Indian Journal of Animal Sciences **92** (6): 751–756, June 2022/Article
https://doi.org/10.56093/ijans.v92i6.108935

Evaluation of INDUSCHIP-1 and selected low-density SNP panel for imputation to higher density in Gir dairy cattle of Gujarat

R S THAKKAR[1]✉, N G NAYEE[2] and P H VATALIYA[1]

Kamdhenu University, Gandhinagar, Gujarat 382 010 India

Received: 28 December 2020; Accepted: 16 November 2021

ABSTRACT

Application of genomic selection (GS) among other selection methods mainly depends upon the cost of genotyping, as cheaper the cost, more animals can be genotyped to increase reference population size. Imputation approaches have been useful in reducing the cost. Imputation strategies and GS have been comprehensively studied in several taurine dairy cattle populations, but very limited information is available on *Bos indicus* populations. Factors that affect the efficiency of imputation are population structure, linkage disequilibrium between markers and marker density in target and reference SNP panels. For the present study, INDUSCHIP-1, a customized Illumina bovine microarray chip for indigenous cattle breeds, designed by the National Dairy Development Board (NDDB), Anand, was used for genotyping Gir cattle in India. The objective of the study was to evaluate the performance of INDUSCHIP-1 for imputation to Illumina Bovine high-density (HD) and also to select a low-density (LD) SNP panel useful to impute at INDUSCHIP-1 level in Gir cattle. Fivefold cross-validation by masking genotypes of animals to keep either INDUSCHIP-1 SNPs or LD SNPs and imputing to either HD level or at INDUSCHIP-1 level respectively was performed. Population-based imputation algorithm BEAGLE was used without including pedigree information. Imputation accuracies were evaluated as concordance level (allele correct rate expressed in percentage). All imputation accuracy, represented as % concordance between actual genotypes and imputed genotypes were shown as an average of five replicates for each chromosome. A mean concordance of 92.33% was observed when INDUSCHIP-1 genotypes were imputed at Illumina HD 777K level. The use of INDUSCHIP-1 as an HD panel resulted in the accuracy of 88.42%. SNP selection criteria used for the LD panel in the present study was very simple and seem effective. The distribution of polymorphic SNPs along the length of chromosomes and across all the chromosomes showed that this custom selected panel is a promising option for developing the LD genotyping chip for Gir cattle.

Keywords: BEAGLE, Concordance, Genotype imputation, Gir, INDUSCHIP-1, LD chip

India is rich in genetic diversity of livestock and possesses a large number of indigenous breeds of cattle. The National Bureau of Animal Genetic Resources (NBAGR) has recognized 43 indigenous breeds of cattle (http://www.nbagr.res.in/regcat.html). Amongst the indigenous breeds, Gir, Sahiwal and Red Sindhi are prominent milch breeds. Gujarat is fortunate of having the origin of the milch breed like Gir. This breed of cattle is an important milch breed for milk in India. Looking at the importance of this breed in the rural economy, it is essential to improve the productivity of Gir cows through modern technological advancement in animal breeding. Genomic selection (GS), one such approach requires genotyping of a large number of animals along with performance records for creating reference population. Marker density is one of the factors affecting the accuracy of GS. However, genotyping a large number of animals at high marker density is a costly proposition. The implementation of GS demands cost-effective strategies for genotyping reference and selection candidates. NDDB has developed a medium-density (50K) custom genotyping chip specially designed for Indigenous cattle breeds, named INDUSCHIP in collaboration with Aarhus University, Denmark. This chip can be a good starting point for genotyping recorded cattle of various breeds in India. There is a need to evaluate designed customized chips, develop new low-density chips for imputation to higher density marker sets for Indigenous breeds (Nayee *et al*. 2018).

Genotype imputation is a powerful tool that allows the inclusion of a large number of animals genotyped with low-density arrays in the genomic evaluation without having to genotype them with more expensive HD panels. Additionally, it can be used to impute missing SNP genotypes to hybridization problems or quality control issues which will help to increase the accuracy of genomic selection (Weigel *et al.* 2010, Zhang and Druet 2010). Imputation is a cost-effective *in-silico* genotyping of missing genotypes.

Present address: [1]Kamdhenu University, Gandhinagar, Gujarat. [2]National Dairy Development Board, Anand, Gujarat. ✉Corresponding author email: thakkarrajesh151@gmail.com

MATERIALS AND METHODS

The present investigation was carried out to evaluate the performance of INDUSCHIP-1 (50K SNPs) in Gir cattle of Gujarat, and the efficiency of INDUSCHIP-1 imputation to Illumina Bovine HD (777K) level. Another objective was to select a subset of around 15K SNPs as LD panel, which can be used for imputation to INDUSCHIP-1 (50K) level to enable implementing GS at lower cost in Gir cattle.

Genotype data used in the present study was made available by NDDB. The genotype data was corrected and checked for quality control with the following criterion using PLINK software for HD as well as INDUSCHIP— (a) SNPs with a MAF < 0.01, (b) SNPs with a call rate per SNP < 0.90, (c) Animals with all SNP call rate < 0.90, (d) SNPs with a p-value 10^{-5} in the Hardy-Weinberg equilibrium, (e) SNPs that were located in non-autosomal regions. After the primary quality control, 117 Gir animals and 5,67,020 SNPs were left in the HD dataset whereas 902 Gir animals and 41,428 SNPs were left in the INDUSCHIP-1 dataset for further analysis. Reference animals used in the present imputation were not related to each other. From HD data, INDUSCHIP-1 SNPs for all 117 animals were extracted using PLINK (Purcell *et al*. 2007) software and merged with INDUSCHIP-1 data which finally resulted in 1019 animal's INDUSCHIP-1 data.

To study the efficiency of INDUSCHIP-1 to impute at HD level, five different data sets were created from this dataset by randomly selecting 12 animals as test individuals and keeping 105 animals as reference (other than test animals). The criteria were to put all the animals in different test data sets, hence every time random 20% samples were dropped and the same animal was not available in the validation sets against that test group. INDUSCHIP-1 data for test animals were extracted for each autosome (chromosome 1 to chromosome 29) separately. These data files were named test1chr1, test1chr2, test1chr3, respectively up to test1chr29 for the first test data set and so on. The HD data of corresponding chromosomes for remaining animals were used as a reference dataset and named as ref1chr1, ref1chr2, and ref1chr3 respectively up to ref1chr29. Only SNPs having MAF > 0.001 were used in reference and test files. MAF criteria for SNP selections were done to remove very less variable SNPs (Fixed SNPs). During LD panel SNP selection, SNPs with high MAF and distance between two adjacent SNPs that remain equal, only those SNPs were included. For comparison of imputation accuracy, validation data sets were generated using test individuals and HD SNPs present in reference data files. The validation files, were named as val1chr1, val1chr2 and val1chr3 respectively up to val1chr29 for comparing with imputed data to calculate concordance.

LD panel, a subset of INDUSCHIP-1 was chosen from SNPs having very high MAF and are at around 50 kbps apart. A total of 12,851 SPNs were selected. For studying the imputation accuracy of this LD panel, fivefold cross-validation was performed using 1019 animals having INDUSCHIP-1 genotype data. Out of these, each time random 15 animal's LD panel genotypes were used as test genotypes to predict their INDUSCHIP-1 genotypes using INDUSCHIP-1 genotype of the rest of the 1004 animals as reference genotypes. To study the effect of several individuals in the reference population, we used INDUSCHIP-1 genotype of 117 animals that had HD genotypes. In this scenario, there were 105 individuals in reference and 12 individuals in the test data set.

For all the three scenarios, imputation of test data sets at reference level was carried out using BEAGLE 3.3.2 (Browning and Browning 2007). BEAGLE is a software package, written in Java and runs on computing platform with a Java version 1.6 interpreter. It is used for analysis of large-scale genetic data sets, can phase genotype data, infer sporadic missing genotype data and impute ungenotyped markers that have been genotyped in a reference panel (http://faculty.washington.edu/browning/beagle/b3.html#intro).

Initially unphased chromosome wise test data sets were phased and that phased files used for creating chromosome wise imputed files using reference data sets. Concordance means imputation efficiency of software used. To know the concordance, imputed file were compared with original validation data sets. The concordance of the SNP genotypes was calculated using R statistical software (R Core Team, 2017) and was presented as % correct genotypes imputed.

RESULTS AND DISCUSSION

This section is broadly divided into two parts first part consist of the performance of INDUSCHIP-1 in comparison to commercially available SNP chips for Gir breed while the second part deals with selection of the low-density chip and study efficiency to input at a higher level.

Performance of INDUSCHIP-1 in Gir cattle population: SNPs were evenly distributed across all the chromosomes on Illumina Bovine HD (777K) as well as on the INDUSCHIP-1 (50k) level. The overall distribution of SNPs was uniform across chromosomes (Supplementary Figs 1A, 1B). The distribution of the percentage of SNPs out of HD panel ranged from 3.61 to 6.53% SNPs across

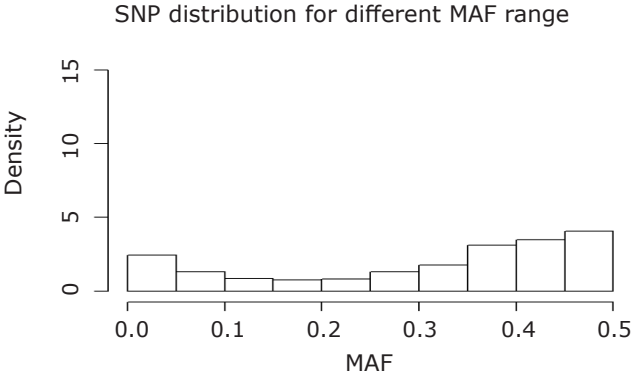

Fig. 1. Minor allele frequency (MAF) based distribution of chromosome wise SNPs available on INDUSCHIP BreadChip [50k; Medium Density (MD)].

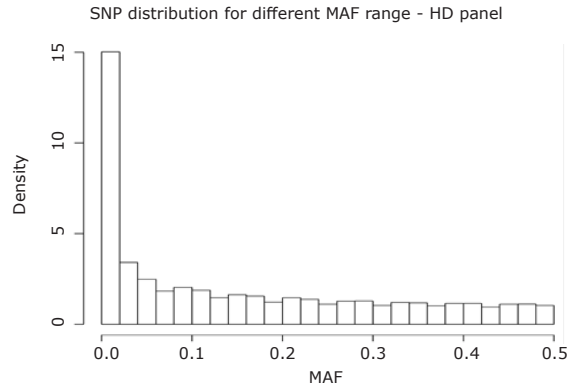

Fig. 2. Minor allele frequency (MAF) based distribution of chromosome wise SNPs available on Illumina Bovine HD BreadChip [777k; High-density (HD)].

different chromosomes. The distribution of selected SNPs across chromosomes reveals that the highest proportion of SNP selected were on chromosome no. 23 (6.53% SNPs) whereas the lowest numbers were selected for sex chromosome XY (3.65%) (Supplementary Fig. 2). The distribution was relative to chromosome size, whereas Supplementary Fig. 2 shows chromosome-wise percentage of BovineHD SNPs available on INDUSCHIP-1. The distribution is similar to commercially available Illumina Bovine SNP 50k chip (Matukumalli *et al*. 2009).

On INDUSCHIP-1, most of the chromosomal regions were covered with high MAF SNPs (Fig. 1). This means all the chromosomal regions will show considerable heterozygosity in the Gir population and thus these SNPs will be useful in studying the associations of various traits. The INDUSCHIP-1 (Fig. 1) showed a high proportion of polymorphic SNPs in the Gir breed compared to BovineHD 777K chip (Fig. 2).

The distribution of SNPs per MB ranged from 16 to 18 SNPs across different chromosomes on INDUSCHIP-1. The filtered SNPs spanned 2499.76 Mb region over all the autosomes with the longest length (158.28 Mb) on chromosome 1 and the shortest (42.8 Mb) on chromosome 25. Chromosome 1 had the highest and chromosome 26 had the lowest number of filtered SNPs.

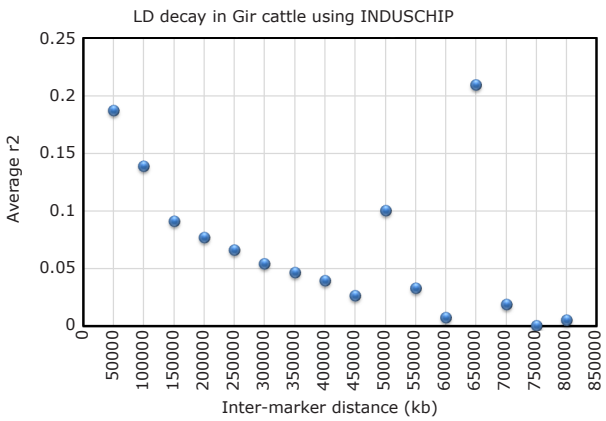

Fig. 3. LD decay shows loss of linkage when subsequent SNPs are located farther from each other on INDUSCHIP.

Table 1. Chromosome-wise distribution of the number of SNPs away from HWE (based on p-value) on INDUSCHIP-1 in comparison to Illumina Bovine HD chip

Chromosome No.	No. of SNPs away from HWE	
	Illuminae Bovine HD	INDUSCHIP-1
1	185	62
2	134	53
3	123	54
4	135	41
5	166	40
6	134	45
7	145	45
8	121	40
9	135	45
10	108	34
11	101	17
12	195	45
13	77	24
14	104	31
15	137	33
16	97	23
17	84	27
18	78	32
19	59	15
20	70	29
21	79	15
22	71	11
23	72	23
24	67	20
25	46	16
26	84	21
27	64	25
28	67	21
29	75	21
X Y	214	14
Total	3227	885

There were 885 SNPs (2.1%) away from Hardy-Weinberg equilibrium (HWE) in INDUSCHIP-1 at $p<0.00001$. The distribution of SNPs deviating from HWE in different chromosomes is given in Table 1.

It is evident from Table 2 and Fig. 3 that on average, the INDUSCHIP-1 SNPs had average Linkage Disequilibrium (LD) value of 0.154 (r^2 value) at 50 kb distance. The SNPs located around 50 kb apart (which is the median distance) among subsequent SNPs had an average r^2 value of 0.187. These results are comparable to other studies also who reported $r^2 < 0.2$ in 52 samples of Sahiwal (19), Tharparkar (17), and Gir (16) (Dash *et al*. 2017).

Effectiveness of INDUSCHIP-1 for imputation of missing SNPs at HD level in Gir cattle breed: Mean concordance of 92.33% was obtained when Illumina HD 777K was used as HD panel and INDUSCHIP-1 as LD panel for all autosomes five-fold cross-validation (Supplementary Table 1). For instance, randomly three chromosomes No 1, 17 and 23 were selected and results of these three

Table 2. Average linkage among subsequent SNPs on INDUSCHIP-1 (The SNPs are classified based on the distance between subsequent SNPs)

Inter marker distance (kb)	Count of r^2 SNPs	Average of r^2
0-50	18481	0.187
50-100	16231	0.138
100-150	3541	0.091
150-200	1187	0.077
200-250	463	0.066
250-300	165	0.054
300-350	77	0.046
350-400	31	0.039
400-450	15	0.026
450-500	13	0.100
500-550	6	0.033
550-600	2	0.007
600-650	4	0.209
650-700	2	0.0192
700-750	1	0.0005
Total	40,222	0.154

chromosome No 1, 17 and 23 during various validation rounds and overall averages is shown in Table 3 fivefold cross-validation table for INDUSCHIP-1 to HD level Imputation [for all autosomes (chromosome 1 to 29)]. The overall mean of concordance across all the chromosome in INDUSCHIP-1, including all the chromosome was 92.33%. There were no specific major regions that showed lower imputation accuracies. For example, how imputation accuracies change across the length of chromosome 1 is presented in Fig. 4.

Imputation accuracy obtained varied from chromosome to chromosome and within individuals also. The concordance ranged from 57.23 to 99.67% for different individuals in chromosome 1, 56.02 to 99.85% in chromosome 17 and 57.83 to 99.74% in chromosome 23. The highest overall concordance was observed on chromosome 1. The observed imputation accuracies are higher than reported by other researchers. Chud *et al*. (2015) observed 89.98% imputation accuracies using BEAGLE as an imputation method. However, the accuracy is lower than that reported by Ma *et al*. (2013) in a population of Swedish and Finnish Red cattle using 2931 animals in reference and 977 animals in test population. They found higher accuracies (96.7%) from the 50K to the HD chip using BEAGLE as imputation software.

The accuracies obtained in the current study are comparatively lower than reported by Boison *et al*. (2014). This can be attributed to less number of animals

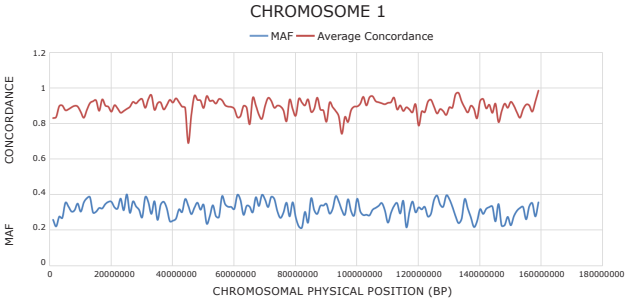

Fig. 4. Chromosomal region-wise Average MAF and Average concordance level for Chromosome 1.

in reference genotype [471 genotyped animals in the study reported by Boison *et al*. (2014), 2200 in the study reported by Júnior *et al*. (2017), 400 in the study reported by Chud *et al*. (2015), 556 in a study reported by Brøndum *et al*. (2012)] which indicates that use of related animals in reference population increases imputation accuracy. The animals used in this study were unrelated to each other hence imputation accuracies are expected to be lower in the current study compared to other studies (Boison *et al*. (2015), Júnior *et al*. (2017), Carvalheiro *et al*. (2014) used related animals as the reference population).

The selected LD chip had less number of polymorphic SNPs that results in low linkage disequilibrium, as compared to INDUSCHIP. We tried to see if we can get good imputation despite keeping low number of SNPs so as to save on cost of genotyping and genotype more animals at less cost.

Imputation accuracy from LD to INDUSCHIP-1 level: Mean concordance was 88.42% when INDUSCHIP-1 50K was used as a HD panel and selected 15K SNP panel was used as LD panel for all autosomes (Supplementary Table 2). Tables 4 and 5 showcase a fivefold cross-validation table for Selected LD to INDUSCHIP level Imputation (Chromosomes No 1, 17 and 23). Mean concordance was 85.49% when less number of reference individuals were used using INDUSCHIP-1 50K as an HD panel and with a selected LD panel (Supplementary Table 3).

The MAF along the length of chromosomes 1, 17 and 23 is shown in Supplementary Figs 3, 4 and 5. The selected panel had reasonably high MAF SNPs for various chromosomes.

The concordance ranged from 64.30 to 99.67% for different individuals in chromosome 1, 61.38 to 99.92% in chromosome 17 and 63.44 to 100% in chromosome 23. The highest overall concordance was observed on chromosome No. 23. The concordance % obtained by imputation from LD to INDUSCHIP-1 level were relatively lower than the imputation done from INDUSCHIP-1 to HD level which

Table 3. Fivefold cross-validation table for INDUSCHIP-1 to HD level Imputation

Chromosome No.	Test 1	Test 2	Test 3	Test 4	Test 5	Overall concordance
1	96.43%	93.23%	89.77%	93.53%	89.91%	92.57%
17	96.25%	93.29%	88.56%	92.72%	90.25%	92.21%
23	95.42%	92.69%	89.80%	93.68%	89.00%	92.12%

Table 4. Fivefold cross-validation table for LD to INDUSCHIP-1 level (1019 animals)

Chromosome No.	Test 1	Test 2	Test 3	Test 4	Test 5	Overall concordance
1	88.64%	89.95%	88.99%	90.76%	89.22%	89.51%
17	87.91%	89.95%	86.10%	87.82%	87.77%	87.91%
23	88.49%	88.90%	87.20%	90.17%	88.24%	88.60%

Table 5. Fivefold cross-validation table for LD to INDUSCHIP-1 level (117 animals)

Chromosome No.	Test 1	Test 2	Test 3	Test 4	Test 5	Overall concordance
1	90.18%	85.41%	83.58%	84.59%	82.90%	85.33%
17	89.36%	85.76%	80.25%	83.50%	81.74%	84.12%
23	89.56%	86.60%	84.29%	86.42%	82.32%	85.84%

could probably be due to the high effective population size of Gir cattle, and unrelated animals in the test and reference population has hampered the imputation ability. The number of SNPs in the LD panel should be increased to obtain higher imputation accuracy. However, this may be not cost-effective and the purpose of developing an LD panel may be defeated if the number of SNPs goes beyond a certain level. Imputation accuracies obtained by Mulder *et al*. (2012), Ma *et al*. (2013) were higher than that obtained in the present study, whereas those reported by Dassonneville *et al*. (2011), Jiménez-Montero *et al*. (2013), Berry *et al*. (2014), Nicolazzi *et al*. (2013), were lower when SNP panels having < 15K SNPs were used to impute SNP panel having around 50K SNPs.

Reference and validation population for different studies were 5304 and 4074 (Mulder *et al.* 2012), 2931 and 977 (Ma *et al.* 2013), 3071 and 966 (Dassonneville *et al.* 2011), 1632 and 834 (Jiménez-Montero *et al.* 2013), 2424 and 31 (Berry *et al.* 2014), 15671 and 4233 (Nicolazzi et al. 2013).

In conclusion, the INDUSCHIP-1 has high polymorphic markers across all chromosomes in Gir cattle. The distribution of MAF along all chromosomes and the length of chromosomes was found to be uniform. The imputation accuracies for imputing SNPs at HD level, obtained using INDUSCHIP-1 panel were high (92.3%), considering only 115 individuals as a reference with HD genotypes. The imputation accuracies obtained using the selected LD panel in the present work are promising. The imputation accuracies for imputing SNPs at INDUSCHIP-1 level, obtained by using selected LD panel were 88.66%, considering 1004 unrelated individuals as the reference population. There was only 3.9% reduction in imputation accuracy compared to imputation from INDUSCHIP to HD. We envisage that accuracies are likely to increase with the increase in the number of individuals in the reference population and if more related animals are genotyped and included in the further study. The study provides evidence that adopting a relatively cheaper option of LD SNP chip is feasible which would help to reduce the cost of implementing GS at field level for Gir cattle. Imputation accuracies are likely to increase with the increase in the number of individuals with the known relationship in the reference population.

ACKNOWLEDGEMENTS

Kamdhenu University, Gandhinagar, Gujarat, India and National Dairy Development Board (NDDB), Anand, Gujarat, India have collaborated to explore the imputation procedure by designing a low-density chip to make it feasible to implement genomic selection at ground level in Indigenous Gir cattle breed. I want to acknowledge and express my cordial gratitude to my guide Dr P. H. Vataliya and NDDB Animal Breeding group. Special thanks to Dr Nilesh Nayee who had advised me in the right direction so that I can always think and work better.

REFERENCES

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. *Journal of Animal Breeding and Genetics* **131**(3): 165–72.

Boison S, Santos D, Utsunomiya A, Carvalheiro R, Neves H, O'Brien A and Silva M D. 2015. Strategies for single nucleotide polymorphism (SNP) genotyping to enhance genotype imputation in Gir (*Bos indicus*) dairy cattle: Comparison of commercially available SNP chips. *Journal of Dairy Science* **98**(7): 4969–89.

Browning S R and Browning B L. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American Journal of Human Genetics* **81**(5): 1084–97.

Brøndum R, Ma P, Lund M and Su G. 2012. Short communication: Genotype imputation within and across Nordic cattle breeds. *Journal of Dairy Science* **95**(11): 6795–6800.

Carvalheiro R, Boison S A, Neves H, Sargolzaei M, Schenkel F, Utsunomiya Y, O'Brien A, Sölkner J, McEwan J, Van Tassell C, Sonstegard T and Garcia J F. 2014. Accuracy of genotype imputation in Nelore cattle. *Genetics Selection Evolution* **46**(1): 69.

Chud T C, Ventura R V, Schenkel F S, Carvalheiro R, Buzanskas M E, Rosa J O, Mudadu M de, da Silva M V, Mokry F B, Marcondes C R, Regitano L C and Munari D P. 2015. Strategies for genotype imputation in composite beef cattle. *BMC Genetics* **16:** 1–10.

Dash S, Singh A, Bhatia A K, Jayakumar S, Sharma A, Singh S, Ganguly I and Dixit S P. 2018. Evaluation of bovine high-density SNP genotyping array in indigenous dairy cattle breeds. *Animal Biotechnology* **29**(2): 129–35.

Dassonneville R, Brøndum R, Druet T, Fritz S, Guillaume F, Guldbrandtsen B, Lund M, Ducrocq V and Su G. 2011. Effect

of imputing markers from a low-density chip on the reliability of genomic breeding values in Holstein populations. *Journal of Dairy Science* **94**(7): 3679–86.

Jiménez-Montero J, Gianola D, Weigel K, Alenda R and González-Recio O. 2013. Assets of imputation to ultra high-density for productive and functional traits. *Journal of Dairy Science* **96**(9): 6047–58.

Júnior G A, Chud T C, Ventura R V, Garrick D J, Cole J B, Munari D P, Ferraz J B S, Mullart E, DeNise S, Smith S and da Silva M V. 2017. Genotype imputation in a tropical crossbred dairy cattle population. *Journal of Dairy Science* **100**(12): 9623–34.

Ma P, Brøndum R, Zhang Q, Lund M and Su G. 2013. Comparison of different methods for imputing genome-wide marker genotypes in Swedish and Finnish Red Cattle. *Journal of Dairy Science* **96**(7): 4666–77.

Matukumalli L K, Lawley C T, Schnabel R D, Taylor J F, Allan M F, Heaton M P, O'Connell J, Moore S, Smith T P, Sonstegard T S and VanTassell C P. 2009. Development and characterization of a high-density SNP genotyping assay for cattle. *PLoS ONE* **4**(4): e5350.

Mulder H, Calus M, Druet T and Schrooten C. 2012. Imputation of genotypes with low-density chips and its effect on reliability of direct genomic values in Dutch Holstein cattle. *Journal of Dairy Science* **95**(2): 876–89.

Nayee N, Sahana G, Gajjar S, Sudhakar A, Trivedi K, Lund M S and Guldbrandtsen B. 2018. Suitability of existing commercial single nucleotide polymorphism chips for genomic studies in *Bos indicus* cattle breeds and their *Bos taurus* crosses. *Journal of Animal Breeding and Genetics* **135**(6): 432–41.

Nicolazzi E, Biffani S and Jansen G. 2013. Short communication: Imputing genotypes using PedImpute fast algorithm combining pedigree and population information. *Journal of Dairy Science* **96**(4): 2649–53.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M A, Bender D, Maller J, Sklar P, de Bakker P LW, Daly M J and Sham P C. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**(3): 559–75.

R Core Team 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Weigel K, VanTassell C P, O'Connell J, VanRaden P and Wiggans G. 2010. Prediction of unobserved single nucleotide polymorphism genotypes of Jersey cattle using reference panels and population-based imputation algorithms. *Journal of Dairy Science* **93**(5): 2229–38.

Zhang Z and Druet T. 2010. Marker imputation with low-density marker panels in Dutch Holstein cattle. *Journal of Dairy Science* **93**(11): 5487–94.