Received: 31 May 2023 | Accepted: 2 June 2023

DOI: 10.1111/age.13334

SHORT COMMUNICATION

ANIMAL GENETICS WILEY

Increased heterozygosity in low-pass sequencing data allows identification of blood chimeras in cattle

Abstract

In about 90% of multiple pregnancies in cattle, shared blood circulation between fetuses leads to genetic chimerism in peripheral blood and can reduce reproductive performance in heterosexual co-twins. However, the early detection of heterosexual chimeras requires specialized tests. Here, we used low-pass sequencing data with a median coverage of 0.64× generated from blood samples of 322 F1 crosses between beef and dairy cattle and identified 20 putative blood chimeras through increased levels of genome-wide heterozygosity. In contrast, for 77 samples with routine SNP microarray data generated from hair bulbs of the same F1s, we found no evidence of chimerism, simultaneously observing high levels of genotype discordance with sequencing data. Fifteen out of 18 reported twins showed signs of blood chimerism, in line with previous reports, whereas the presence of five alleged singletons with strong signs of chimerism suggests that the *in-utero* death rate of co-twins is at the upper limit of former estimates. Together, our results show that low-pass sequencing data allow reliable screening for blood chimeras. They further affirm that blood is not recommended as a source of DNA for the detection of germline variants.

Blood chimeras are individuals that contain genetically distinct blood cell lines. In cattle, this is commonly caused by the development of placental vascular anastomoses between dizygotic twins, enabling the exchange of hematopoietic stem cells (Gurgul et al., 2014; Niku et al., 2007; Owen, 1945). Through their establishment in the bone marrow, a mixture of genetically distinct blood cells is produced throughout the lifetime (Niku et al., 2007; Ohno et al., 1962). In contrast, chimerism is rare in other tissues like skin or germ cell lines (Dunn et al., 1979; Ohno et al., 1962; Plante et al., 1992; Ron et al., 1995). For dizygotic twins of both sexes, the transmission of masculinization molecules, produced by the testes of the male co-twin, through anastomoses early in pregnancy can lead to abnormal development of external genitalia of the female co-twin. The resulting disorder, known as freemartinism, is the most common type of disorder of sex development in cattle and diagnosed by physical examination, cytogenetic detection of XX/XY leukocyte chimerism or molecular detection of Y chromosome-specific DNA (Esteves et al., 2012; Komisarek & Dorynek, 2002; Lillie, 1916; Padula, 2005). In contrast, the effects on reproductive performance for male co-twins are less clear, as some studies showed reduced fertility parameters in chimeric bulls, whereas others found no differences from control animals (Kozubska-Sobocińska et al., 2016; Padula, 2005).

With the frequency of multiple births in cattle ranging between 1 and >5% (Cady & Van Vleck, 1978; Echternkamp & Gregory, 2002; Komisarek & Dorynek, 2002; Rutledge, 1975; Silva del Río et al., 2007; Widmer et al., 2022), about 95% of twins being dizygotic (Erb & Morrison, 1959; Fricke, 2001; Silva del Río et al., 2006) and 82–97% of dizygotic twins being found to be blood chimeras (Buoen et al., 1992; Plante et al., 1992), about 1.5–8.8% of calves can be expected to show blood chimerism. *In-utero* death of one co-twin, occurring at a frequency of 5–20% of diagnosed twin pregnancies in cattle (Andreu-Vazquez et al., 2012; Cockcroft & Sorrell, 2015; Echternkamp & Gregory, 2002; López-Gatius & Hunter, 2005; Silva del Río et al., 2009), can result in additional single-born blood chimeras. Although this makes blood an error-prone source of DNA for the detection of germline variants, it allows the detection of chimeric individuals with potentially reduced reproductive performance. Here, we used low-pass sequencing data to (i) assess genotype calling accuracy when blood is used as a source of DNA, (ii) investigate the utility of low-pass sequencing data to detect blood chimeras during routine genomic analysis, without requiring further specialized tests and (iii) estimate the frequency of blood chimeras in twins and single-born calves.

A total of 322 F1 crosses between 13 Aberdeen Angus (AAN), 28 Limousin (LIM) and 13 Simmental (SIM) sires and 311 Brown Swiss (BSW) cows (Table 1) were included in this study. Of those, 71 F1s (22%) were the result of insemination with a balanced mixture of

Cross	*N* F1s	*N* SILIAN inseminations	*N* twins	*N* sires	*N* dams
AAN × BSW	96 (29)	24 (24)	7 (4)	8 (8)	61 (61)
LIM × BSW	130 (10)	8 (8)	5 (1)	18 (18)	58 (56)
SIM × BSW	96 (38)	39 (37)	6 (2)	8 (8)	65 (64)
Total	322 (77)	71 (69)	18 (7)	34 (34)	184 (181)

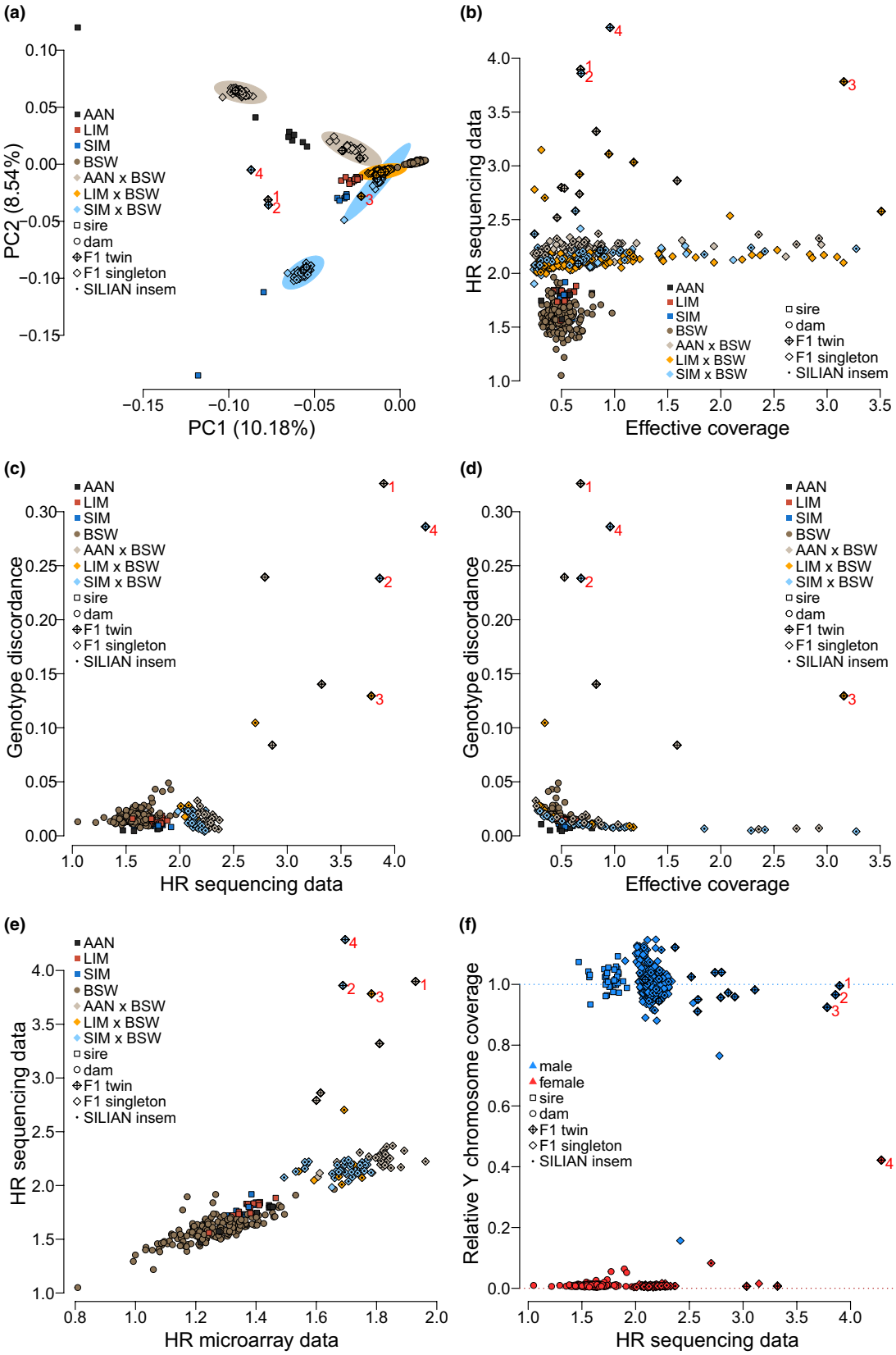
TABLE 1 Numbers of sequenced F1s, Aberdeen Angus (AAN), Limousin (LIM) or Simmental (SIM) sires, and Brown Swiss (BSW) dams for each cross type (number of samples with additional microarray genotypes are in parentheses).

semen obtained from three different bulls, SIM, LIM and AAN sires (called SILIAN inseminations hereafter). In SILIAN inseminations, bulls from different breeds may have sired one of each co-twin, facilitating the detection of blood chimeras with sequencing data owing to a larger number of paternally inherited alleles that differ between cell lines. According to the BrunaNet herdbook (https://brunanet.ch/BrunaNet/), 18 F1s (5.6%) were recorded as twins (including six twin sets), nine had no information available and the remaining F1s (91.6%) were recorded as singletons. One female had a male co-twin and thus represents a potential freemartin, whereas the remaining twins were of the same sex as their co-twin. Five twins, including the potential freemartin, resulted from SILIAN inseminations. For all 322 F1s, DNA was extracted from EDTA blood using the Promega Maxwell RSC DNA system, sequenced as 150 bp paired-end reads on an Illumina NovoSeq 6000 at Neogen, and imputed separately for each sample to a cattle reference panel comprising 946 animals (Snelling et al., 2020) with Gencove's LOIMPUTE pipeline v0.1.5 (Wasik et al., 2021; https://gitlab.com/gencove/loimpute-public) and with the ARS-UCD1.2 reference genome (Rosen et al., 2020), resulting in a median of 0.64× (0.24–3.51×) effective coverage (Li et al., 2021). In addition, DNA from the semen of 34 bulls and hair bulbs of 184 cows, being the sires and dams of 273 and 186 F1s, respectively, was sequenced and imputed as aforementioned to a median of 0.49× (0.31–0.98×) effective coverage. The parental samples allowed us to simulate genotypes of non-chimeric and chimeric progeny based on the same parents as in the real data. Single-sample VCF files of F1s, sires and dams, each comprising 59 204 166 SNPs and 13 indels across 29 autosomes and the X and Y chromosome, were combined. Only biallelic SNPs with less than 5% low-confidence genotypes (i.e. maximum genotype probability <0.9), total depth per site below twice its mean over all sites and minor allele frequency larger than or equal to 5% were retained, leaving 8 255 351 autosomal variants for further analysis unless noted otherwise.

To ascertain the population structure of F1s and their parents, principal component analysis (PCA) was computed with PLINK v1.90b6.26 (Chang et al., 2015; www.cog-genomics.org/plink) --pca, using a subset of 192 437 autosomal SNPs that were retained after linkage disequilibrium-pruning of filtered data with PLINK v2.00a3.7 --indep-pairwise 100 10 0.2 (i.e. using sliding windows of 100 SNPs, steps of 10 SNPs and pruning variants within a window with squared correlation greater than 0.2). The results showed a distinct separation among parental breeds but also among crosses (Figure 1a). Some subclustering occurred for AAN × BSW and SIM × BSW crosses, which resulted from a high number of descendants from two individual bulls with 69 and 73 offspring, respectively. Despite the predominant grouping, four F1s (tagged with red numbers in Figure 1a) were positioned outside their expected clusters. As these four F1s were twins from SILIAN inseminations, we suspected that they were blood chimeras of dizygotic twins with sires of different breed origins.

As chimeras carry an unbalanced mix of distinct cell lines, an increased number of heterozygous sites in comparison with non-chimeric samples is expected (Gurgul et al., 2014). In addition, for a subset of apparent heterozygotes, allele counts will be imbalanced (i.e. deviate from their expected 1:1 ratio) when homozygous and heterozygous cell lines are mixed (Conlin et al., 2010; Gurgul et al., 2014). With low-pass sequencing data, accurately inferring a heterozygous genotype is difficult because of the reduced probability of sequencing both alleles, and this difficulty should be more pronounced at a locus with

FIGURE 1 Characteristics of low-pass sequencing vs. microarray data. (a) Visualization of the first and second principal components (PC1 and PC2) from linkage disequilibrium-pruned autosomal SNPs of sequencing data. Ellipses highlight clusters of F1s of the three different crosses. Four F1s that are positioned outside their expected clusters are tagged with red numbers. (b) Heterozygosity to non-reference homozygosity ratio (HR) of sequencing data shows little dependence on effective sequencing coverage but differs among breeds and crosses. The majority of F1s recorded as twins and a few F1s recorded as singletons show unusually high HRs. (c) Samples with an unusually high HR in sequencing data have high genotype discordances between sequencing and microarray data. (d) Genotype discordance between sequencing and microarray data tends to decrease with effective sequencing coverage, but this cannot explain unusually high genotype discordances for a subset of samples. (e) The HRs of microarray and sequencing data are correlated except for a subset of samples with unusually high HRs in sequencing data. (f) Intermediate relative Y chromosome coverage suggests a mixture of XX and XY cell lines in some samples with high HR in sequencing data. Symbols are colored by breed or cross in (a)–(e), or by sex in (f). Sires, dams and F1s are indicated with squares, circles and diamonds, respectively. For F1s, additional crosses highlight twins, dots show SILIAN inseminations and red numbers tag putative blood chimeras between breeds. AAN, Aberdeen Angus; BSW, Brown Swiss; LIM, Limousin; SIM, Simmental.



imbalanced allele counts. We thus investigated whether heterozygosity was increased in PCA outliers compared with other samples, which would support a chimeric origin. To identify unusually high levels of heterozygosity within samples, we used the heterozygosity to non-reference homozygosity ratio (HR hereafter), which is expected to be 2.0 for whole-genome sequencing variants in Hardy–Weinberg equilibrium within an individual

(Guo et al., 2013), but can be lower or higher depending on ancestry and admixture (Samuels et al., 2016; Wang et al., 2014). The measure has been shown to be consistent across genome regions or functionalities and can be used for quality control (Wang et al., 2014), for example, to identify sample contamination (Zhao et al., 2022). We found median HRs of 1.61 for parental samples and of 2.15 for F1s, where individual values ranged from 1.05 to 1.96 and from 1.90 to 4.29, respectively (Figure 1b; Figure S1), with only a minor increase in HR with sequencing coverage (parentals: Pearson's r 0.03, p-value 0.613; F1s: Pearson's r 0.17, p-value 0.002).

Increased HRs are expected for F1s; however, we observed a clear change in the distribution of values at about 2.38, separating 20 samples from the majority of F1s. These samples were identical to those identified as high outliers with 'Tukey's fences' (Tukey, 1977) and formed a distinct group with model-based clustering using the Mclust function in R (Scrucca et al., 2016). Fifteen out of 20 samples with HR above 2.38 were recorded as twins, including the four suspected blood chimeras, which showed the highest HR of all samples. One further twin showed a slightly increased HR of 2.37, whereas the two remaining twins had values of 2.09 and 2.22 in the same range as the majority of F1s. HRs within twin sets were highly correlated (Pearson's r 0.96, p-value 0.002), which has been previously observed for blood chimeras (Basrur & Kanagawa, 1969; Marcum et al., 1972) and thus made sample contamination unlikely.

To determine whether an unusually high HR, in comparison with the majority of samples, can indeed indicate chimerism, we performed numerical simulations based on alternative allele dosages of parental samples for 30 randomly selected trios per cross type. We simulated non-chimeric progeny and chimeras resulting from a balanced mixture of genetic material between siblings (Appendix S1). The simulations demonstrated that for our dataset, HRs exceeding 2.38 can most likely be explained by a mixture of genotypes in chimeras (Figures S2 and S3).

Chimerism confined to a specific tissue will result in increased genotype discordance when other tissues are compared with the chimeric tissue. Although we could not obtain different tissues for F1s, SNP microarray genotypes from DNA obtained from semen or hair bulbs were provided by Swiss breeding associations for 77 F1s, 34 sires and 181 dams (Table 1). These samples were genotyped using seven different microarrays encompassing 26K–139K SNP markers (Table S1). Genotype discordance between the two data sources, separately for each microarray, were computed with PLINK v2.00a3.7 --sample-diff using default settings. Depending on the microarray, 12 333–70 585 autosomal variants that could be matched between the microarray and the filtered low-pass sequencing data and with call rates higher than 0.5 in either and call rates higher than 0.2 in the combined dataset were kept for discordance calculations. Our results showed that all samples with unusually high HR also had high genotype discordance (Figure 1c), which was independent from sequencing coverage (Figure 1d), whereas the HR computed from microarray data of semen or hair bulbs was not increased (Figure 1e). Therefore, these results further supported that the samples with high HR observed by low-pass sequencing were most likely blood chimeras.

To identify potential chimeras between heterosexual co-twins, we computed the relative Y chromosome coverage as the ratio of the average read coverage of loci on the Y to the average read coverage of loci on the X chromosome. This ratio is expected to be zero for females and one for males, whereas intermediate values will point to a mixture of XX and XY cell lines within chimeras. To increase power, we computed coverages for all 6141 and 2 136 885 SNPs on the Y and X chromosome, respectively, prior to filtering. Our results showed that the single documented heterosexual co-twin in our data, which had a high HR, exhibited an intermediate relative Y chromosome coverage of 0.42, which confirmed freemartinism. Our results further revealed that three out of five reported singletons with high HR had intermediate relative Y chromosome coverages as well (Figure 1f), which supported their chimeric origin.

In summary, low-pass sequencing data allowed us to screen for blood chimeras, without requiring further specialized tests like karyotyping, blood grouping or PCR-based detection of Y chromosome-specific DNA (Kozubska-Sobocińska et al., 2016; Padula, 2005; Qiu et al., 2018; Szczerbal et al., 2019; Young & Kirkpatrick, 2016). However, control samples from the same breed or cross are required to establish the expected HR for non-chimeric samples, which depends not only on ancestry and admixture, but also on sequencing coverage. As our data were imputed, the effect of sequencing coverage on HR was minor, in contrast to the HR computed on genotypes called from raw (i.e. unimputed) read counts (Figure S4a). Nevertheless, it remains to be investigated how imputation of chimeric genotypes is affected across a range of different sequencing coverages. Other measures that may be used to identify chimeric samples, like allelic imbalance (Conlin et al., 2010; Gurgul et al., 2014), require high sequencing coverage and were thus not applicable to our low-pass data (Figure S4b). Our finding that 15 out of 18 reported twins (83%) showed unusually high HR suggests that they were indeed blood chimeras, which is in line with previously recorded frequencies. Interestingly, five out of 304 F1s (1.6%) not reported as twins showed similarly high HR, suggesting that a non-negligible number of singletons might be unknown chimeras. In conclusion, the *in-utero* death of co-twins may occur at frequencies of >20%, which together with the high incidence of blood chimerism in cattle, argues against the use of blood as a source of DNA for the detection of germline variants. Instead, other tissues with a lower risk

of chimerism, for example hair roots or skin cells, would be recommended.

KEYWORDS

blood, bovine, chimerism, disorders of sex development, freemartin, heterozygosity ratio

ACKNOWLEDGEMENTS

We would like to thank Isabelle Morel and Sylvain Lerch for designing the RegioBeef project from which we recruited our F1s, and Raphael Siegenthaler for selecting animals and collecting sample information. We would like to thank Braunvieh Schweiz for sharing existing SNP microarray data of F1s and dams, and DNA and hair samples of dams, and providing birth records, parentage information, and database access, Mutterkuh Schweiz for sharing existing SNP microarray data of sires, Swissgenetics for semen samples of sires, IFN Schönow e.V. for sending DNA extractions of dams, and Qualitas AG for facilitating data exchange. We thank the reviewers for their valuable comments on the manuscript. Open access funding provided by Agroscope.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Low-pass sequencing data of F1s are available at NCBI Sequence Read Archive with the BioProject accession number PRJNA930154.

Dorothea Lindtke[1]
Franz R. Seefried[2]
Cord Drögemüller[3]
Markus Neuditschko[1]

[1]*Agroscope, Posieux, Switzerland*
[2]*Qualitas AG, Zug, Switzerland*
[3]*Institute of Genetics, University of Bern, Bern, Switzerland*

Correspondence
Dorothea Lindtke, Agroscope, Rte de la Tioleyre 4, 1725 Posieux, Switzerland.
Email: dorothea.lindtke@agroscope.admin.ch

ORCID

Dorothea Lindtke https://orcid.org/0000-0003-1339-1410
Cord Drögemüller https://orcid.org/0000-0001-9773-522X

REFERENCES

Andreu-Vazquez, C., Garcia-Ispierto, I., Ganau, S., Fricke, P.M. & Lopez-Gatius, F. (2012) Effects of twinning on the subsequent reproductive performance and productive lifespan of high-producing dairy cows. *Theriogenology*, 78(9), 2061–2070.

Basrur, P.K. & Kanagawa, H. (1969) Parallelism in chimeric ratios in heterosexual cattle twins. *Genetics*, 63(2), 419–425.

Buoen, L.C., Zhang, T.Q., Weber, A.F. & Ruth, G.R. (1992) Non-freemartin rate in Holstein heterosexual twins. American Association of Bovine Practitioners Conference, 1, 300-303.

Cady, R.A. & Van Vleck, L.D. (1978) Factors affecting twinning and effects of twinning in Holstein dairy cattle. *Journal of Animal Science*, 46, 950–956.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M. & Lee, J.J. (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4(1), 7.

Cockcroft, P.D. & Sorrell, E.J. (2015) Twinning in Holstein-Friesian dairy cows: proportion carried to term and calf sex ratios. *Veterinary Sciences*, 2(3), 131–134.

Conlin, L.K., Thiel, B.D., Bonnemann, C.G., Medne, L., Ernst, L.M., Zackai, E.H. et al. (2010) Mechanisms of mosaicism, chimerism and uniparental disomy identified by single nucleotide polymorphism array analysis. *Human Molecular Genetics*, 19(7), 1263–1275.

Dunn, H.O., McEntee, K., Hall, C.E., Johnson, R.H., Jr. & Stone, W.H. (1979) Cytogenetic and reproductive studies of bulls born co-twin with freemartins. *Journal of Reproduction and Fertility*, 57(1), 21–30.

Echternkamp, S.E. & Gregory, K.E. (2002) Reproductive, growth, feedlot, and carcass traits of twin vs single births in cattle. *Journal of Animal Science*, 80(E-suppl_2), 1–10.

Erb, R.E. & Morrison, R.A. (1959) Effects of twinning on reproductive efficiency in a Holstein-Friesian herd. *Journal of Dairy Science*, 42, 512–519.

Esteves, A., Båge, R. & Payan Carreira, R. (2012) Freemartinism in cattle. In: Mendes, R. E. (Ed.) *Ruminants: anatomy, behavior and diseases*, pp. 99–120. New York, NY: Nova Science Publishers.

Fricke, P.M. (2001) Twinning in dairy cattle. *The Professional Animal Scientist*, 17(2), 61–67.

Guo, Y., Ye, F., Sheng, Q., Clark, T. & Samuels, D.C. (2013) Three-stage quality control strategies for DNA re-sequencing data. *Briefings in Bioinformatics*, 15(6), 879–889.

Gurgul, A., Rubiś, D. & Bugno-Poniewierska, M. (2014) The effect of using DNA obtained from blood of cattle with genetic chimerism on Illumina's Beadchip assay performance. *Annals of Animal Science*, 14(2), 279–286.

Komisarek, J. & Dorynek, Z. (2002) Genetic aspects of twinning in cattle. *Journal of Applied Genetics*, 43(1), 55–68.

Kozubska-Sobocińska, A., Danielak-Czech, B. & Rejduch, B. (2016) Cytogenetic and molecular diagnostics of XX/XY chimerism in cattle, sheep, and goats – a review. *Annals of Animal Science*, 16(4), 989–1005.

Li, J.H., Mazur, C.A., Berisa, T. & Pickrell, J.K. (2021) Low-pass sequencing increases the power of GWAS and decreases measurement error of polygenic risk scores compared to genotyping arrays. *Genome Research*, 31(4), 529–537.

Lillie, F.R. (1916) The theory of the free-martin. *Science*, 43(1113), 611–613.

López-Gatius, F. & Hunter, R.H.F. (2005) Spontaneous reduction of advanced twin embryos: its occurrence and clinical relevance in dairy cattle. *Theriogenology*, 63(1), 118–125.

Marcum, J.B., Lasley, J.F. & Day, B.N. (1972) Variability of sex-chromosome chimerism in cattle from heterosexual multiple births. *Cytogenetics*, 11(5), 388–399.

Niku, M., Pessa-Morikawa, T., Taponen, J. & Iivanainen, A. (2007) Direct observation of hematopoietic progenitor chimerism in fetal freemartin cattle. *BMC Veterinary Research*, 3(1), 29.

Ohno, S., Trujillo, J.M., Stenius, C., Christian, L.C. & Teplitz, R.L. (1962) Possible germ cell chimeras among newborn dizygotic twin calves (*Bos taurus*). *Cytogenetics*, 1, 258–265.

Owen, R.D. (1945) Immunogenetic consequences of vascular anastomoses between bovine twins. *Science*, 102(2651), 400–401.

Padula, A.M. (2005) The freemartin syndrome: an update. *Animal Reproduction Science*, 87(1), 93–109.

Plante, Y., Schmutz, S.M., Lang, K.D.M. & Moker, J.S. (1992) Detection of leucochimaerism in bovine twins by DNA fingerprinting. *Animal Genetics*, 23(4), 295–302.

Qiu, Q., Shao, T., He, Y., Muhammad, A.U., Cao, B. & Su, H. (2018) Applying real-time quantitative PCR to diagnosis of freemartin in Holstein cattle by quantifying SRY gene: a comparison experiment. *PeerJ*, 6, e4616.

Ron, M., Blank, Y. & Band, M.R. (1995) Determination of the optimal tissue source and number of microsatellites for detection of zygotic origin of cattle twins. *Animal Biotechnology*, 6, 27–39.

Rosen, B.D., Bickhart, D.M., Schnabel, R.D., Koren, S., Elsik, C.G., Tseng, E. et al. (2020) De novo assembly of the cattle reference genome with single-molecule sequencing. *GigaScience*, 9(3), 1–9.

Rutledge, J.J. (1975) Twinning in cattle. *Journal of Animal Science*, 40(5), 803–815.

Samuels, D.C., Wang, J., Ye, F., He, J., Levinson, R.T., Sheng, Q. et al. (2016) Heterozygosity ratio, a robust global genomic measure of autozygosity and its association with height and disease risk. *Genetics*, 204(3), 893–904.

Scrucca, L., Fop, M., Murphy, T.B. & Raftery, A.E. (2016) mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *R Journal*, 8(1), 289–317.

Silva del Río, N., Colloton, J.D. & Fricke, P.M. (2009) Factors affecting pregnancy loss for single and twin pregnancies in a high-producing dairy herd. *Theriogenology*, 71(9), 1462–1471.

Silva del Río, N., Kirkpatrick, B.W. & Fricke, P.M. (2006) Observed frequency of monozygotic twinning in Holstein dairy cattle. *Theriogenology*, 66(5), 1292–1299.

Silva del Río, N., Stewart, S., Rapnicki, P., Chang, Y.M. & Fricke, P.M. (2007) An observational analysis of twin births, calf sex ratio, and calf mortality in Holstein dairy cattle. *Journal of Dairy Science*, 90(3), 1255–1264.

Snelling, W.M., Hoff, J.L., Li, J.H., Kuehn, L.A., Keel, B.N., Lindholm-Perry, A.K. et al. (2020) Assessment of imputation from low-pass sequencing to predict merit of beef steers. *Genes (Basel)*, 11(11), 1312.

Szczerbal, I., Nowacka-Woszuk, J., Albarella, S. & Switonski, M. (2019) Technical note: droplet digital PCR as a new molecular method for a simple and reliable diagnosis of freemartinism in cattle. *Journal of Dairy Science*, 102(11), 10100–10104.

Tukey, J.W. (1977) *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Wang, J., Raskin, L., Samuels, D.C., Shyr, Y. & Guo, Y. (2014) Genome measures used for quality control are dependent on gene function and ancestry. *Bioinformatics*, 31(3), 318–323.

Wasik, K., Berisa, T., Pickrell, J.K., Li, J.H., Fraser, D.J., King, K. et al. (2021) Comparing low-pass sequencing and genotyping for trait mapping in pharmacogenetics. *BMC Genomics*, 22(1), 197.

Widmer, S., Seefried, F.R., von Rohr, P., Häfliger, I.M., Spengeler, M. & Drögemüller, C. (2022) Associated regions for multiple birth in Brown Swiss and original Braunvieh cattle on chromosomes 15 and 11. *Animal Genetics*, 53(5), 557–569.

Young, A.S. & Kirkpatrick, B.W. (2016) Frequency of leukochimerism in Holstein and Jersey twinsets. *Journal of Animal Science*, 94(11), 4507–4515.

Zhao, S., Jiang, L., Yu, H. & Guo, Y. (2022) GTQC: automated genotyping array quality control and report. *Journal of Genomics*, 10, 39–44.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.