



Simulation Study on Parentage Analysis with SNPs in the Japanese Black Cattle Population

Takeshi Honda*, Tomohiro Katsuta¹ and Fumio Mukai¹

Food Resources Education and Research Center, Graduate School of Agricultural Science,
Kobe University, Uzurano Kasai, Hyogo 675-2103, Japan

ABSTRACT : Parentage tests using polymorphic DNA marker are commonly performed to avoid incorrect recording of the parental information of livestock animals, and single-nucleotide polymorphisms (SNPs) are becoming the method of choice. In Japanese Black cattle, parentage tests based on the exclusion method using microsatellite markers are currently conducted; however, an alternative SNP system aimed at parentage tests has recently been developed. In the present study, two types of simulations were conducted using the pedigree data of two subpopulations in the breed (subpopulations of Hyogo and Shimane prefectures) in order to examine the effect of actual genetic and breeding structures. The first simulation (simulation 1) investigated the usefulness of SNPs for excluding a close relative of the true sire; the second one (simulation 2) investigated the accuracy of sire identification tests for multiple full-sib putative sires by a combined method of exclusion and paternity assignment based on the LOD score. The success rates of excluding a single full-sib and sire of the true sires were, respectively, 0.9915 and 0.9852 in Hyogo and 0.9848 and 0.9852 in Shimane, when 50 SNPs with minor allele frequency (MAF: q) of $0.25 \leq q \leq 0.35$ were used in simulation 1. The success rates of sire identification tests based solely on the exclusion method were relatively low in simulation 2. However, assuming that 50 SNPs with MAF of $0.25 \leq q \leq 0.35$ or $0.45 \leq q \leq 0.5$ were available, the total success rates including achievements due to paternity assignment were, respectively, 0.9430 and 0.9681 in Hyogo and 0.8999 and 0.9399 for Shimane, even when each true sire was assumed to compete with 50 full-sibs. (**Key Words :** Parentage Analysis, SNPs, Simulation, Japanese Black, Pedigree Analysis)

INTRODUCTION

The production of livestock animals through such reproductive technologies as artificial insemination and embryo transfer is common. Although the use of these technologies has many advantages, it involves the risk of intentional or accidental recording of incorrect parental information; from the perspective of consumer protection, this risk should be minimized to the greatest possible extent. The accumulation of correct pedigree information is also fundamental to the precise prediction of genetic merits of animals. For the avoidance of the incorrect pedigree recording, importance of parentage analysis using polymorphic DNA markers is recognized in the field of livestock production (Heaton et al., 2002; Werner et al., 2004; Dodds et al., 2005; Eenennaam, 2007; Gomez-Raya

et al., 2008). The use of SNPs for analysis is recommended because of its low mutation and genotyping error rates, abundance across the genome, and adaptability to low-cost genotyping (Anderson and Garza, 2006; Rohrer et al., 2007; Baruch and Weller, 2008).

Parentage analysis techniques are largely classified as exclusion and likelihood-based methods. Exclusion methods exclude the parentage of a putative parent who has an incompatible genotype as a parent, based on Mendelian rules of inheritance; likelihood-based methods assign parentage to the putative parent having the largest likelihood ratio among several putative parents. The accuracy of both techniques depends on such specific breeding and genetic structure in the population of interest as allele frequencies, genotype frequencies, mating systems, the number of putative parents, and the degree of their relatedness. Although there are a number of recent theoretical developments in likelihood-based parentage analysis for livestock (Dodds et al., 2005; Hill et al., 2008) and natural populations (Marshall et al., 1998; Jones and Ardren, 2003; Kalinowski et al., 2007), most methods are

* Corresponding Author: Takeshi Honda. Tel: +81-0-790-49-3122, Fax: +81-0-790-49-0343, E-mail: takhonda@diamond.kobe-u.ac.jp

¹ Wagyu Registry Association, Kyoto 603-8555, Japan.

Received December 8, 2008; Accepted May 18, 2009

evaluated by simulations assuming randomly mating population under Hardy-Weinberg equilibrium (HWE) or nearly so. Therefore, the results of such evaluations are not simply applicable to the livestock population under intensive selection.

Currently, in the Japanese Black cattle, an exclusion method using microsatellites is carried out for candidate AI sires before selection and offspring produced by embryo transfer, only to confirm that their fathers or mothers are correctly reported. However, an alternative SNP system aimed at parentage analysis for the breed has recently been developed, and nearly 80 loci are now available (Dr. Mannen of Kobe University, personal communication). Evaluation of accuracy prior to the introduction of a new DNA marker set is necessary. In addition, since the introduction of the likelihood-based method widens the applicable scope of parentage analysis (e.g., when several putative parents cannot be excluded by a given marker set), the accuracy of this type of parentage analysis is also of interest.

In this study, we used computer simulations to examine the precision of several parentage tests using SNP system. In the simulations, pedigree data of subpopulations of Japanese Black cattle were used to examine breeding and genetic structure largely deviating from the randomly mating population.

MATERIALS AND METHODS

Data

We simulated situations in which parentage tests were carried out for all the offspring born in each of the two Hyogo and Shimane prefecture subpopulations in 2004. We examined whether the difference in degree of relatedness between individuals in each population affects test precision. Pedigrees of Japanese Black cattle, which are recorded by the Wagyu Registry Association, can be traced to the ancestral population of the foundation period (1944 or before). All available pedigree records of the two subpopulations were used in the simulation. According to conventional pedigree analysis, ancestors with unknown parents are referred to as founders and all their descendants as non-founders. Table 1 presents the numbers of offspring in the two subpopulations, along with those of their sires, dams, and all the ancestors that appeared in the pedigree data. Since the subpopulation of Hyogo has been closed to other populations for a long time, the amount of inbreeding and relationships among animals is considerably higher than for the other subpopulations in the breed (cf. Honda et al., 2002). Average inbreeding coefficients and average coancestries of the two offspring populations are also presented in Table 1.

Table 1. Numbers of offspring born in each subpopulation of Hyogo and Shimane prefecture in 2004 (N_o), and those of their sires (N_s) and dams (N_d), and the total number of ancestors appearing in each pedigree file (N_a).

	Hyogo	Shimane
N_o	1,316	850
N_s	55	74
N_d	1,307	803
N_a	9,749	1,2370
\bar{F}_o^1	0.196	0.067
\bar{f}_o^2	0.226	0.098

¹ Average inbreeding coefficient of the offspring population.

² Average coancestry within the offspring population.

Simulation

We conducted two simulations. Simulation 1 investigated the precision of current parentage tests carried out in the Japanese Black cattle, with the assumption that the SNP system was used instead of microsatellites. Simulation 2 examined the extent to which the likelihood-based method is useful if used with the SNP system.

In the simulation, all the offspring and parents were assumed to be genotyped at n_l ($= 10, 30, \text{ and } 50$) unlinked SNP loci segregating two alleles with minor allele frequencies (MAF) of q ($= 0.05 \leq q \leq 0.15, 0.25 \leq q \leq 0.35, \text{ and } 0.45 \leq q \leq 0.5$) in each parental population of sires and dams (i.e., nine conditions in terms of marker information were simulated). To generate possible genotype patterns of all the offspring and their parents, slightly modified gene dropping simulation (MacCluer et al., 1986) was applied to the pedigree data. For a single locus, the simulation randomly generated the genotype of each founder by sampling two alleles from a hypothetical base population, in which frequencies of the two types of alleles followed the Dirichlet distribution with all parameters set to one (i.e., the marginal distribution of each allele frequency is uniform). Following the Mendelian segregation rules, alleles assigned to the founders were then transmitted along the pedigree, and all the genotypes of the non-founders in the pedigree were determined. Mutations were ignored because SNPs have considerably low mutation rates and would not substantially affect the results. This process was repeated until n_l loci having MAF of q in the parental populations of each sex were generated.

We assumed in the simulation that the dam of each offspring is known but the sire is unknown, and investigated the results of excluding a putative sire that is not the true sire of the offspring (simulation 1), and of sire identification among multiple putative sires related to each other (simulation 2).

Simulation 1 : After the genotypes of all the offspring

and actual parents in the pedigree were determined at n_l loci, one hypothetical full-sib was generated for each actual sire, using the simulated genotypes of sires and dams of the actual sires (i.e., paternal grand sire and grand dam of the offspring). Hypothetical paternity tests were then carried out for each offspring, assuming that the full-sib of the true sire was a putative sire. If incompatibility of genotypes between the putative sire and offspring-dam set was found at one or more loci, the test was regarded as successful. The success rate of this paternity test in each subpopulation was computed as

$$acc1_{FS} = \frac{n_{FS}}{N_T}$$

where N_T is the number of tested trios (set of offspring, dam, and putative sire), which is equivalent to the number of offspring in each subpopulation, and n_{FS} is the number of trios in which a putative sire can be excluded from paternity.

Because paternal grand sires of the offspring also seem to have comparative relationships to the actual sires, another hypothetical test was carried out assuming that the paternal grand sire of the offspring was a putative sire. The success rate was computed as

$$acc1_{FATH} = \frac{n_{FATH}}{N_T}$$

where n_{FATH} is the number of trios in which a putative sire (paternal grand sire of the offspring) can be excluded from paternity.

Simulation 2 : In a manner similar to that used in simulation 1, N_{FS} (= 1 to 50) hypothetical full-sibs were generated for all the actual sires after genotype determination, and hypothetical paternity identification tests were carried out for each offspring, assuming that $N_{FS}+1$ males, including the true sire of the offspring, were putative sires. Each identification test began with the exclusion method, so that every putative sire having an incompatible genotype as a parent at one or more loci was excluded from paternity. If all the putative sires except the true sire were excluded, the test was considered to be correctly accomplished. If several sires, including the true sire, could not be excluded, the LOD score (natural logarithm of likelihood ratio, cf. Meagher, 1986; Marshall et al., 1998; Jones and Ardren, 2003) was calculated for each non-excluded putative sire, and paternity was assigned to the putative sire with the highest LOD score.

To examine different types of performances, three success rates were computed in this simulation. The first one investigated the performance of the exclusion method,

and was computed by

$$acc2_{EX} = \frac{n_{EX}}{N_T}$$

where n_{EX} is the number of trios, in which the true sire was identified by only the exclusion method. The second one investigated the comprehensive performance in paternity assignment, which was computed by

$$acc2 = \frac{n_{EX} + n_{LOD}}{N_T}$$

where n_{LOD} is the number of trios in which the true sire was identified by assigning paternity to the putative sire with the highest LOD score among non-excluded putative sires. The third one, which was computed by

$$acc2_{LOD} = \frac{n_{LOD}}{N_T - n_{EX}}$$

investigated the accuracy of the paternity assignment based solely on LOD scores.

Simulations for all the parameter sets were run with 50,000 replicates; all the success rates, except for $acc2_{LOD}$, were averaged over all the replicates. Here, $acc2_{LOD}$ was averaged over all the replicates, except for the replicates in which sire identification succeeded in all the trios by only the exclusion method (i.e., except for the replicates in which $N_T = n_{EX}$).

RESULTS AND DISCUSSION

To investigate the realized SNP marker conditions in each subpopulation, expected heterozygosity (H_e), observed heterozygosity (H_o), and $F_{IS} = 1 - (H_o/H_e)$ of Nei (1977) were averaged over $n_l = 50$ loci in the parental populations of each sex and their offspring (Figure 1). Because almost the same values were obtained, results for $n_l = 10$ and 30 are not indicated in Figure 1. The differences of expected heterozygosity (H_e) within the true sires and dams were very small between the two subpopulations, indicating that the quality of the SNP marker sets available for the parentage tests was almost equivalent in each of them. Averaged F_{IS} were smaller than zero for all the cases suggesting heterozygote excess, but the deviations from zero were not so large. Therefore, marker genotype frequencies in each population are considered not to be greatly deviated from those expected under HWE. But, it should be noted that the two subpopulations are different from a randomly mating population under HWE from the

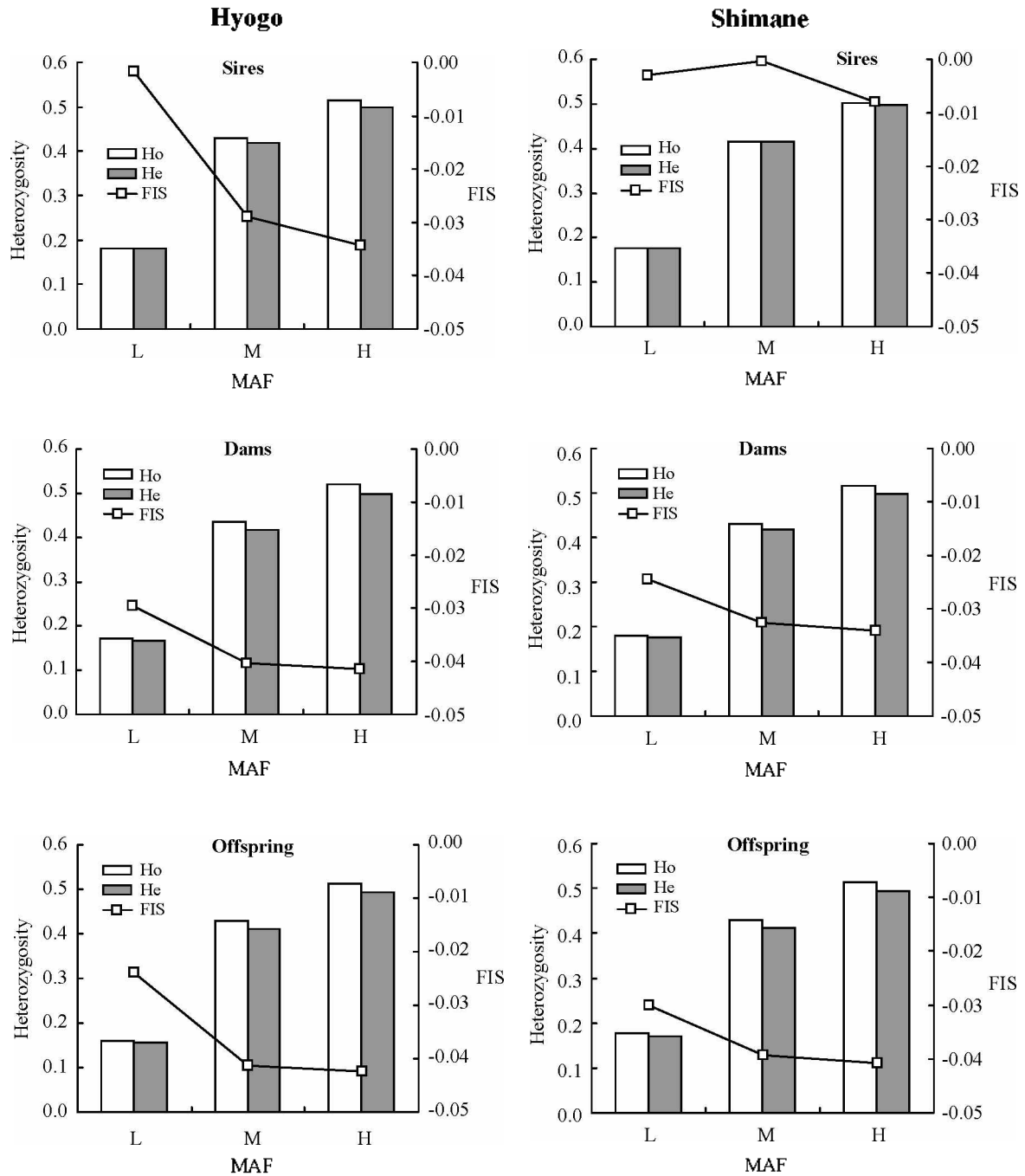


Figure 1. Mean of realized expected heterozygosity (H_e), observed heterozygosity (H_o), and F_{IS} of Nei (1977) averaged over $n_l = 50$ loci computed within parental populations of each sex and their offspring. The symbols L, M, and H in each figure represent the minor allele frequencies (MAF) in the parental populations of each sex: L = low ($0.05 \leq q \leq 0.15$), M = middle ($0.25 \leq q \leq 0.35$), and H = high ($0.45 \leq q \leq 0.5$). Standard errors were ≤ 0.00005 for H_e , ≤ 0.00005 for H_o , and ≤ 0.00011 for F_{IS} .

perspective of an association between mating frequencies of parents (especially those of sires) and their marker genotypes.

The success rates of excluding a single relative of the true sire ($acc1_{FS}$ for full-sib and $acc1_{FATH}$ for sire) computed for various combinations of the number of SNP loci (n_l) and MAF (q) are presented in Table 2. For a given number of

marker loci (n_l), an increase of MAF inflated both success rates, but the differences between the results for intermediate MAF and those for high MAF were not so large. For example, when $n_l = 50$, an increase of MAF from $0.05 \leq q \leq 0.15$ to $0.25 \leq q \leq 0.35$ inflates $acc1_{FS}$ by 0.179 and $acc1_{FATH}$ by 0.1232 in the subpopulation of Hyogo. However, corresponding increments are only 0.0042 and

Table 2. Mean success rates of paternity tests based on the exclusion method assuming that a single full-sib ($acc1_{FS}$) or sire ($acc1_{FATH}$) of the true sire was a putative sire, in the two subpopulations of Hyogo and Shimane prefectures, simulated with various conditions of SNP marker sets

MAF (q)	Hyogo						Shimane						
	$acc1_{FS}$			$acc1_{FATH}$			$acc1_{FS}$			$acc1_{FATH}$			
	L ¹	M ¹	H ¹	L	M	H	L	M	H	L	M	H	
10 loci	Mean	0.3017	0.6159	0.6663	0.3404	0.5801	0.6251	0.3106	0.5731	0.6270	0.3479	0.6022	0.6387
	SE	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003	0.0003
30 loci	Mean	0.6446	0.9429	0.9624	0.7036	0.9234	0.9438	0.6665	0.9207	0.9455	0.7131	0.9283	0.9422
	SE	0.0005	0.0001	0.0001	0.0004	0.0001	0.0001	0.0004	0.0002	0.0001	0.0004	0.0001	0.0001
50 loci	Mean	0.8125	0.9915	0.9957	0.8620	0.9852	0.9906	0.8340	0.9848	0.9915	0.8673	0.9852	0.9885
	SE	0.0004	0.0000	0.0000	0.0003	0.0000	0.0000	0.0003	0.0001	0.0000	0.0002	0.0000	0.0000

¹ The symbols L, M, and H represent the minor allele frequencies (MAF) in the parental populations of each sex:

L = low ($0.05 \leq q \leq 0.15$), M = middle ($0.25 \leq q \leq 0.35$), and H = high ($0.45 \leq q \leq 0.5$).

0.0054 when MAF increases from $0.25 \leq q \leq 0.35$ to $0.45 \leq q \leq 0.5$. This tendency can also be found in the theoretical results obtained by calculating the conventional exclusion probability of a random putative sire in a randomly mating population under HWE (Pr_{RND} ; cf. Dodds et al., 1996) (Table 3). A similar tendency was observed in the effect of n_i for a given MAF.

In the subpopulation of Hyogo, the success rates of $acc1_{FS}$ were consistently higher than those of $acc1_{FATH}$ when MAF was $0.25 \leq q \leq 0.35$ and $0.45 \leq q \leq 0.5$, while such a tendency could not be observed in the subpopulation of Shimane. Since the average inbreeding coefficient of hypothetical full-sibs (0.173) was higher than that of the sires of the true sires (0.133) in Hyogo, the full-sibs would form homozygote with higher frequency than the paternal grand sire. This would be one of the reasons for the tendency described above because SNPs were assumed to be biallelic so that putative sires could be excluded only when they formed homozygotes. No difference was observed between the corresponding average inbreeding coefficients in Shimane, both of which were 0.101.

Assuming a randomly mating population under HWE, Double et al. (1997) theoretically demonstrated that the exclusion probability of a full-sib of the true sire with a single biallelic locus (Pr_{FS}) is simply computed by $Pr_{FS} = Pr_{RND}/2$. The probabilities computed with corresponding n_i and MAF used in the simulation are presented in Table 3.

When MAF is intermediate or high (i.e., $0.25 \leq q \leq 0.35$ or $0.45 \leq q \leq 0.5$), $acc1_{FS}$ in Hyogo was higher than the corresponding exclusion probability (Pr_{FS}), and $acc1_{FATH}$ of Hyogo and both success rates in Shimane were almost the same as the corresponding Pr_{FS} . These results indicated that deviation from HWE and random mating in an actual population does not necessarily degrade the performance of excluding a close relative of the true sires. A strict parentage test based on the exclusion method seems to require at least 50 loci, each having MAF of 0.3 or above.

Two success rates ($acc2_{EX}$ and $acc2$) in the two subpopulations were simulated with various conditions of marker set and numbers of putative sires ($N_{FS}+1$, where N_{FS} is the number of full-sibs of the true sire) (Figure 2). For almost all the combinations of the parameters simulated, both success rates in Hyogo were higher than those in Shimane. The success rate differences between the two subpopulations were generally negligible; however, nontrivial differences were sometimes observed. For example, the largest difference in $acc2_{EX}$ was 0.1 for $n_i = 50$, $0.25 \leq q \leq 0.35$, and $N_{FS} = 50$, and the corresponding value in $acc2$ was 0.083 when $n_i = 30$, $0.25 \leq q \leq 0.35$, and $N_{FS} = 50$. One might expect that the accuracy in distinguishing the true sire from its close relatives using polymorphic DNA markers would be higher in the subpopulation of Shimane, which has a lower degree of relatedness between the individuals within the population (cf. Table 1). In general,

Table 3. Exclusion probability of a random sire (Pr_{RND}) and a full-sib of the true sire (Pr_{FS}) in a randomly mating population under Hardy-Weinberg equilibrium, computed for various conditions of SNP marker sets

MAF (q)	Pr_{RND} ²			Pr_{FS} ²		
	L ¹	M ¹	H ¹	L	M	H
10 loci	0.5745	0.8370	0.8746	0.3417	0.5793	0.6263
30 loci	0.9230	0.9957	0.9980	0.7147	0.9256	0.9478
50 loci	0.9861	0.9999	1.0000	0.8764	0.9868	0.9927

¹ The symbols L, M, and H represent the minor allele frequencies (MAF) in the population:

L = low ($q = 0.1$), M = middle ($q = 0.3$), and H = high ($q = 0.5$).

² Pr_{RND} was computed following Dodds et al. (1996), and Pr_{FS} was computed following Double et al. (1997).

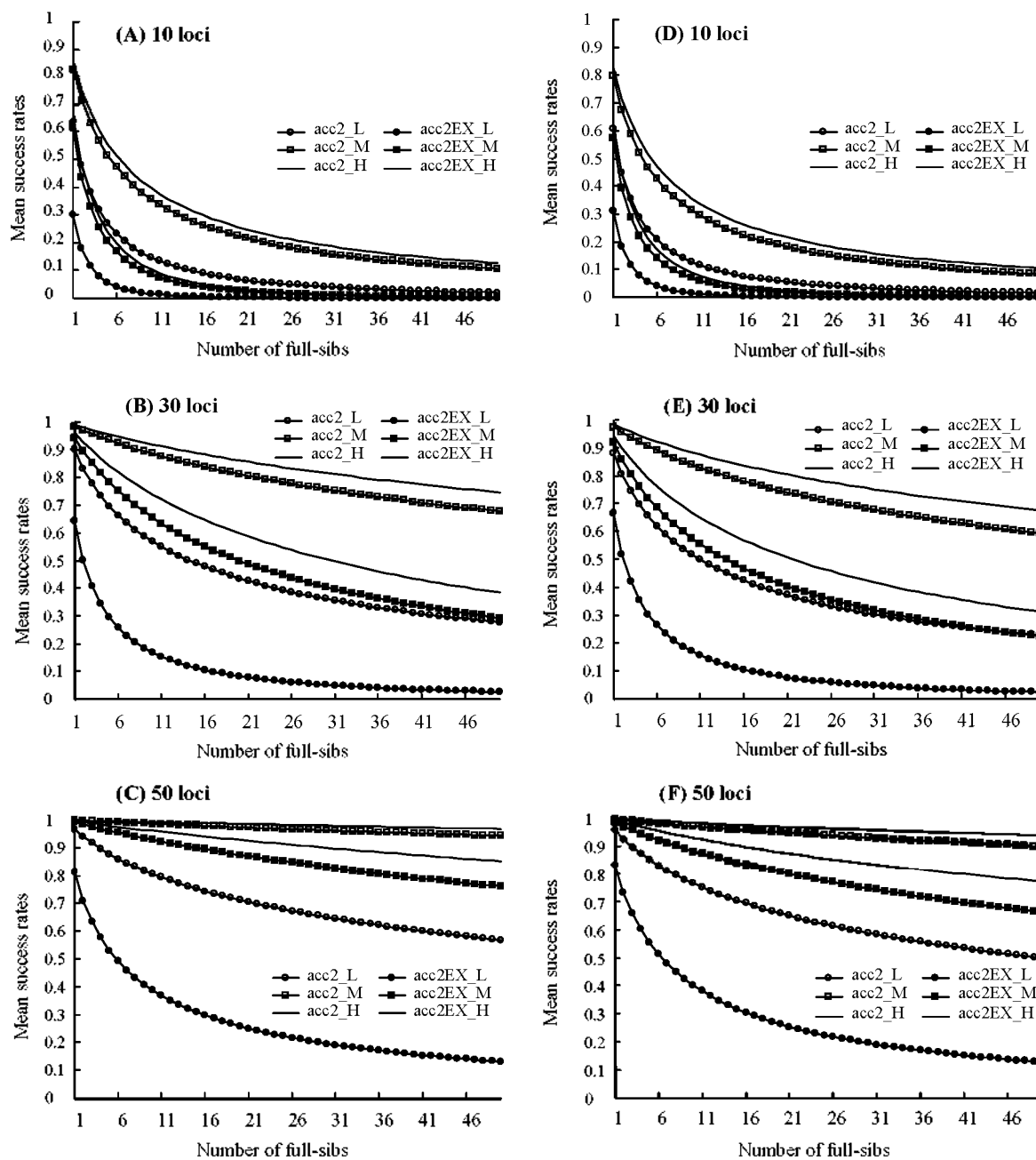


Figure 2. Mean success rates of sire identification tests solely by the exclusion method ($acc2_{EX}$) and those by the combined method of exclusion and the paternity assignment based on the LOD score ($acc2$), in the Hyogo ((A), (B), and (C)) and Shimane ((D), (E), and (F)) subpopulations, simulated with various conditions of SNP marker sets and the number of full-sib putative sires. The symbols L, M, and H in each figure represent the MAF in the parental populations of each sex: L = low ($0.05 \leq q \leq 0.15$), M = middle ($0.25 \leq q \leq 0.35$), and H = high ($0.45 \leq q \leq 0.5$). The standard error of $acc2_{EX}$ was ≤ 0.0005 and that of $acc2$ was ≤ 0.0006 for Hyogo. The standard error of $acc2_{EX}$ was ≤ 0.0005 and that of $acc2$ was ≤ 0.0004 for Shimane.

the degree of relatedness within a population, which is frequently expressed by average coancestry of the population, is interpreted as the expected proportion of the reduced expected heterozygosity (H_e) relative to the base population (Caballero and Toro, 2000; Frankham et al., 2002). Therefore, in a practical field, the difference in the degree of relatedness within each subpopulation might be

reflected in the amount of information contained in SNPs sampled, which is associated with H_e . In that sense, a parentage test might be more difficult in the subpopulation in Hyogo. However, simulation in the present study assumes that the presented marker sets had already been detected in each subpopulation, so that the higher degree of relatedness within Hyogo did not simply result in lower

accuracy. The accuracies of the different parentage tests in each subpopulation were dependent on the specific genotype patterns of the tested parents, which were affected by historical events of selection and mating, and on the current breeding structure in the parental generation.

Except for $n_l = 50$ and $0.25 \leq q \leq 0.35$ or $0.45 \leq q \leq 0.5$, the success rates of $acc2_{EX}$ in both subpopulations rapidly diminished with increasing N_{FS} , especially when N_{FS} was small. According to Chakraborty et al. (1998), in a randomly mating population under HWE, the probability of excluding $N-1$ random sires except the true sire of the offspring among N putative sires can be simply calculated as the exclusion probability of a random sire to the power of $N-1$ (i.e., Pr_{RND}^{N-1}). Computation of this probability with Pr_{RND} (Table 3) suggests that only $n_l = 30$ SNPs with $q = 0.3$ can exclude 10 random sires with a probability of 0.9578 in an idealized population. However, for both subpopulations in this study, success rates of $acc2_{EX} \approx 0.9$ or above were observed only when $n_l = 50$ and $0.25 \leq q \leq 0.35$ or $0.45 \leq q \leq 0.5$, indicating the difficulty in distinguishing the true sire from close relatives using only the exclusion method.

The success rates of paternity assignment in the two subpopulations solely by the likelihood-based method ($acc2_{LOD}$) are given in Table 4. Similar to the results obtained by simulation 1 (Table 2), increasing the MAF from low to intermediate values (i.e., from $0.05 \leq q \leq 0.15$ to $0.25 \leq q \leq 0.35$) is much more beneficial than increasing it from intermediate to high values (i.e., from $0.25 \leq q \leq 0.35$ to $0.45 \leq q \leq 0.5$). Similar results were obtained in the simulation study of Hill et al. (2008), in which sire identification tests using LOD scores were carried out assuming five full-sib putative sires.

For $n_l = 30$ or 50 and $0.25 \leq q \leq 0.35$ or $0.45 \leq q \leq 0.5$,

decreasing rates of $acc2_{LOD}$ with the increase of N_{FS} were gradual, while those for the other marker conditions were fairly rapid (Table 4). Even when $N_{FS} = 50$ full-sibs of the true sire competed, the tests making use of $n_l = 50$ loci with MAF of $0.25 \leq q \leq 0.35$ succeeded at a rate of 0.75 in the subpopulation of Hyogo and approximately 0.7 in the subpopulation of Shimane. Marshall et al. (1998) examined the accuracy of likelihood-based paternity tests using the actual allele frequency data at three protein and nine microsatellite loci sampled from the red deer population. One of their simulation results revealed that the reduction in test accuracy was not substantial even when the number of putative sires in the population varied from 2 to 101.

Differences in the marker conditions generated large differences in the comprehensive success rates of sire identification ($acc2$), which is the most important index in this simulation (see Figure 1). For all cases except $n_l = 50$ and $0.25 \leq q \leq 0.35$ or $0.45 \leq q \leq 0.5$, the success rates rapidly diminished as the number of putative sires (N_{FS}) increased. However, when $n_l = 50$ loci with $0.25 \leq q \leq 0.35$ or $0.45 \leq q \leq 0.5$, the decrease in success rate was much more gradual, with $acc2$ of 0.9430 and 0.9681 in the subpopulation of Hyogo, and 0.8999 and 0.9399 in Shimane, even for $N_{FS} = 50$. Therefore, good performance in the sire identification tests may be expected with the detection of more than $n_l = 50$ SNPs having MAF larger than 0.3, and the introduction of the combined method of the exclusion and the paternity assignment based on the LOD score. To realize high performance in each prefectural subpopulation unit, SNPs having a large amount of information should be selected in each of them.

In livestock species, a large number of breeds or subpopulations under selection are expected to deviate from an idealized randomly mating population under HWE. In

Table 4. Mean success rates¹ of the sire identification tests based on the LOD score ($acc2_{LOD}$), in the two subpopulations of Hyogo and Shimane prefectures, simulated with various conditions of SNP marker sets and the number of full-sib putative sires (N_{FS})

	N_{FS}	Hyogo						Shimane					
		1	10	20	30	40	50	1	10	20	30	40	50
MAF (q)													
10 loci	L ²	0.4677	0.1329	0.0658	0.0402	0.0275	0.0199	0.4232	0.1143	0.0553	0.0331	0.0223	0.0161
	M ²	0.5319	0.2965	0.2024	0.1530	0.1228	0.1023	0.5142	0.2624	0.1721	0.1270	0.0998	0.0820
	H ²	0.5357	0.3264	0.2312	0.1788	0.1458	0.1229	0.5219	0.2936	0.2014	0.1528	0.1227	0.1025
30 loci	L	0.6981	0.4793	0.3836	0.3267	0.2873	0.2582	0.6341	0.4171	0.3256	0.2717	0.2358	0.2090
	M	0.6982	0.6606	0.6194	0.5879	0.5620	0.5403	0.6776	0.6153	0.5653	0.5288	0.5000	0.4761
	H	0.6975	0.6795	0.6484	0.6234	0.6025	0.5843	0.6827	0.6425	0.6024	0.5722	0.5474	0.5266
50 loci	L	0.7895	0.6767	0.6095	0.5639	0.5296	0.5022	0.7213	0.6005	0.5319	0.4862	0.4519	0.4246
	M	0.7721	0.7895	0.7801	0.7709	0.7626	0.7551	0.7498	0.7519	0.7352	0.7217	0.7102	0.6998
	H	0.7761	0.7970	0.7931	0.7878	0.7828	0.7779	0.7553	0.7666	0.7566	0.7471	0.7387	0.7311

¹ Standard errors were ≤ 0.0013 for Hyogo and ≤ 0.0012 for Shimane.

² The symbols L, M, and H represent the minor allele frequencies (MAF) in the parental populations of each sex: L = low ($0.05 \leq q \leq 0.15$), M = middle ($0.25 \leq q \leq 0.35$), and H = high ($0.45 \leq q \leq 0.5$).

general, the variance of progeny number among parents should be larger than that expected under random mating, and together with the degree of relatedness between the mates, the difference of allele frequency between male and female parents generated by genetic drift affect genotype frequencies in the offspring population (Robertson, 1965; Wang, 1996). Results obtained in the present study illustrated that application of the reasonable SNP system to the parentage test will result in good performance, even in intensively selected livestock populations.

Recently, Hill et al. (2008) developed a novel likelihood ratio for SNP systems, especially dealing with the problem of genotyping errors. Furthermore, they proposed posterior probability, considering the expected mating frequencies of the animals, to accommodate the likelihood-based method to the populations in which a large variance of progeny number among parents is expected. Since one of our interests was the effect of deviation from random mating and HWE on the accuracy of parentage tests, the effect of genotyping errors was not considered to prevent it from masking the former effect. Together with the performance of posterior probability proposed by Hill et al. (2008), the effect of genotyping errors will be examined in a future study.

REFERENCES

- Anderson, E. C. and J. C. Garza. 2006. The power of single-nucleotide polymorphisms for large-scale parentage inference. *Genetics* 172:2567-2582.
- Baruch, E. and J. I. Weller. 2008. Estimation of the number of SNP genetic markers required for parentage verification. *Anim. Genet.* 39:474-479.
- Caballero, A. and M. A. Toro. 2000. Interrelations between effective population size and other pedigree tools for management of conserved populations. *Genet. Res. Camb.* 75: 331-343.
- Chakraborty, R., T. R. Meagher and P. E. Smouse. 1988. Parentage analysis with genetic markers in natural populations. I. The expected proportion of offspring with unambiguous paternity. *Genetics* 118:527-536.
- Dodds, K. G., M. L. Tate, J. C. McEwan and A. M. Crawford. 1996. Exclusion probabilities for pedigree testing farm animals. *Theor. Appl. Genet.* 92:966-975.
- Dodds, K. G., M. L. Tate and J. A. Sise. 2005. Genetic evaluation using parentage information from genetic markers. *J. Anim. Sci.* 83:2271-2279.
- Double, M. C., A. Cockburn, S. C. Barry and P. E. Smouse. 1997. Exclusion probabilities for single-locus paternity analysis when related males compete for matings. *Mol. Ecol.* 6:1155-1166.
- Eenennaam, A. L. V., R. L. Weaber, D. J. Drake, M. C. T. Penedo, R. L. Quaas, D. J. Garrick and E. J. Pollak. 2007. DNA-based paternity analysis and genetic evaluation in a large, commercial cattle ranch setting. *J. Anim. Sci.* 85:3159-3169.
- Frankham, R., J. D. Ballou and D. A. Briscoe. 2002. *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge.
- Gomez-Raya, L., K. Priest, W. M. Rauw, M. Okomo-Adhiambo, D. Thain, B. Bruce, A. Rink, R. Torell, L. Grellman, R. Narayanan and C. W. Beattie. 2008. The value of DNA paternity identification in beef cattle: Examples from Nevada's free-range ranches. *J. Anim. Sci.* 86:17-24.
- Heaton, M. P., G. P. Harhay, G. L. Bennett, R. T. Stone, W. M. Grosse, E. Casas, J. W. Keele, T. P. L. Smith, C. G. Chitko-McKown and W. W. Laegreid. 2002. Selection and use of SNP markers for animal identification and paternity analysis in US beef cattle. *Mamm. Genome* 13:272-281.
- Hill, W. G., B. A. Salisbury and A. J. Webb. 2008. Parentage identification using single nucleotide polymorphism genotypes: Application to product tracing. *J. Anim. Sci.* 86: 2508-2517.
- Honda, T., T. Nomura, Y. Yamaguchi and F. Mukai. 2002. Pedigree analysis of genetic subdivision in a population of Japanese Black cattle. *Anim. Sci. J.* 73:445-452.
- Jones, A. and W. R. Ardren. 2003. Methods of parentage analysis in natural populations. *Mol. Ecol.* 12:2511-2523.
- Kalinowski, S. T., M. L. Taper and T. C. Marshall. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* 16:1099-1106.
- MacCluer, J. W., J. L. VandeBerg, B. Read and O. A. Ryder. 1986. Pedigree analysis by computer simulation. *Zoo. Biol.* 5:147-160.
- Marshall, T. C., J. Slate, L. E. B. Kruuk and J. M. Pemberton. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* 7:639-655.
- Meagher, T. R. 1986. Analysis of paternity within a natural population of *Chamaelirium luteum*. I. Identification of most-likely male parents. *Am. Nat.* 128:199-215.
- Nei, M. 1977. *F*-statistics and analysis of gene diversity in subdivided populations. *Ann. Hum. Genet.* 41:225-233.
- Robertson, A. 1965. The interpretation of genotypic ratios in domestic animal populations. *Anim. Prod.* 7:319-324.
- Rohrer, G. A., B. A. Freking and D. Nonneman. 2007. Single nucleotide polymorphisms for pig identification and parentage exclusion. *Anim. Genet.* 38:253-258.
- Wang, J. 1996. Deviation from Hardy-Weinberg proportions in finite populations. *Genet. Res., Camb.* 68:249-257.
- Werner, F. A. O., G. Durstewitz, F. A. Habermann, G. Thaller, W. Krämer, S. Kollers, J. Buitkamp, M. Georges, G. Brem, J. Mosner and R. Fries. 2004. Detection and characterization of SNPs useful for identity control and parentage testing in major European dairy breeds. *Anim. Genet.* 35:44-49.