Genotypes imputation as a supporting method in pedigree control in Polish Holstein-Friesian cattle

Abstract no.: 17384

Żukowski Kacper, Gurgul Artur, Rubiś Dominika, Radko Anna
National Research Institute of Animal Production, Kraków, Poland
kacper.zukowski@izoo.krakow.pl

Objectives

Utilize microsatellite (MS) and single nucleotide polymorphism (SNP) information to predict MS genotypes through imputation methods from selected SNP panels for parentage testing in cattle.

Material and Methods

- MS and SNP genotypes from <u>321 Polish Holstein-Friesian bulls</u>, the same bulls were utylize in training and validation dataset
- <u>11 MS markers</u> recommended by ICAR/ISAG for parental control testing
- SNP genotypes obtained from Illumina Bovine50SNP
- each MS were flanked by <u>10 or 20 markers</u> (Figure 1)
- <u>three models of genotypes missingness</u>: without 50% of MS (ms), without of 100% of MS (no.ms), without 50% of MS and 50% of SNP markers (snp.ms) in validation dataset
- <u>multiallelic imputation software</u>: PHASE (Stephens et al., 2001) and <u>BEAGLE</u> (Browning and Browning, 2009).

Figure 1
The idea of the panel consisting of a centrally located MS marker.

Results

Figure 2
Heatmaps of imputation accuracy between the reference and validation datasets presented across 11 chromosomes for different software, windows size and model of data missingness.
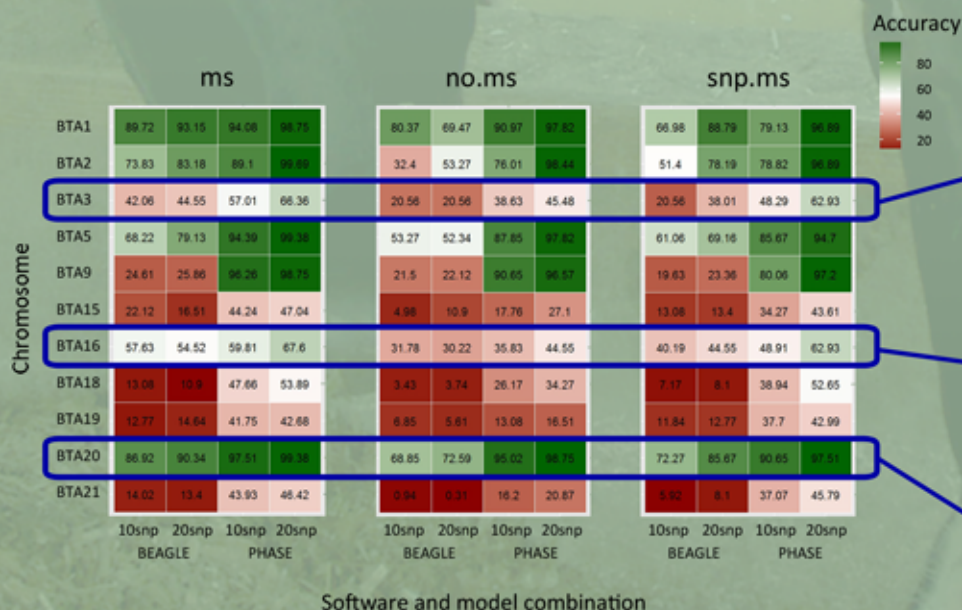

Chromosome	ms				no.ms				snp.ms			
	10snp BEAGLE	20snp BEAGLE	10snp PHASE	20snp PHASE	10snp BEAGLE	20snp BEAGLE	10snp PHASE	20snp PHASE	10snp BEAGLE	20snp BEAGLE	10snp PHASE	20snp PHASE
BTA1	89.72	93.15	94.08	98.75	80.37	69.47	90.97	97.82	66.98	88.79	79.13	96.89
BTA2	73.83	83.18	89.1	99.69	32.4	53.27	76.01	98.44	51.4	78.19	78.82	96.89
BTA3	42.06	44.55	57.01	66.36	20.56	20.56	38.63	45.48	20.56	38.01	48.29	62.93
BTA5	68.22	79.13	94.39	99.38	53.27	52.34	87.85	97.82	61.06	69.16	85.67	94.7
BTA9	24.61	25.86	96.26	98.75	21.5	22.12	90.65	96.57	19.63	23.36	80.06	97.2
BTA15	22.12	16.51	44.24	47.04	4.98	10.9	17.76	27.1	13.08	13.4	34.27	43.61
BTA16	57.63	54.52	59.81	67.6	31.78	30.22	35.83	44.55	40.19	44.55	48.91	62.93
BTA18	13.08	10.9	47.66	53.89	3.43	3.74	26.17	34.27	7.17	8.1	38.94	52.65
BTA19	12.77	14.64	41.75	42.68	6.85	5.61	13.08	16.51	11.84	12.77	37.7	42.99
BTA20	86.92	90.34	97.51	99.38	68.85	72.59	95.02	98.75	72.27	85.67	90.65	97.51
BTA21	14.02	13.4	43.93	46.42	0.94	0.31	16.2	20.87	5.82	8.1	37.07	45.79

Software and model combination

Figure 3
Linkage disequilibrium plots for SNPs localized on BTA3 (top), BTA16 (middle) and BTA20 (bottom).
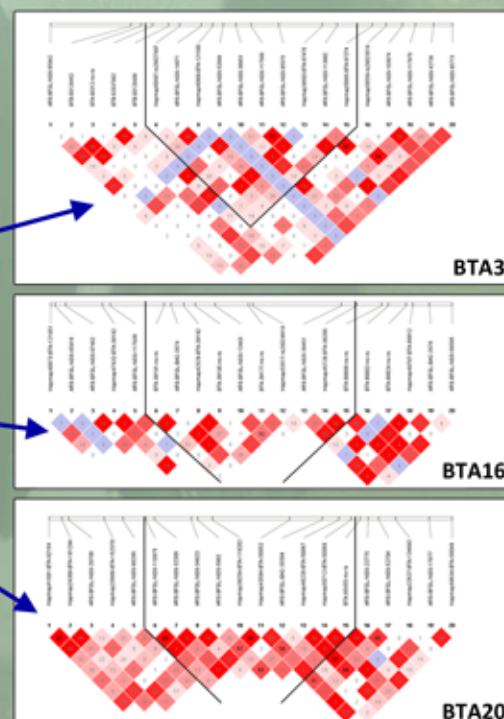
Conclusions

The PHASE imputation for multiallelic markers is more accurate, however the approach works only for chromosomes on which markers were in stronger LD blocks. It is recommended to use higher density SNP panels, for example SNPs from microarray Illumina BovineHD.