Accuracy of Genotype Imputation in Canadian Yorkshire Pigs Using FImpute Software

Alison Kinsman¹, Mehdi Sargolzaei¹-², Flavio Schenkel¹
Gordon Vander Voort¹, Mohsen Jafarkia³, Andy Robinson¹

¹University of Guelph, 50 Stone Rd E, Guelph ON Canada N1G 2W1
²L’Alliance Boviteq, 19320 Grand rang Saint-François, Saint-Hyacinthe QC Canada J2T 5H1
³Canadian Center for Swine Improvement Inc., Bldg. #54-960 Carling Ave., Ottawa ON Canada K1A 0C6

Introduction

Problem: The relatively high cost of the current 60K genotyping panel is a limitation for large scale commercial application of genomic evaluation of pigs.

Solution: Imputing genotypes from a smaller, cheaper panel to the higher density 60K panel.

Objective

Analyze factors to be taken into account when choosing a smaller panel.

Predict the most cost effective panel size.

Materials & Methods

Animals: A total of 907 Canadian Yorkshire pigs genotyped for the illumina PorcineSNP60 panel by DNA Landmarks.

SNP filtration: 39,526 SNPs out of 62,163 SNPs were used in the analysis. The following SNPs were excluded:
- 1,328 SNPs located on the sex chromosomes
- 13,985 SNPs without map location
- 7,324 SNPs with minor allele frequency of less than 1% or call rate of less than 85%

Low-density panels: Panels of 395, 6K (5,647), 10K (9,881), and 20K (19,763) evenly spaced SNPs were selected from the 60K SNP panel for testing.

Reference Populations: Three different reference population sizes (Ref. Size) of 658, 500 and 200 pigs were tested.

Validation population: A total of 249 younger pigs were used for validation of imputation accuracy from each of the low-density SNP panels. SNPs not on the low-density panel were masked.

Genotype information of parents: Scenarios were tested with neither, one or both parents genotyped for the 60K panel.

Software: All imputation scenarios were run using FImpute version 2 software (Sargolzaei et al., 2011). Fimpute accounts for population and pedigree information.

Results

Overall accuracy of panels: Accuracy of panels increased as the density of panels increased from 395 to 20K (Figure 1). A minimum of 6K SNPs on a panel was needed to achieve an imputation accuracy greater than 90%. Accuracy of imputation increased to 96% with the 10K panel.

Reference panel size: The accuracy of imputation improved by increasing the number of animals in the reference population, especially from 200 to 500 (Figure 2).

Parents genotype information: Including the 60K SNP genotypes from parents in the reference population increased the accuracy of imputation, especially for lower density panels (Figure 3).

Discussion

Selecting SNPs: SNPs in this study were equally spaced. Other methods of choosing SNPs should be tested as well.

Panel size: A 6K or 10K SNP panel might be a practical choice for the pig industry. The cost effectiveness of the panels should be evaluated based on the accuracy of genomic evaluation for different panels.

Further investigation: Results from this preliminary study should be verified using other breeds and more animals.

Implication

Large-scale application of genomic evaluation seems feasible by using imputation.

Collaboration among research groups to pool 60K genotypes could result in the development of a more effective lower density panel and more accurate imputation. Its cost could also be lowered with a higher volume of genotyping, if the lower density panel is adopt more broadly.

Genotyping parents for 60K SNPs and imputation of progeny from a lower density panel could lower the cost of routine application of genomic evaluations with minimum lost in accuracy.

Acknowledgments

Genotypes for this research were provided by the Canadian Centre for Swine Improvement. Genotypes were accumulated with financial support from Atlantic Canada Swine Centre, Ontario Swine Improvement, Centre de Développement du porc du Quebec inc., Western Swine Testing Association, PigGen Canada, Canadian Swine Research and Development Cluster, a Growing Canadian Ag-Innovation Program – Canadian Agri-Science Cluster Initiative of Agriculture and Agri-Food Canada, Agriculture Adaptation Councils of Quebec, New Brunswick, Nova Scotia, Manitoba & Ontario, le ministére de l’Agriculture, Pêcheries et de l’alimentation du Québec and Fédération des producteurs de porcs du Québec. Additional genotypes were also provided by Fast Genetics.