J. Dairy Sci. 102:1806–1810
https://doi.org/10.3168/jds.2018-15076


# Pedigree verification and parentage assignment using genomic information in the Mexican Holstein population

**A. García-Ruiz,[1] G. R. Wiggans,[2]* and F. J. Ruiz-López[1]†**
[1]Centro Nacional de Investigación en Fisiología y Mejoramiento Animal, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Ajuchitlán, Querétaro 76280, México
[2]USDA, Agricultural Research Service, Animal Genomics and Improvement Laboratory, Beltsville, MD 20705-2350

## ABSTRACT

Genealogical information is an essential tool for carrying out any genetic improvement program. The objective of this study was to determine the accuracy of pedigree information in the Mexican registered Holstein population using genomic data available in Mexico and for the US Holstein population. The study included 7,508 animals (158 sires and 7,350 cows) that were born from 2002 through 2014, registered with Holstein de México, and genotyped with single nucleotide polymorphism arrays of different densities. Parentage could not be validated for 17% of sires of cows and 12% of sires of bulls. Most (79%) of the dams of cows and the dams of bulls had no genotype available and could not be validated. A parentage test was possible for only 6,104 sires of cows, 139 sires of bulls, 1,519 dams of cows, and 33 dams of bulls. Of the animals with a parentage test, parent assignment was confirmed for 89% of sires of cows, 92% of dams of cows, 95% of sires of bulls, and 97% of dams of bulls. Parent discovery was possible for some animals without confirmed parents: 17% for sires of cows, 2.5% for dams of cows, 43% for sires of bulls, and 0% for dams of bulls. Of the 7,795 progeny tests, 777 had parent conflicts, which is an error rate of 9.97% for parental recording in the population, a rate that is similar to those recently reported for other populations. True parents for some progeny conflicts (15%) were discovered for the Mexican population, and the remaining parents were assigned as unknown. Expected effects of misidentification on rate of genetic gain could be decreased by half if genealogical errors were decreased to 5%. This study indicates that genotyping and genealogy recovery may help in increasing rates of genetic improvement in the Mexican registered Holstein population.

Received May 16, 2018.
Accepted October 22, 2018.
*Retired; current address: Council on Dairy Cattle Breeding, 4201 Northview Dr., Suite 302, Bowie, MD 20716.
†Corresponding author: ruiz.felipe@inifap.gob.mx

**Key words:** parentage verification, Mexican Holstein, ancestor genotype, parent discovery

## INTRODUCTION

Genealogical information is an essential tool for carrying out any genetic breeding program. Its accuracy and completeness influence the reliability of EBV as well as the estimation and control of inbreeding rates (Israel and Weller, 2000; Harder et al., 2005; Heaton et al., 2014) and genetic standard deviation (Banos et al., 2001).

Different methodologies have been used for parentage verification of cattle since its start in the 1960s. Initially, blood typing was used throughout the world as a regular part of cattle breeding programs (Stormont, 1967). Although this method was effective for progeny exclusion (approximately 96%; Rendel, 1958), it was not easy to apply because of the complexity associated with handling of blood samples; blood parentage testing was only possible with fresh blood, which required specialized assistance and made the test expensive. Additionally, retrospective analysis was not possible (Bowling, 2001). In the early 1990s, short tandem repeats of DNA known as microsatellites began being used for parentage verification (Usha et al., 1995); their success was based on their high degree of polymorphism, which made them useful for paternity testing (Eggleston-Stott et al., 1997) with high accuracy (~99%) when a set of microsatellites was used (Usha et al., 1995). Blood typing and microsatellites worked for parentage exclusion, but validation or discovery of ancestry was not possible (Usha et al., 1995; Bowling, 2001). More recently, a different type of DNA marker known as SNP has been adopted; SNP are highly attractive because they are abundant, genetically stable, and amenable to high-throughput automated analysis (Heaton et al., 2002). In cattle, a set of approximately 100 SNP is used for parentage verification with an accuracy of >99%. However, the usefulness of a particular set of parentage SNP varies by breed depending on the minor allele frequency of each SNP and context sequences, as

these intrinsic molecular properties affect SNP testing (Heaton et al., 2014).

Single nucleotide polymorphisms are not exclusively used for parentage verification. For genomic selection they are also used as a third source of information in genetic improvement programs in different species. In Mexico, pedigree accuracy needed to be evaluated for the genomic selection program that was initiated in 2012. Because the Mexican Holstein population is largely derived genetically from US and Canadian Holstein populations (García-Ruiz et al., 2015), genomic data from US Holstein cattle have been used for the Mexican Holstein parentage verification program. The objective of our study was to determine the accuracy of pedigree information from the Mexican Holstein population using genomic data available for Mexican and US Holstein populations.

## MATERIALS AND METHODS

The parentage testing program included 7,508 animals registered with Holstein de México, AC (Querétaro, México); the 158 sires and 7,350 cows were born from 2002 through 2014 and genotyped with SNP arrays of different densities (Table 1). Genotypes were used after quality control analysis with parameters established for the US dairy cattle population (Wiggans et al., 2011), which excluded genotypes with a call rate of <90% and a departure from Hardy-Weinberg equilibrium. Procedures used to assess and discover parent-progeny relationships were those used for the US population (Wiggans et al., 2010). All SNP in common between parent and progeny were compared. An imputed genotype was used for the dam if she had not been genotyped but had sufficient genotyped progeny. Opposite homozygous alleles between parent and progeny were designated as a SNP conflict. A heterozygous progeny with both parents homozygous for the same allele also was designated as a SNP conflict. The threshold for declaring a parent-progeny conflict was when 0.47% of the SNP in common were conflicts.

**Table 1.** Percentage of genotyped animals in the Mexican Holstein population that were included in parentage testing by chip density and animal sex

| Chip density (no. of SNP) | Female | Male |
|---|---|---|
| 200 | 33.76 | 0.00 |
| 9,000 | 6.22 | 0.00 |
| 29,000 | 20.40 | 0.46 |
| 50,000 | 9.12 | 0.35 |
| 77,000 | 23.01 | 1.29 |
| 150,000 | 3.93 | 0.03 |
| 777,000 | 1.43 | 0.00 |

Mexican and US Holstein ancestor genotype databases were used for parentage testing and discovery. Results were classified as confirmed, conflict, or not testable. The not testable class included animals with no genotyped ancestors. The conflict class included animals for which the reported dam or sire was not correct, although parent discovery was possible later for some of these animals.

The effect of parentage error on genetic gain was measured as in Visscher et al. (2002). Two levels of parentage error were examined (the current level and a possible level) as well as different heritabilities (0.10 to 0.30) and different numbers of progeny per bull (10, 20, or 50).

## RESULTS AND DISCUSSION

### *Parentage Validation Not Possible Because of Incomplete Information*

Table 2 shows results of parentage testing of Mexican animals for 4 relationship categories: sires of cows, dams of cows, sires of bulls, and dams of bulls. Among the animals included in the study, 17 and 12% of sires of cows and bulls, respectively, did not have a parent genotype available for parentage validation. Although most of the Holstein genetic material used in Mexico comes from the United States and Canada, the Mexican Holstein population has also received genetic material from Italy, France, England, Spain, New Zealand, and other countries. Genotypes for some of that material were not available for parentage testing. If a databank of genotypes could be assembled for international paternity tests for AI bulls as in the GENOEX initiative presented by Interbull (Dürr et al., 2014), more extensive parentage validation and even parent discovery might be possible, which would increase gain in genetic improvement of dairy cattle worldwide. For most dams of cows and bulls (79% for both sources), a genotype was not available for testing because most dams of cows are from Mexico and have not been genotyped as a result of limited genomic testing programs for females in Mexico, lack of biological samples, and the short time that SNP markers have been used as part of genetic improvement programs. The males that were tested for parentage were Mexican sires, and no biological samples were available for their dams.

### *Parentage Validation*

Parentage testing was only possible for 6,104 sires of cows, 139 sires of bulls, 1,519 dams of cows, and 33 dams of bulls. Of the 7,795 parentage tests, 777 were conflicts, which is a 9.97% rate of parental recording

**Table 2.** Numbers of animals with confirmed parentage after testing, animals with discovered parents after parentage conflicts, and animals not tested because of lack of parental genomic information by source of information

| Information source | Animals with confirmed parents (no.) | Animals with parentage conflicts (no.) | | Animals without parentage testing (no.) |
|---|---|---|---|---|
| | | Parent discovered | Parent not discovered | |
| Sires of cows | 5,457 | 110 | 537 | 1,246 |
| Dams of cows | 1,397 | 3 | 119 | 5,831 |
| Sires of bulls | 132 | 3 | 4 | 19 |
| Dams of bulls | 32 | 0 | 1 | 125 |

error for animals that could be tested. Errors in recording of parents can result from many sources: on the farm, at AI centers, in genotyping laboratories, or even genotype format (McClure et al., 2018). Determining the cause of parentage errors for the Mexican Holstein population was not possible with the data available for our study.

Cow sire was confirmed for 89.4% of the animals. Sire discovery was possible for 17% of cows with sire conflicts (10.6% of cows with parentage tests) because sire genotype was available and use of that sire was reported for that herd. For the remaining 83% of cows with sire conflicts, sire discovery was not possible because the sire's genotype was not available in the genome bank, probably because the sire was from another country or was a nongenotyped local sire. The conflict rate for sires of cows in the Mexican population was similar to the 7 to 12% reported recently for other populations. An incorrect sire was detected for 7% of the German Angeln dairy cattle population (Sanders et al., 2006), and a rate of 7 to 9% was reported for pedigree errors nationally in Ireland (McClure et al., 2018). For Holstein cows sired by AI bulls, the error rate was 11% in the United States (Banos et al., 2001) and 12% in Israel (Weller et al., 2004).

Cow dam was confirmed for 92% of animals. The percentage of parentage conflicts for dams of cows was lower (8%) than that for sires of cows. However, the percentage of discovery was also lower (2.5%), perhaps because of the limited number of Mexican cows with genotypes and the wide use of sires for which genotypes were not available. For dams of cows, parentage confirmation and discovery results from other countries were not available for comparison.

For bulls, parent confirmation was 95% for sires and 97% for dams. Parentage testing was limited for bulls and the 5% conflict rate for sires of bulls was half of that for sires of cows. Nevertheless, sire discovery was possible for 43% of conflicts, perhaps because only bulls that are sons of elite sires are kept for reproductive purposes. For dams of bulls, parentage was confirmed for 97% of animals tested, but the dam could not be discovered for the 1 conflict observed. Differences in parentage error rates between cows and bulls could be because bulls usually are the result of specific matings and greater attention is given to that cow at AI service and calving (Spelman, 2002).

### *Effect of Paternal Errors on Rate of Genetic Gain*

Pedigree errors have negative effects on genetic selection and improvement and, thus, directly affect the improvement rate for economic value of germplasm. With a paternal error rate of 10% (as currently found in the Mexican population), heritability of 0.10, and 20 progeny per bull, a 7% loss of genetic gain is expected; that loss decreases as heritability or progeny per bull increases (Figure 1). If the implementation of genomic testing caused the paternal error rate to decrease to 5%, the expected loss of genetic gain would decrease by half, on average, especially for higher heritability and larger numbers of progeny per bull (Figure 2). Similar results have been reported for other populations. A stochastic simulation with a 10% paternal error rate and a trait with a heritability of 0.25 showed that genetic gain decreased by 4.3% per year (Israel and Weller, 2000). In a simulation with the same heritability and error rate, genetic variance decreased approximately 8% for progeny-tested bulls and approximately 14% for progeny-test bulls (Harder et al., 2005). In other studies with similar conditions, a decrease in EBV reliability was observed (Harder et al., 2005; Sanders et al., 2006).

Banos et al. (2001) determined the effect of paternal rate error for US Holstein cows (11%) on genetic evaluation and estimation of genetic parameters and EBV for bulls across countries. Their results showed a higher decrease (approximately 11% for production traits) in national genetic gain than that reported by Israel and Weller (2000) and up to 18% in losses of genetic gain for international evaluations. Additionally, standard deviations of sire transmitting ability decreased by 8 to 9% and estimates of inbreeding coefficients were reduced by approximately 7 to 14% (Banos et al., 2001).

Based on the results in simulation studies and for other dairy populations, a decrease in genetic gains of approximately 4 to 11% was expected in Mexico, with

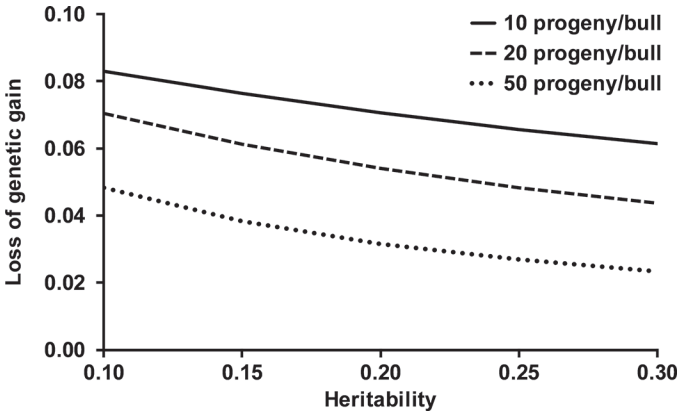


**Figure 1.** Loss of genetic gain in a population with 10% paternity errors by heritability and number of progeny per bull.

the negative effect directly dependent on trait heritability. A trait with lower heritability will have higher losses of genetic gain than a trait with higher heritability (Harder et al., 2005).

Although just a few of the parental conflicts in the Mexican population were corrected (approximately 15% discovery), the use of SNP parental tests will still decrease the adverse effects of incorrect pedigree information by assigning missing parents to problem animals. The effect of missing information is less than that of wrong parentage assignment (Sanders et al., 2006).

### *Program to Increase Parentage Validation for Mexican Holsteins*

Holstein de México samples 1 of every 300 registered females in the herdbook, all embryo donors, at least 1 embryo-transfer product per flush, and all registered sires. These animals are genotyped and tested for parentage if the ancestor genotype is available. Unfortunately, parentage testing is not always possible because most semen used in Mexico is imported and no agreement exists for providing genotype information for parentage testing of imported sires. Collaboration between foreign AI companies and Holstein de México could help to improve the possibility of progeny testing. Another incentive for AI companies that distribute genetic material in Mexico to provide genotypes is the opportunity for increased sales if breeders know that they can perform parentage tests. Another viable option for Holstein de México is to participate in the GENOEX program, which has a parentage SNP-exchange database that member countries can access to perform parentage verification tests based on 200 SNP specified by the International Society for Animal Genetics (2013). Benefits of GENOEX participation currently may not outweigh the additional costs of becoming a service user because the proportion of sires with information useful for the Mexican Holstein population presently is unknown.

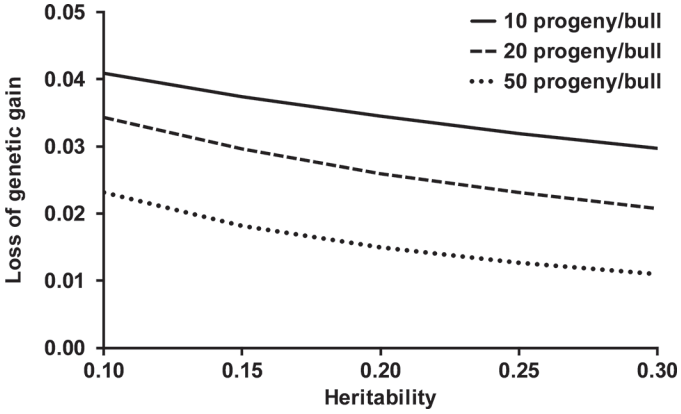


**Figure 2.** Loss of genetic gain in a population with 5% paternity errors by heritability and number of progeny per bull.

The percentage of Mexican animals with a parentage test will increase in future years as breeders use genomic testing, a service that includes parentage testing. Additionally, as part of genomic evaluation service, Holstein de México has created a bank of collected hair samples, which could increase the possibility of future parentage tests, especially for dams of cows.

Benefits of increasing the accuracy of pedigree information for the registered population are expected to extend to commercial herds. The pedigree error rate for dams in commercial herds can be substantial, which has a negative and additive effect on the rate of genetic gain (Harder et al., 2005).

## CONCLUSIONS

Percentage of pedigree errors reported in the Mexican Holstein population (approximately 10%) was in the range reported in other populations. This error rate is expected to result in a loss of genetic gain from 2 to 8%, depending on the heritability of the trait and the number of progeny per bull. This loss of genetic gain could be decreased to 1 to 4% if the error rate was decreased to 5%. These results show the advantage of improving the quality of pedigree information through the application of a continuous parentage verification and discovery program. Because of the extensive use of international genetic material in Mexico, access to an international genotype database is necessary to support parentage testing and discovery or at least acquisition of genotype information for Holstein animals imported into Mexico. Additionally, the Mexican Holstein population should continue routine parentage testing and,

if possible, increase the proportion of animals being tested. Genotyping of Mexican cows must continue to encourage testing of dams of cows and bulls. These actions could help improve the Mexican genetic evaluation program and allow breeders to make more accurate selection decisions.

## ACKNOWLEDGMENTS

This study was supported by Consejo Nacional de Los Recursos Genéticos Pecuarios, Instituto Nacional de Investigaciones Forestales, Agricolas y Pecuarias (Ciudad de México, México), research project 11513634465, "Study of Inbreeding and Its Effect on Productive and Reproductive Traits in Holstein Cattle." We thank Association Holstein de México (Querétaro, México), the Cooperative Dairy DNA Repository (Columbia, MO), and the Council on Dairy Cattle Breeding (Bowie, MD) for providing data. Funding for G. R. Wiggans was solely from USDA Agricultural Research Service appropriated project 8042-31000-101-00-D, "Improving Genetic Predictions in Dairy Animals Using Phenotypic and Genomic Information." The authors thank Suzanne Hubbard (Animal Genomics and Improvement Laboratory, ARS, USDA) for manuscript review and 2 anonymous reviewers for suggestions for manuscript improvement. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by USDA; USDA is an equal opportunity provider and employer.

## REFERENCES

Banos, G., G. R. Wiggans, and R. L. Powell. 2001. Impact of paternity errors in cow identification on genetic evaluations and international comparisons. J. Dairy Sci. 84:2523–2529. https://doi.org/10.3168/jds.S0022-0302(01)74703-0.

Bowling, A. T. 2001. Historical development and application of molecular genetic tests for horse identification and parentage control. Livest. Prod. Sci. 72:111–116. https://doi.org/10.1016/S0301-6226(01)00271-8.

Dürr, J. W., H. Jorjani, and R. Reents. 2014. International genotype exchange platform (GENOEX). 39th ICAR Session, Berlin (Germany). Accessed Oct. 3, 2018. https://www.icar.org/wp-content/uploads/2016/07/Berlin-2014-Durr-Jorjani-Reents-International-Genotype-Exchange-Platform-GENOEX.pdf.

Eggleston-Stott, M. L., A. DelValle, M. Bautista, S. Dileanis, E. Wictum, and A. T. Bowling. 1997. Nine equine dinucleotide repeats at microsatellite loci UCDEQ136, UCDEQ405, UCDEQ412, UCDEQ425, UCEDQ437, UCDEQ467, UCDEQ487, UCDEQ502 and UCDEQ505. Anim. Genet. 28:370–371.

García-Ruiz, A., F. de J. Ruiz-López, C. P. Van Tassell, H. H. Montaldo, and H. J. Huson. 2015. Genetic differentiation of Mexican Holstein cattle and its relationship with Canadian and U.S. Holsteins. Front. Genet. 6:7. https://doi.org/10.3389/fgene.2015.00007.

Harder, B., J. Bennewitz, N. Reinsch, M. Mayer, and E. Kalm. 2005. Effect of missing sire information on genetic evaluation. Arch. Anim. Breed. 48:219–232. https://doi.org/10.5194/aab-48-219-2005.

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 U.S. beef cattle. Mamm. Genome 13:272–281. https://doi.org/10.1007/s00335-001-2146-3.

Heaton, M. P., K. A. Leymaster, T. S. Kalbfleisch, J. W. Kijas, S. M. Clarke, J. McEwan, J. F. Maddox, V. Basnayake, D. T. Petrik, B. Simpson, T. P. L. Smith, C. G. Chitko-McKown, and the International Sheep Genomics Consortium. 2014. SNPs for parentage testing and traceability in globally diverse breeds of sheep. PLoS One 9:e94851. https://doi.org/10.1371/journal.pone.0094851.

International Society for Animal Genetics. 2013. ISAG cattle core + additional SNP panel. Accessed Oct. 3, 2018. https://www.isag.us/Docs/Cattle-SNP-ISAG-core-additional-panel-2013.xlsx.

Israel, C., and J. I. Weller. 2000. Effect of misidentification on genetic gain and estimation of breeding value in dairy cattle populations. J. Dairy Sci. 83:181–187. https://doi.org/10.3168/jds.S0022-0302(00)74869-7.

McClure, M. C., J. McCarthy, P. Flynn, J. C. McClure, E. Dair, D. K. O'Connell, and J. F. Kearney. 2018. SNP data quality control in a national beef and dairy cattle system and highly accurate SNP based parentage verification and identification. Front. Genet. 9:84. https://doi.org/10.3389/fgene.2018.00084.

Rendel, J. 1958. Studies of cattle blood groups. II. Parentage tests. Acta Agric. Scand. A Anim. Sci. 8:131–161. https://doi.org/10.1080/00015125709441453.

Sanders, K., J. Bennewitz, and E. Kalm. 2006. Wrong and missing sire information affects genetic gain in the Angeln dairy cattle population. J. Dairy Sci. 89:315–321. https://doi.org/10.3168/jds.S0022-0302(06)72096-3.

Spelman, P. 2002. Utilisation of molecular information in dairy cattle breeding. Commun. 22–02 in Proc. 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France. INRA, Castanet-Tolosan, France.

Stormont, C. 1967. Contribution of blood typing to dairy science progress. J. Dairy Sci. 50:253–260. https://doi.org/10.3168/jds.S0022-0302(67)87401-0.

Usha, A. P., S. P. Simpson, and J. L. Williams. 1995. Probability of random sire exclusion using microsatellite markers for parentage verification. Anim. Genet. 26:155–161. https://doi.org/10.1111/j.1365-2052.1995.tb03155.x.

Visscher, P. M., J. A. Woolliams, D. Smith, and J. L. Williams. 2002. Estimation of pedigree errors in the UK dairy population using microsatellite markers and the impact on selection. J. Dairy Sci. 85:2368–2375. https://doi.org/10.3168/jds.S0022-0302(02)74317-8.

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

Wiggans, G. R., P. M. VanRaden, L. R. Bacheller, M. E. Tooker, J. L. Hutchison, T. A. Cooper, and T. S. Sonstegard. 2010. Selection and management of DNA markers for use in genomic evaluation. J. Dairy Sci. 93:2287–2292. https://doi.org/10.3168/jds.2009-2773.

Wiggans, G. R., P. M. VanRaden, and T. A. Cooper. 2011. The genomic evaluation system in the United States: Past, present, future. J. Dairy Sci. 94:3202–3211. https://doi.org/10.3168/jds.2010-3866.