3608.

Brumfield R.T., Beerli P., Nickerson D.A. and Edwards S.V., 2003. The utility of single nucleotide polymorphisms in inferences of population history. Trends Ecol Evol 18: 249-256.

Chen X. and Sullivan P.F., 2003. Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. Pharmacogenom J 3: 77-96.

Demontis D., Larsen P.F., Baekgaard H., Soenderup M., Hansen B.K., Nielsen V.H., Loeschcke V., Zalewski A., Zalewska H., and Pertoldi C., 2011. Inbreeding affects fecundity of American mink (*Neovison vison*) at Danish farm mink. Animal Genetics (in press).

Larsen, P. F., Bækgaard, H., Sønderup, M., Møller, H. H., Hansen, B. K., Nielsen, V. H., Demontis, D., Loeschcke, V. and Perfoldi, C., 2010. Level of inbreeding and genetic distance in Danish farm mink (*Neovison vison*). Annual report 2009, 27-31, Danish Fur Breeders Research Center, Holstebro, Denmark

Larsen, P.F., 2011. The effect of inbreeding on the breeding result in Danish mink (In Danish: Indavls betydning for avlsresultat i danske mink). Internal report Foulum September 2011.

Michalska-Parda A., BrzeziÑski M., Zalewski A., and Kozakiewicz M., 2009. Genetic variability of feral and ranch American mink *Neovison vison* in Poland. Acta Theriologica, 54: 1-10.

Morin P.A. and McCarthy M., 2007. Highly accurate SNP genotyping from historical and low-quality samples. Molecular Ecology Notes, 7: 937-046.

Peakall R. and Smouse P.E., 2005. genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes, 6: 288-295.

Quintáns B., Alvarez-Iglesias V., Phillips C., Lareu M.V. and Carracedo A., 2004. Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing. Forensic Sci Int, 140: 251-257.

Raymond M. and Rousset F., 1995. GENEPOP (version 1.2): Population genetic software for exact tests and ecumenicism. Journal of Heredity, 86: 248-249.

Ritland K., 1996. Marker-based method for inferences about quantitative inheritance in natural populations. Evolution; 50: 1062-1073.

Ryynänen H.J., Tonteri A., Vasemägi A. and Primmer C.R., 2007. A comparison of biallelic markers and microsatellites for the estimation of population and conservation genetic parameters in Atlantic Salmon (*Salmo salar*). J Hered, 7: 692-704.

Seddon J.M., Parker H.G., Ostrander E.A. and Ellegren H., 2005. SNPs in ecological and conservation studies: a test in the Scandinavian wolf population. Mol Ecol, 14: 503-511.

Sobrino B., Brión M. and Carracedo A., 2005. SNPs in forensic genetics: a review on SNP typing methodologies. Forensic Sci. Int. 154: 181-94.

Troell K., Mattsson J.G., Alderborn A. and Höglund J., 2003. Pyrosequencing analysis identifies discrete populations of *Haemonchus contortus* from small ruminants. Int J Parasitol, 33: 765-771.

Tschentscher F., Frey U.H. and Bajanowski T., 2008. Amelogenin sex determination by pyrosequencing of short PCR products. Int J Legal Med. 122: 333-335.

Vignal A., Milan D., SanCristobal M. and Eggen A., 2002. A review on SNP and other types of molecular markers and their use in animal genetics. Genet Sel Evol, 34: 275-305.

Wang J., Lin M., Crenshaw A., Hutchinson A., Hicks B., Yeager M., Berndt S., Huang W-Y., Hayes R.B., Chanock S.J., Jones R.C. and Ramakrishnan R., 2009. High-throughput single nucleotide polymorphism genotyping using nanofluidic Dynamic Arrays. BMC Genomics, 10: 561-574.

Weir B.S. and Cockerham C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution, 38: 1358-1370.