BioOne COMPLETE

Analysis of validated and population-specific single nucleotide polymorphism parentage panels in pedigreed and commercial beef cattle populations

Authors: Buchanan, Justin W., Woronuk, Grant N., Marquess, F. Leigh, Lang, Kevin, James, Steven T., et al.

Source: Canadian Journal of Animal Science, 97(2) : 231-240

Published By: Canadian Science Publishing

URL: https://doi.org/10.1139/cjas-2016-0143

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Downloaded From: https://bioone.org/journals/Canadian-Journal-of-Animal-Science on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

NRC Research Press

ARTICLE

Analysis of validated and population-specific single nucleotide polymorphism parentage panels in pedigreed and commercial beef cattle populations

Justin W. Buchanan, Grant N. Woronuk, F. Leigh Marquess, Kevin Lang, Steven T. James, Heather Deobald, Bryan T. Welly, and Alison L. Van Eenennaam

Abstract: A pedigreed population containing 71 calves and 8 sires was used to compare sire qualification using three genotyping platforms [14 microsatellite, real-time quantitative PCR, and 100, 200, 500, and 1000 single nucleotide polymorphism (SNP) arrays]. Parentage was also qualified in an unknown-pedigree population containing 8480 calves with 460 sires using SNP arrays. The three platforms qualified the true sire in the known-pedigree population with zero mismatches. The 100 and 200 SNP arrays yielded specificities of 0.92 and 0.99 with a 1% mismatch rate in the known-pedigree population, respectively. In the larger population, SNP panels of the 500 and 1000 highest minor allele frequency SNPs were also evaluated. The 1000 SNP panel qualified paternity to a single sire for 82.1% of calves with 1% or 2% mismatches. Not all commercial sires were genotyped, which accounts for missing paternity for some calves. In this larger population, the 100 SNP array qualified multiple sires to 0.42% of calves and single sires to 80.84% of calves without mismatches. The 200 SNP array assigned unique paternity, and 79.8% of calves were qualified to a sire without mismatches. With a 2% mismatch rate, sire qualifications agreed with the 1000 SNP array. This study highlights the interplay among population size, genotyping error rates, and the specificity and sensitivity of parentage platforms.

Key words: SNP parentage panel, sensitivity (parentage), specificity (parentage), Charolais, cattle (beef).

Résumé : Une population à pedigree contenant 71 veaux et 8 géniteurs mâles a été utilisée pour comparer la qualification des géniteurs mâles au moyen de trois plateformes de génotypage (14 microsatellites, PCR quantitative en temps réel, ainsi que des réseaux de 100, 200, 500, et 1000 SNPs). La parenté a aussi été qualifiée dans une population de pedigree inconnu contenant 8480 veaux et 460 géniteurs mâles au moyen des réseaux de SNPs. Les trois plateformes ont qualifié le véritable géniteur dans la population à pedigree connu avec aucun écart. Les réseaux à 100 et 200 SNPs ont rendu des spécificités de 0,92 et 0,99 avec un taux d'écart de 1 % dans la population à pedigree connu, respectivement. Dans la plus grande population, les panels de SNPs de 500 et 1000 SNPs de fréquence d'allèles mineurs les plus élevés ont aussi été évalués. Le panel de 1000 SNPs a qualifié la paternité à un seul géniteur pour 82,1 % des veaux avec seulement 1 % à 2 % d'écarts. Tous les géniteurs mâles commerciaux n'ont pas été génotypés ce qui explique la paternité manquante pour certains veaux. Dans cette population plus grande, le réseau à 100 SNPs a qualifié de multiples géniteurs à 0,42 % des veaux et un seul géniteur à 80,84 % des veaux sans écarts. Le réseau à 200 SNPs a assigné une paternité unique et 79,8 % des veaux ont été qualifiés à un géniteur sans écart. Avec un taux d'écart de 2 %, la qualification des géniteurs était en accord avec le réseau de 1000 SNPs. Cette étude souligne l'interdépendance entre la taille de la population, les taux d'écarts de génotypage ainsi que la spécificité et la sensibilité des plateformes de parenté. [Traduit par la Rédaction]

Mots-clés : panel de parenté SNP, sensibilité (parenté), spécificité (parenté), Charolais, bovins (boeuf).

Received 22 July 2016. Accepted 18 November 2016.

J.W. Buchanan, B.T. Welly, and A.L. Van Eenennaam. Department of Animal Science, University of California, Davis, CA 95616, USA.

G.N. Woronuk, F.L. Marquess, K. Lang, S.T. James, and H. Deobald. Quantum Genetix, 101 Research Drive, Saskatoon, SK S7N 3R3, Canada.

Corresponding author: Alison L. Van Eenennaam (email: alvaneenennaam@ucdavis.edu).

Copyright remains with the author(s) or their institution(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Downloaded From: https://bioone.org/journals/Canadian-Journal-of-Animal-Science on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Introduction

Parentage qualification in commercial beef cattle production is important for monitoring reproduction and estimating genetic merit. In large commercial beef cattle systems with multisire breeding pastures, parentage recording is often inaccurate or unfeasible, and incorrect parentage qualification can affect genetic gain (Munoz et al. 2014). The incorporation of genomic data into modern commercial beef cattle production systems creates the opportunity for reconstruction of a pedigree through parentage qualification based on single nucleotide polymorphism (SNP) markers (Tokarska et al. 2009; Fernandez et al. 2013). Although single-sire parentage qualifications from SNP genotypes can be accomplished by excluding potential sires when mismatching, or opposing homozygous, loci are identified (Hayes 2011), the presence of mismatching markers between true parent–offspring individuals is also a reality of array-based genotyping platforms due to known error rates (Hong et al. 2012). Multiple challenges related to array genotyping error rate and panel size need to be addressed when dealing with large populations, composite breeds, missing sires, and candidate sires with a high degree of relatedness.

The objective of this study was to evaluate the sensitivity and specificity of different DNA-based platforms and different-sized SNP panels, including the International Society for Animal Genetics (ISAG) panels, for qualifying parentage for both a small cattle population of known pedigree and a larger Charolais-sired commercial cattle population with unknown pedigree typical of field data sets. In addition, we evaluated two population-derived SNP panels composed of the 500 and 1000 SNP that had the highest minor allele frequency in the Charolais-sired commercial data set. This selection approach for population-derived SNP panels could be extended to other large populations with SNP array data that require more than 200 SNP markers to ensure both sensitivity and specificity in parentage assignments.

Materials and Methods

Cattle populations

A commercial cattle population and a smaller known-pedigree population were used as reference populations to assess parentage qualification using the opposing homozygous SNP comparison method. A known-pedigree population consisting of 71 calves and eight sires, including two half-sib sire pairs, was assembled based on veterinary records of artificial insemination. To assemble the commercial population, ear-punch tissue samples were collected from 8840 Charolais-sired calves and 460 Charolais bulls from seven separate ranch locations in Idaho and Washington over a period of 2 yr. Each sire potentially produced calves at multiple ranch locations over two breeding seasons which required all potential sire–calf comparisons to be tested. Not all potential sires could be genotyped due to the scale and organization of this operation. In addition, DNA samples were collected from all calves undergoing parentage assignment at weaning that allowed for potential sires to be differentiated from offspring.

Genotyping platforms

The known-pedigree population was genotyped for parentage markers using three different platforms. To call microsatellite genotypes, a panel of 14 fluorescently labelled short tandem repeat markers (12 ISAG standard markers plus two additional markers: *BM1818*, *BM1824*, *BM2113*, *CSSM036*,*ETH3*, *ETH10*, *ETH225*, *HEL1*, *INRA023*, *SPS115*, *TGLA53*, *TGLA122*, *TGLA126*, and *TGLA227*) were genotyped for all samples using standard PCR techniques. The raw data files were analyzed with the GeneMapper Software version 4.1 (Applied Biosystems, Foster, CA, USA), and the genotypes were assigned to sizing bins that have been adjusted to conform to the ISAG standards.

A 109 SNP subset of the ISAG SNP panel was selected for genotyping by real-time quantitative PCR (qPCR) for paternity use in the small known-pedigree population. The majority of genotypes were generated using hydrolysis probe chemistry, with the remaining generated using a bipartite probe system (KBioscience Ltd., Hoddesdon, UK). All qPCRs were conducted on a BioRad CFX Real-Time PCR Detection System, and data were analyzed using the BioRad CFX Manager version 3.1 (Bio-Rad Laboratories, Hercules, CA, USA) qPCR analysis software.

High-throughput SNP genotypes were obtained using the GeneSeek Genomic Profiler Low Density (GGP-LD) version 1.1 (20k) SNP array. The 100 SNP core and 200 SNP parentage panels from the ISAG Cattle Molecular Markers and Parentage Testing Committee were present. Parentage qualification was carried out by comparing the number of opposing homozygous, or mismatched, SNPs between each calf and all potential sires using the SEEKPARENTF90 software (Aguilar et al. 2014). Potential sires and calves were flagged within the program to facilitate qualification. Any marker with more than 5% missing SNP calls across the population was excluded from consideration from any parentage panel. Animals with more than 10% of their SNP markers missing from any parentage panel were removed from the parentage qualification process.

The commercial population was genotyped using the GeneSeek Genomic Profiler GGP_LD (20k, 26; versions 1.1–1.4) array for the calves and the GeneSeek Genomic Profiler GGP-HD version 1.9 (76k) SNP array for the potential sires. Two additional population-derived parentage panels of 500 and 1000 SNP selected from high minor allele frequencies specific to the Charolais-sired population used in this study were also evaluated. No more than 5% missing SNP calls were allowed for each SNP marker across the population, and sex chromosome

Downloaded From: https://bioone.org/journals/Canadian-Journal-of-Animal-Science on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Table 1. Overlap among the four single nucleotide polymorphism (SNP) arrays used in parentage qualification.

	ISAG 200	500 SNP	1000 SNP
ISAG 100	100	9	13
ISAG 200	—	17	28
500 SNP	—	—	500

Note: ISAG, International Society for Animal Genetics.

markers were eliminated prior to panel selection. The lowest minor allele frequency observed in the 500 SNP panel was 0.470 and 0.354 in the 1000 SNP panel. Overlap among the four marker panels is displayed in Table 1. There were 172 markers (Supplementary material)[1] from the ISAG 200 panel that were not included in the 500/1000 SNP panels due to low minor allele frequency.

Assessment of genotyping platforms

Microsatellite accuracy was assessed by ascertaining marker exclusion probabilities according to ISAG standards (http://www.isag.us/Docs/consignmentforms/Exclusion_probability.pdf). Microsatellite calls were assisted by manual intervention requiring technician oversight. When qualifying parentage, no exclusions were allowed between potential sire–calf pairs.

Accuracy of SNP calls between qPCR and SNP array genotypes was assessed by comparing the number of missing calls and the number of mismatched calls in the known-pedigree population for the ISAG 100 panel. The SNP calls were only compared when calls were not missing across both platforms. Accuracy was assessed as a percentage of matching SNP calls across all SNP calls for the 79 animals in the known-pedigree population.

Evaluation of SNP panel performance

Sensitivity and specificity were evaluated for the microsatellite marker, qPCR, and SNP array platforms in the known-pedigree population to assess each platform's ability to declare the single true sire correctly, as well as estimate the impact of allowing a range of mismatches. Sensitivity was defined as the ratio of the number of calves that were qualified to the correct sire to the total number of calves, and specificity was defined as the ratio of the number of calves with a single candidate sire (correct or incorrect) to the total number of calves. The four different-sized (100, 200, 500, and 1000) SNP arrays were also evaluated in the larger commercial population in which parentage was unknown. Performance was assessed in this population by reporting the agreement between panel results when allowing a range of mismatches. In addition, the agreement among results (single sire not excluded, more than one sire not excluded, or all sires excluded) of the different panels when allowing a single 1% mismatch rate was assessed.

Results

SNP call agreement between qPCR and SNP array

Because representative markers from the ISAG SNP panels were used in both arrays and qPCR in the known-pedigree population, agreement between SNP array calls and qPCR SNP calls was evaluated. One animal was removed from the analysis due to a SNP call rate (percentage of SNP calls) less than 0.90 for the set of ISAG markers. In total, 13 844 SNP markers from 78 samples were available. Accuracy of SNP calls for markers contained in the ISAG 100 SNP panel was evaluated across the two platforms. SNP call rates of 99.5% and 72.6% were observed in the SNP array and qPCR SNP calls, respectively, enabling 4763 SNP call comparisons across both platforms. The qPCR SNP call was the same as the SNP array call in 97.8% of all SNP calls, with 103 mismatching SNP calls across platforms. In no case did these mismatches change the paternity assignment. Previous evaluation of the accuracy of SNP array calls suggests that samples with a SNP array call rate greater than 0.90 have a genotyping accuracy above 0.99 (Cooper et al. 2013). The SNP call error rate for the SNP array was within the expected range, but this error needs to be accounted for in subsequent parentage analysis using SNP genotypes.

Panel performance

Sensitivity, specificity, and sire–calf qualifications produced by the microsatellite markers, qPCR, and ISAG 100/200 SNP array in the known-pedigree population as well as the commercial population-derived 500 and 1000 SNP arrays analyzed in the commercial populations are displayed in Table 2. Sensitivity and specificity of the platforms and panel sizes, respectively, are reported in the known-pedigree population by comparing sire–calf qualifications with the recorded pedigree. In this population, the ISAG 100 and 200 SNP panels had a sensitivity of 1.00 at 0%, 1%, and 2% mismatch rates, although the sensitivity and specificity of the qPCR and microsatellite markers were sufficient to correctly identify the true sires and exclude all incorrect sires with no mismatches allowed. When conducting this parentage study, three samples from the pedigreed population were mishandled, leading to an incorrect bull–calf match in the array laboratory which was identified after the true pedigrees were revealed. Subsequent resampling and retesting of the mishandled samples resulted in the correct sire–calf match.

[1]Supplementary material is available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/cjas-2016-0143.

Downloaded From: https://bioone.org/journals/Canadian-Journal-of-Animal-Science on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Table 2. Sire qualification results from microsatellite markers, International Society for Animal Genetics (ISAG) single nucleotide polymorphism (SNP) arrays, and population-derived SNP arrays in the known-pedigree of 70 calves and eight sires and commercial population of 8480 calves and 460 potential sires at varying levels of allowed mismatch rates.

	Known-pedigree population		Commercial population		
Mismatch rate	Sensitivity	Specificity	Single-sire qualified[a] (%)	No sire qualified (%)	>1 sire qualified (%)
Microsatellite PCR					
0%	1.00	1.00	na	na	na
SNP qPCR					
0%	1.00	1.00	na	na	na
ISAG 100 SNP array					
0%	1.00	1.00	80.84	18.74	0.42
1%	1.00	0.92	77.76	17.33	4.91
2%	1.00	0.87	56.19	14.46	29.35
ISAG 200 SNP array					
0%	1.00	1.00	79.81	20.19	0.00
1%	1.00	0.99	81.86	18.12	0.01
1.5%	1.00	0.97	81.84	18.05	0.11
2%	1.00	0.93	82.10	17.51	0.39
2.5%	1.00	0.87	80.92	17.62	1.39
500 SNP array					
0%	1.00	1.00	74.85	25.15	0.00
1%	1.00	1.00	82.08	17.92	0.00
1.5%	1.00	1.00	82.08	17.92	0.00
2%	1.00	1.00	82.09	17.89	0.02
2.5%	1.00	0.97	82.12	17.81	0.07
1000 SNP array					
0%	1.00	1.00	72.18	27.82	0.00
1%	1.00	1.00	82.12	17.88	0.00
1.5%	1.00	1.00	82.12	17.88	0.00
2%	1.00	1.00	82.12	17.88	0.00
2.5%	1.00	0.99	82.12	17.88	0.00

Note: qPCR, real-time quantitative PCR; na, not available.
[a]Percentage of 8480 calves qualified to a single sire.

In the larger commercial population, pedigree was not recorded so SNP array performance of each panel size was assessed by reporting the percentage of qualifications to a single sire, percentage of qualifications to multiple sires, and percentage of failed qualifications observed when allowing a range of mismatches. Individuals were identified either as a calf or as a potential sire, and all possible parentage scenarios were tested between each of the 8840 calves and 460 sires.

Specificity of ISAG 100 SNP panel at different mismatch rates

When zero mismatches were allowed, the ISAG 100 SNP array qualified parentage correctly for all 70 sire–calf pairs with no false-negative qualifications. However, allowing for a 1% mismatch rate resulted in failing to exclude nonsire bulls in six cases, reducing the specificity of the array panel to 0.92 (Fig. 1). Increasing the allowed mismatch rate to 2% resulted in 11 such failures to exclude nonsire bulls, resulting in a specificity of 0.87.

Allowing zero mismatches in the commercial population qualified sires to 80.84% of calves, with 0.42% of calves qualified to more than one candidate sire. A significant proportion of calves were not qualified to a sire because some proportion of the commercial bulls in this operation were not genotyped. Allowing 1% and 2% mismatch rates increased the multiple candidate sire rate to 4.91% and 29.35%, respectively, thereby concomitantly decreasing the percentage of calves qualified to a single sire to 77.76% and 56.19%, respectively. Results from the ISAG 100 SNP array in the larger commercial population at the 1% mismatch rate are displayed in Fig. 2.

Specificity of ISAG 200 SNP panel at different mismatch rates

Similar to the ISAG 100 SNP array, the 200 SNP array called correct single-sire matches with zero SNP

Downloaded From: https://bioone.org/journals/Canadian-Journal-of-Animal-Science on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Fig. 1. Qualifications observed for eight candidate sires using the International Society for Animal Genetics (ISAG) 100 single nucleotide polymorphism (SNP) array in the known-pedigree population. Calves with >1 sire qualified noted at a 1% mismatch rate allowed.

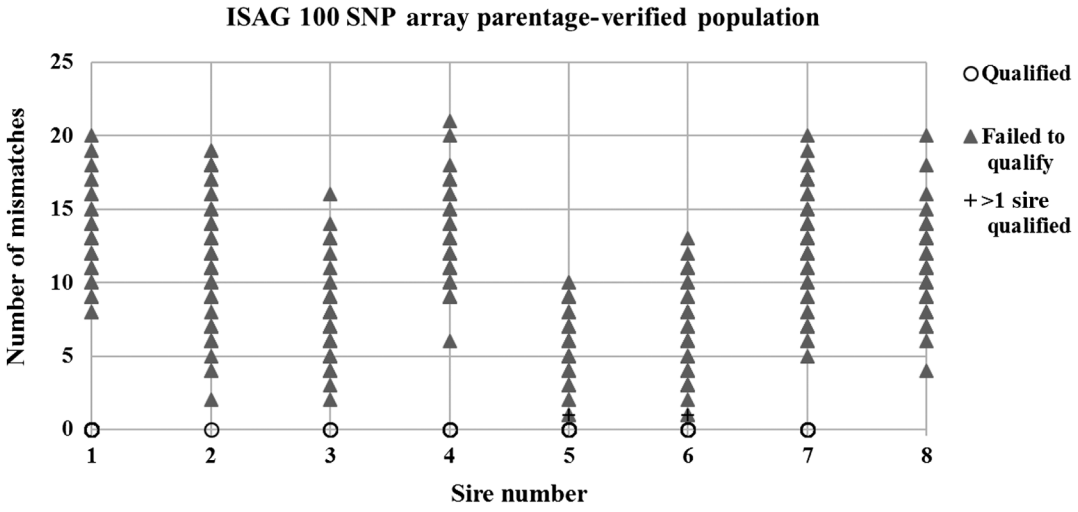
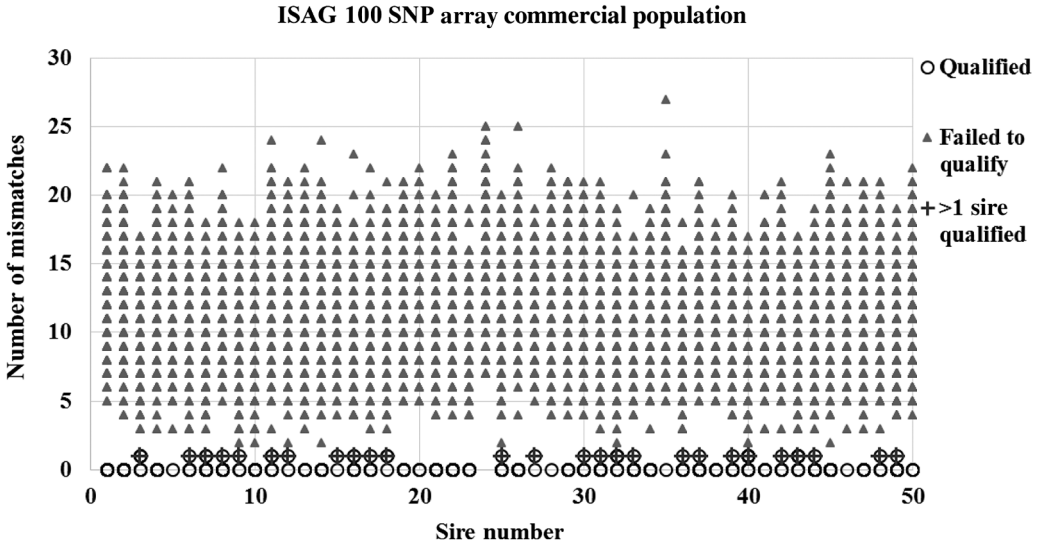


Fig. 2. Qualifications observed for 50 randomly chosen sires of the 460 candidate sires using the International Society for Animal Genetics (ISAG) 100 single nucleotide polymorphism (SNP) array in the commercial population. Calves with >1 sire qualified noted at a 1% mismatch rate allowed.



mismatches in the known-pedigree population. Allowing a 1% mismatch rate resulted in a case where a calf was qualified to two half-sib sires, with the incorrect sire having two allowed matches with the calf, decreasing the specificity to 0.99 (Fig. 3). Allowing a 2% mismatch rate increased the number of multiple qualified candidate sires to five, resulting in a specificity of 0.87.

Applying the ISAG 200 SNP panel to the commercial population resulted in 79.81% of calves qualified to a single sire, with no multiple sire qualifications, at zero mismatches. Allowing a 1% mismatch rate increased single-sire qualifications to 81.86%, but multiple candidate sire qualifications increased to 0.01% (Fig. 4). Increasing the mismatch rate to 2% with this SNP array increased the rate of multiple sires being qualified to a single calf to 0.39%, whereas the single-sire qualification rate increased to 82.12%.

Population-derived SNP panel performance

Results from the 1000 SNP arrays applied to both the known-pedigree (Fig. 5) and commercial populations (Fig. 6) at the 1% mismatch rate show a clear

Downloaded From: https://bioone.org/journals/Canadian-Journal-of-Animal-Science on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Fig. 3. Qualifications observed for eight candidate sires using the International Society for Animal Genetics (ISAG) 200 single nucleotide polymorphism (SNP) array in the known-pedigree population. Calves with >1 sire qualified noted at a 1% mismatch rate allowed.

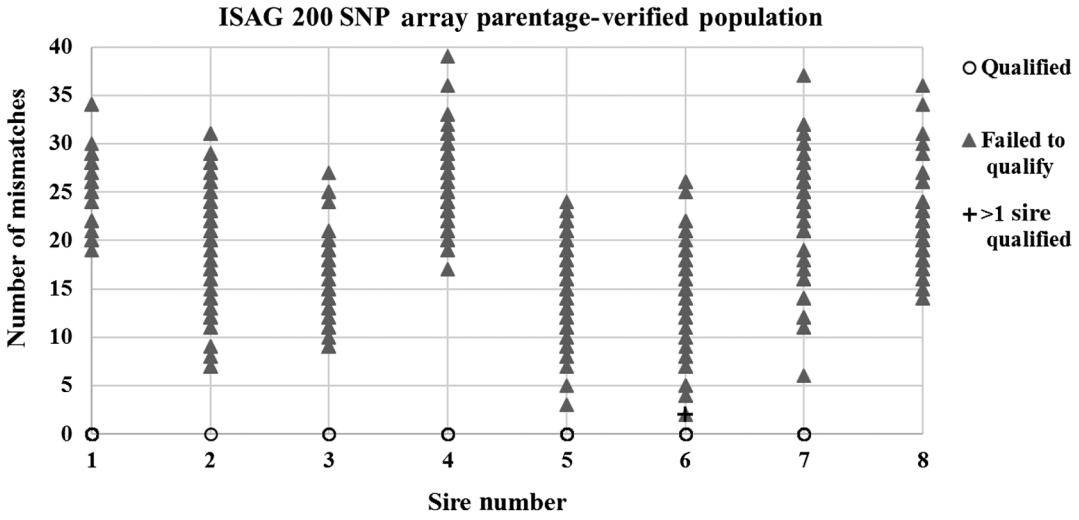
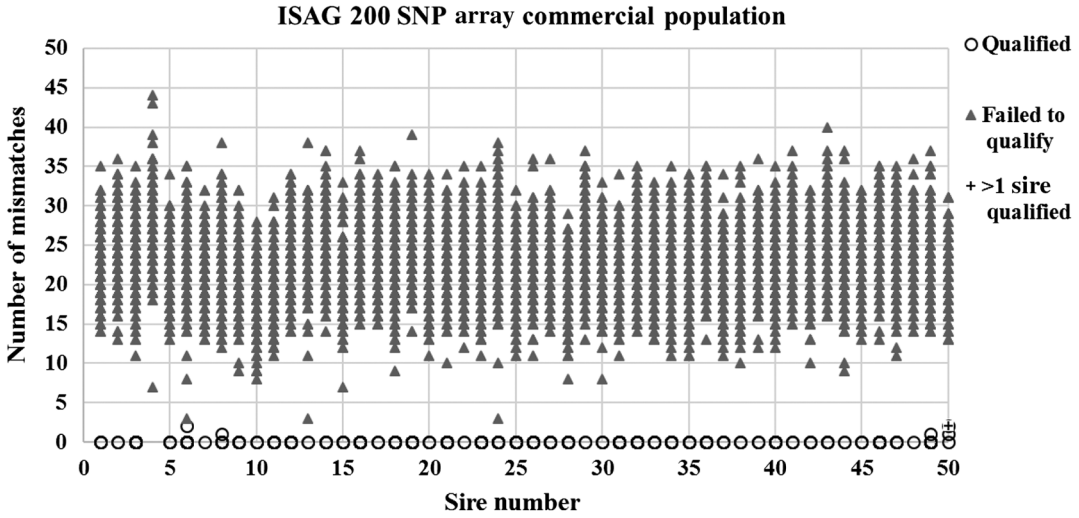


Fig. 4. Qualifications observed for 50 randomly chosen sires of 460 candidate sires using the International Society for Animal Genetics (ISAG) 200 single nucleotide polymorphism (SNP) array in the commercial population. Calves with >1 sire qualified noted at a 1% mismatch rate allowed.



differentiation between qualifying sires and nonqualifying sires. In contrast to the smaller ISAG panels, multiple sires qualifying to a single calf were not observed in the known-pedigree population until a 2.5% mismatch rate was allowed for both the 500 and 1000 SNP arrays. In the commercial population, using the 0% mismatch rate with the 500 and 1000 panels increased the number of calves with no sire assigned by approximately 10% relative to the assignments observed when allowing a 1% mismatch rate with these larger SNP panels. Presumably genotyping error results in a low but unknown number of apparent mismatches between true sire–calf pairs. Multiple sire qualifications started occurring at the 2% mismatch rates with the 500 SNP array, but this was not seen with the 1000 SNP array even at the 2.5% mismatch rate (Table 2).

Agreement of qualification across panels and platforms

Agreement of parentage qualification across specific panels and platforms was also assessed. In the known-pedigree population, the correct parentage matches were obtained by microsatellite, ISAG 100 SNP qPCR, and ISAG 100 and 200 SNP arrays when allowing zero mismatches. However, allowing even a 1% mismatch rate in these platforms confounded the pedigree with

Downloaded From: https://bioone.org/journals/Canadian-Journal-of-Animal-Science on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Fig. 5. Qualifications observed for eight candidate sires using the 1000 single nucleotide polymorphism (SNP) array selected on high minor allele frequency and call rate (MAF/CR) in the known-pedigree population.

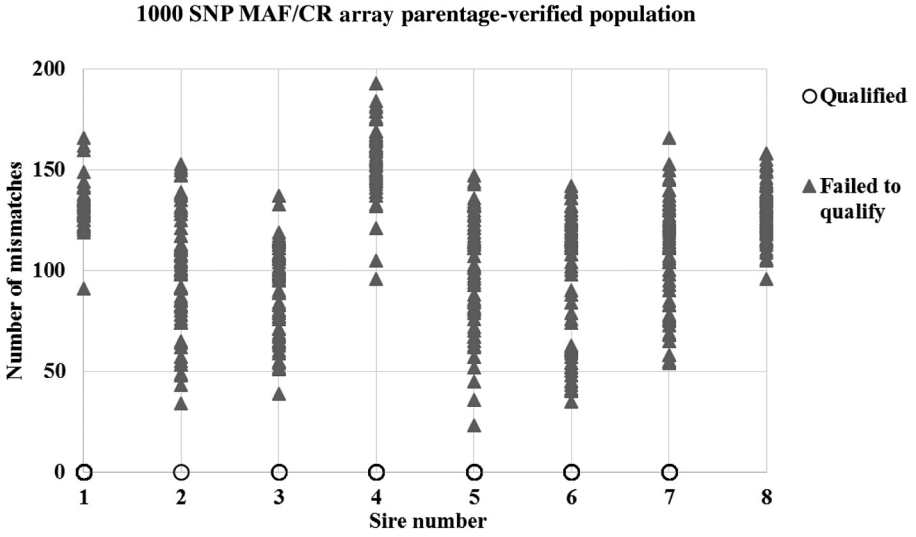
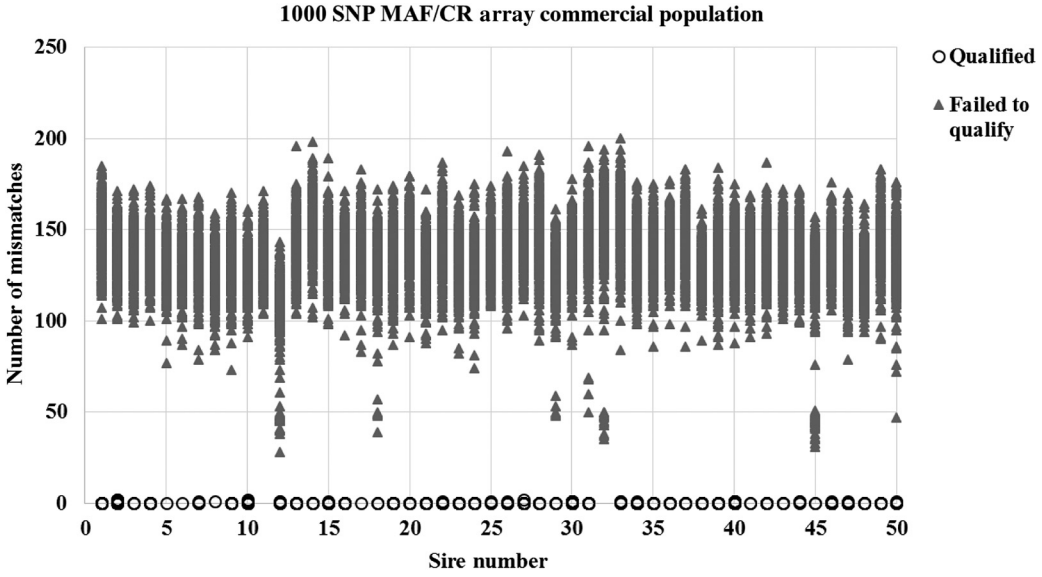


Fig. 6. Qualifications observed for 50 randomly chosen sires of the 460 candidate sires using the 1000 single nucleotide polymorphism (SNP) array selected on high minor allele frequency and call rate (MAF/CR) in the commercial population.



multiple sire qualifications. With the 500 and 1000 SNP arrays, correct parentage was observed when allowing up to a 2% mismatch rate in the pedigree population.

Qualified pedigree agreement at a 1% mismatch rate for commercial calves across the four parentage SNP arrays is displayed in Table 3. Agreement between the ISAG 100 and 200 SNP arrays for qualifications including single sire, multiple sires, and no sire present was 94.7% in the commercial population, with the highest agreement observed between the 500 and 1000 SNP arrays at 99.9%. The ISAG 200 SNP array had 99.7% overall agreement with the 1000 SNP panel. When comparing single-sire qualification, the ISAG 200 SNP array was in 100% agreement with the qualifications made by the 1000 SNP panel. When comparing single-sire qualifications, agreement of the sire qualified across the four panels was consistently greater or equal to 99.9% in the commercial population.

The genetic relationship between candidate sires in a parentage test can affect the ability of a set of markers to distinguish between true and false sire–calf relationships. Figure 1 shows that there is very little separation between the qualified and disqualified calves from related sires 5 and 6, which have an estimated genomic relationship of 0.28. When allowing a 1% mismatch rate with the ISAG 100 SNP panel, it is impossible to

Downloaded From: https://bioone.org/journals/Canadian-Journal-of-Animal-Science on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Table 3. Qualified pedigree agreement among all qualification (single sire, >1 sire, and no sire qualified) results (above the diagonal) and single-sire qualification (below the diagonal) at a 1% allowed mismatch rate in a commercial cattle population of 8480 calves and 460 potential sires.

	ISAG 100	ISAG 200	500 SNP	1000 SNP
ISAG 100	100	94.7	94.5	94.6
ISAG 200	99.9	100	99.7	99.7
500 SNP	99.9	99.9	100	99.9
1000 SNP	99.9	100	100	100

Note: ISAG, International Society for Animal Genetics; SNP, single nucleotide polymorphisms.

distinguish the correct sire–calf relationship in four cases between sires 5 and 6. This is in contrast to sires 1 and 4 which have a genomic relationship near zero and had easily defined sire–calf qualifications with all panels. Figure 5 demonstrates that as markers are added to the parentage test, it becomes easier to discern true parentage between even closely related sires.

Discussion

The overall goal of conducting parentage testing is to correctly identify one true sire from a group of candidate bulls given the true sire is present and to assign no sire if the true sire is absent. To attain this goal, researchers have developed several genetic marker panels, the most notable of which was developed by the ISAG cattle parentage testing committee. Evaluation of marker sets and genotyping platforms to be used as a parentage panel should include measures of both sensitivity and specificity when the true pedigree is known (Pu and Linacre 2008). When the number of allowed mismatching calls is increased, there is a risk of qualifying an incorrect sire, thereby reducing the sensitivity of the platform (Weller et al. 2004; Hong et al. 2012). Likewise, when too few markers are used, the specificity is decreased because calves may have their true sire fail to be uniquely qualified by the panel as larger sets of markers are needed to qualify parentage in populations with many potential sires (McClure et al. 2015; Strucken et al. 2015). Although simply removing mismatching calls has been discussed as a strategy to address genotyping error in parentage platforms (Hayes 2011), the impacts of such a practice on correctly assigning sire–calf parentage in nature are unknown.

It is generally agreed that PCR-based methods, such as hydrolysis probe qPCR and microsatellites (short tandem repeats), are among the most sensitive DNA assessment tools. However, relative to array-based platforms, conventional PCR-based methods have a lower practical limit in the number of markers assayed, and considerable investment is required to boost PCR capacity to that of common arrays. This is important because the number of markers needed to qualify the true candidate sire may be insufficient due to rate of throughput.

Microsatellite genotyping is a PCR-based method that has been extensively used in parentage qualification (Tian et al. 2008; Carolino et al. 2009; Radko and Slota 2009); it requires a dedicated platform and skilled interpretation for accuracy (McClure et al. 2013; Berry et al. 2014). The polymorphic nature of microsatellite markers facilitates parentage qualification for moderately sized populations; however, larger populations require proportionally larger marker sets. Similarly, although PCR provides an accurate method for genotyping SNP parentage markers (Clarke et al. 2014), investment in large numbers of markers, as well as application of automation where possible, is required to assess larger populations.

The ISAG SNP parentage panels of 100 and 200 markers have been validated across multiple breeds of beef cattle (http://www.isag.us/Docs/Cattle-SNP-ISAG-core-additional-panel-2013.xlsx). Recent qualification attempts with these 200 SNPs using an array platform found that this panel contained too few markers to produce single-sire qualifications in larger populations (Strucken et al. 2014, 2015). These results suggest that parentage platforms are needed that exhibit both specificity and sensitivity across large, genetically diverse populations with both missing and potentially related sires.

The population-derived 500 and 1000 SNP marker panels obtained as a byproduct of SNP array genotyping described in this paper provided unambiguous results as to which calves had missing sires and single-sire assignments when allowing for a 1% mismatch rate in the large commercial ranch data set. It is likely that a different set of markers would be selected when analyzing a different field population and selected panels would have limited utility beyond the breed makeup of the population from which they were derived. In addition, proposed methodology for selecting population-derived marker sets has included preselection against markers with a high frequency of mismatches in cases of true paternity (Weller et al. 2010), but this also requires a validation population. Such population-derived SNP sets may provide a powerful parentage tool for researchers working with SNP array data on field data sets with a large number of potential or missing sires. In this particular commercial application, specificity could have been improved if more information about the population structure was available allowing the potential sire pool to be narrowed down and partitioned for each particular calf crop. Alternative methods, including statistical likelihood-based procedures, are also able to qualify parentage but were not explored in these data. Likelihood procedures offer an alternative methodology to this problem but still require extensive calibration for allele and genotype frequencies as well as

Downloaded From: https://bioone.org/journals/Canadian-Journal-of-Animal-Science on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

population structure and the number of possible sires (Dodds et al. 2005; Hill et al. 2008). Further, principal component analysis of the SNP data can be used to identify half-sib or full-sib groups even in the absence of parental information.

In addition to accuracy of the methodology, cost is another important consideration when performing parentage qualification in cattle populations. Standalone parentage qualification from leading genotyping companies ranges from $13 to $19 per sample, depending on volume for tests that likely use a variation of the ISAG panels. Parentage qualification from SNP array data is more appropriately utilized as an additional feature when performing genomic selection. The cost of conducting qPCR-based parentage is becoming more affordable as reagent costs continue to decline, whereas ongoing advances in laboratory automation and SNP discovery are contributing to expanded laboratory capacities, making rapid and accurate parentage declarations increasingly achievable, even when assessing larger herd populations.

Required accuracy of the pedigree needs consideration. In commercial ranch applications, terminal sire–calf parentage accuracy has a higher tolerance for error than in applications such as registered purebred parentage. In production scenarios where a low level of error may be tolerable for sire–calf parentage, these results have demonstrated that the ISAG 100 SNP panel, when assessed by a sensitive platform, would be an economical alternative if thousands of SNP genotypes are not available for parentage. However, if the cost of genotyping continues to decline and thousands of SNPs are available for every animal, it is likely that parentage qualification methodologies, including both opposing homozygous and statistical likelihood procedures, will continue to focus on better utilization of the data from high-throughput SNP genotyping platforms.

Conclusions

In conducting parentage analyses, one must carefully consider platform sensitivity and panel specificity to ensure true sire–calf relationships are determined. Given the chain reaction nature of qPCR-based platforms, microsatellites and SNP–qPCR markers are so sensitive that even a one-marker mismatch is sufficient to confidently disqualify a bull regardless of panel size. However, specificity can be a limiting factor if the qPCR marker panel is of insufficient size, which is increasingly important as herd population size increases. However, array platforms can easily assess thousands of markers in a given run, yet suffer from intrinsic error rates that compromise platform sensitivity and any resulting parentage interpretation. One approach to overcome these intrinsic errors is to increase panel size and allow mismatches. Assessing sensitivity and specificity of alternative methodologies such as statistical likelihood procedures would be another alternative to account for intrinsic error rate and possibly improve the accuracy of parentage qualification. Previous research has focused on identifying a static set of markers to utilize as a standard parentage panel in cattle. However, the results of this study emphasize that static parentage panels are not required, and population-specific marker sets can be developed using the described methods. These results indicate that parentage validation from SNP markers is a process that requires careful consideration of marker selection, genotyping platform, and intrinsic error rates to ensure accuracy.

Acknowledgements

We thank Quantum Genetix for supplying the known-pedigree population and the qPCR and Microsatellite data preparation. This project was partially supported by funds from the USDA Agriculture and Food Research Initiative Competitive Grant No. 2011-68004-30367 from the USDA National Institute of Food and Agriculture.

References

Aguilar, I., Misztal, I., Tsuruta, S., Legarra, A., and Wang, H. 2014. PREGSF90–POSTGSF90: computational tools for the implementation of single-step genomic selection and genome-wide association with ungenotyped individuals in BLUPF90 programs. Paper 680 in Proc. 10th World Congress on Genetics Applied in Livestock Production, Vancouver, BC, Canada, 17–22 Aug. 2014. American Society of Animal Science. [Online] Available from https://asas.org/docs/default-source/wcgalp-posters/680_paper_9756_manuscript_1666_0.pdf?sfvrsn=2.

Berry, D.P., McClure, M.C., and Mullen, M.P. 2014. Within- and across-breed imputation of high-density genotypes in dairy and beef cattle from medium- and low-density genotypes. J. Anim. Breed. Genet. **131**: 165–172. doi:10.1111/jbg.12067. PMID:24906026.

Carolino, I., Sousa, C.O., Ferreira, S., Carolino, N., Silva, F.S., and Gama, L.T. 2009. Implementation of a parentage control system in Portuguese beef-cattle with a panel of microsatellite markers. Genet. Mol. Biol. **32**: 306–311. doi:10.1590/S1415-47572009005000026.

Clarke, S.M., Henry, H.M., Dodds, K.G., Jowett, T.W., Manley, T.R., Anderson, R.M., and McEwan, J.C. 2014. A high throughput single nucleotide polymorphism multiplex assay for parentage assignment in New Zealand sheep. PLoS ONE, **9**: e93392. doi:10.1371/journal.pone.0093392. PMID:24740141.

Cooper, T.A., Wiggans, G.R., and VanRaden, P.M. 2013. Short communication: relationship of call rate and accuracy of single nucleotide polymorphism genotypes in dairy cattle. J. Dairy Sci. **96**: 3336–3339. doi:10.3168/jds.2012-6208. PMID:23497995.

Dodds, K.G., Tate, M.L., and Sise, J.A. 2005. Genetic evaluation using parentage information from genetic markers. J. Anim. Sci. **83**: 2271–2279. doi:10.2527/2005.83102271x. PMID:16160036.

Fernandez, M.E., Goszczynski, D.E., Liron, J.P., Villegas-Castagnasso, E.E., Carino, M.H., Ripoli, M.V., Rogberg-Munoz, A., Posik, D.M., Peral-Garcia, P., and Giovambattista, G. 2013. Comparison of the effectiveness of microsatellites and SNP panels for genetic identification, traceability and assessment of parentage in an inbred Angus herd. Genet. Mol. Biol. **36**: 185–191. doi:10.1590/S1415-47572013000200008. PMID:23885200.

Downloaded From: https://bioone.org/journals/Canadian-Journal-of-Animal-Science on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Hayes, B.J. 2011. Efficient parentage assignment and pedigree reconstruction with dense single nucleotide polymorphism data. J. Dairy Sci. **94**: 2114–2117. doi:10.3168/jds.2010-3896.

Hill, W.G., Salisbury, B.A., and Webb, A.J. 2008. Parentage identification using single nucleotide polymorphism genotypes: application to product tracing. J. Anim. Sci. **86**: 2508–2517. doi:10.2527/jas.2007-0276. PMID:18502887.

Hong, H., Xu, L., Liu, J., Jones, W.D., Su, Z., Ning, B., Perkins, R., Ge, W., Miclaus, K., Zhang, L., Park, K., Green, B., Han, T., Fang, H., Lambert, C.G., Vega, S.C., Lin, S.M., Jafari, N., Czika, W., Wolfinger, R.D., Goodsaid, F., Tong, W., and Shi, L. 2012. Technical reproducibility of genotyping SNP arrays used in genome-wide association studies. PLoS ONE, **7**: e44483. doi:10.1371/journal.pone.0044483. PMID:22970228.

McClure, M.C., McCarthy, J., Flynn, P., Weld, R., Keane, M., O'Connel, K., Mullen, M.P., Waters, S., and Kearney, J.F. 2015. SNP selection for nationwide parentage verification and identification in beef and dairy cattle. ICAR Tech. Ser. No. 19. pp. 175–181. International Committee for Animal Recording (ICAR), Rome, Italy.

McClure, M.C., Sonstegard, T.S., Wiggans, G.R., Van Eenennaam, A.L., Weber, K.L., Penedo, C.T., Berry, D.P., Flynn, J., Garcia, J.F., Carmo, A.S., Regitano, L.C., Albuquerque, M., Silva, M.V., Machado, M.A., Coffey, M., Moore, K., Boscher, M.Y., Genestout, L., Mazza, R., Taylor, J.F., Schnabel, R.D., Simpson, B., Marques, E., McEwan, J.C., Cromie, A., Coutinho, L.L., Kuehn, L.A., Keele, J.W., Piper, E.K., Cook, J., Williams, R., and Van Tassell, C.P. 2013. Imputation of microsatellite alleles from dense SNP genotypes for parentage verification across multiple *Bos taurus* and *Bos indicus* breeds. Front. Genet. **4**: 176. doi:10.3389/fgene.2013.00176. PMID:24065982.

Munoz, P.R., Resende, M.F.R., Huber, D.A., Quesada, T., Resende, M.D.V., Neale, D.B., Wegrzyn, J.L., Kirst, M., and Peter, G.F. 2014. Genomic relationship matrix for correcting pedigree errors in breeding populations: impact on genetic parameters and genomic selection accuracy. Crop Sci. **54**: 1115–1123. doi:10.2135/cropsci2012.12.0673.

Pu, C.E., and Linacre, A. 2008. Systematic evaluation of sensitivity and specificity of sibship determination by using 15 STR loci. J. Forensic Leg. Med. **15**: 329–334. doi:10.1016/j.jflm.2007.12.018. PMID:18511010.

Radko, A., and Slota, E. 2009. Application of 19 microsatellite DNA markers for parentage control in Borzoi dogs. Pol. J. Vet. Sci. **12**: 113–117. PMID:19459448.

Strucken, E.M., Gudex, B., Ferdosi, M.H., Lee, H.K., Song, K.D., Gibson, J.P., Kelly, M., Piper, E.K., Porto-Neto, L.R., Lee, S.H., and Gondro, C. 2014. Performance of different SNP panels for parentage testing in two East Asian cattle breeds. Anim. Genet. **45**: 572–575. doi:10.1111/age.12154. PMID:24730981.

Strucken, E.M., Lee, S.H., Lee, H.K., Song, K.D., Gibson, J.P., and Gondro, C. 2015. How many markers are enough? Factors influencing parentage testing in different livestock populations. J. Anim. Breed. Genet. **133**: 13–23. doi:10.1111/jbg.12179. PMID:26234440.

Tian, F., Sun, D., and Zhang, Y. 2008. Establishment of paternity testing system using microsatellite markers in Chinese Holstein. J. Genet. Genomics, **35**: 279–284. doi:10.1016/S1673-8527(08)60040-5. PMID:18499071.

Tokarska, M., Marshall, T., Kowalczyk, R., Wojcik, J.M., Pertoldi, C., Kristensen, T.N., Loeschcke, V., Gregersen, V.R., and Bendixen, C. 2009. Effectiveness of microsatellite and SNP markers for parentage and identity analysis in species with low genetic diversity: the case of European bison. Heredity (Edinb), **103**: 326–332. doi:10.1038/hdy.2009.73. PMID:19623210.

Weller, J.I., Feldmesser, E., Golik, M., Tager-Cohen, I., Domochovsky, R., Alus, O., Ezra, E., and Ron, M. 2004. Factors affecting incorrect paternity assignment in the Israeli Holstein population. J. Dairy Sci. **87**: 2627–2640. doi:10.3168/jds.S0022-0302(04)73389-5. PMID:15328288.

Weller, J.I., Glick, G., Ezra, E., Zeron, Y., Seroussi, E., and Ron, M. 2010. Paternity validation and estimation of genotyping error rate for the BovineSNP50 BeadChip. Anim. Genet. **41**: 551–553. doi:10.1111/j.1365-2052.2010.02035.x. PMID:20331599.

Downloaded From: https://bioone.org/journals/Canadian-Journal-of-Animal-Science on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use