Project number: 6162
Funding source: Genomes Canada

Date: Nov 2014
Project dates: Nov 2010 – Nov 2014

Whole genome selection through genome wide imputation in beef cattle

Key external stakeholders:

International scientists
Geneticists
The Irish Cattle Breeding Federation (ICBF)
Irish breeding industry

Practical implications for stakeholders:

This study proved that

- Full genome sequence data can be accurately imputed from lower density genotypes using family and population based imputation algorithms; imputation accuracy is better if a multi-breed reference population is used
- The genotypes of un-genotyped parental can be accurately using genotypes from at least 6 progeny thus saving genotyping costs but also generating informative genomic information where biological material may not exist for older animals

Main results:

- The accuracy of imputing full genome sequence data from high density genotypes varied from 0.86 to 0.95 across breeds and was best when the sequences of multiple breeds were included in the reference population
- The accuracy of predicting daily gain, feed intake and feed efficiency was 0.42, 0.68, and 0.56 using genomic information from 6796 cattle.
- Where biological material or genotypes are not available on individual animals, at least five progeny can be genotyped (on either a medium or high-density genotyping platform) and the parental alleles imputed with, on average, $\geqslant$96% accuracy.

Opportunity / Benefit:

Full genome sequence data can be accurately imputed from available genotype data thereby avoiding the necessity to sequence all animals; such information could be useful to increase the accuracy of genomic predictions in future research

Collaborating Institutions:

See page two of the full Technology Update

Contact   Donagh Berry   Email: donagh.berry@teagasc.ie.

Teagasc project team: Donagh Berry (PI)

External collaborators:

University of Alberta (Dr Stephen Moore)
University of Guelph (Dr Stephen Miller)
UDSA-MARC (Dr Gary Bennett)
Victorian Department of Primary Industries (Dr Mike Goddard)
Scottish Agricultural College, UK (Dr. Mike Coffey)
University of Alberta (Dr Ellen Goddard)
University of Alberta (Dr Paul Stothard)
University of Alberta (Dr Graham Plastow)
University of Alberta (Dr Zhiquan Wang)
University of Guelph (Dr Flavio Schenkel)
AB Ag. and Rural Dev. (Dr John Basarab)
University of Laval (Dr Marc-Andre Sirard)
Beefbooster (Dr Elisa Marques)
Agriculture and Agrifood (Dr Changxi Li)

1.   Project background:

Genetic improvement is a key determinant of profitability for beef producers, with associated benefits for consumers and the community. The dairy industry, with its more simplified breed structure and sex-limited traits, has been quick to capitalise on the potential of genomic selection, typically doubling rates of annual gain. The beef industry, however, faces additional challenges because of its breed structures and the multitude of traits that affect profitability. It is widely accepted by the scientific community that to achieve the number of animals required for developing reliable genomic tests, international collaboration is essential, especially for important traits like feed efficiency, animal health and product quality. Furthermore, theoretical and methodological developments are necessary in order to develop affordable yet accurate DNA assays. In particular, there is a need to discover more specific gene tags (causal mutations) correlated to economically important traits and possibly design breed or population specific assays using knowledge of the haplotype structure within and across breeds.

- **Questions addressed by the project:**
- The extent of structural variants in the bovine genome and their prevalence in beef populations
- The accuracy of imputing full genome sequence from lower density genotypes generated using arrays
- The accuracy of imputing the genotype of non-genotyped animals from genotyped progeny
- The optimal imputation algorithms and logistics to maximise imputation accuracy to sequence
- The accuracy of genomic predictions for performance traits in growing beef cattle
- The social and economic benefits and costs of using genomic technologies in livestock improvement;

This project will deliver benefits for breeders by providing selection methodologies and it will also lay the foundation for the next generation of technologies, including improved capability for ensuring the environmental sustainability of beef and dairy cattle production.

3.   The experimental studies:

- This study was desktop based
- A total of 315 bulls were sequenced at an average coverage of 8-10X.
- In collaboration with international partners and the 1000 Bull Genomes Project, access to more than 1,000 complete genome sequences were available
- Over 10,000 animals were genotyped
- Feed intake and efficiency data were available on over 8000 growing beef cattle

4.   Main results:

- Imputation algorithms that used both family and population based information was superior at imputing higher density genotypes than algorithms that used just population-wide linkage disequilibrium; the former was also computationally faster and was superior in imputing rare alleles
- Imputation from high density to sequence was substantially more accurate than from lower density genotypes
- Accuracy of imputation from high density genotypes to sequence was greater when a multi-breed reference population was used.
- The median proportion of genotypes per animal that could not be imputed in the imputation process

decreased as the number of genotyped half-sib progeny increased; values for the medium-density panel ranged from a median of 0.015 with a half-sib progeny group size of 4 to a median of 0.0014 to 0.0015 with a half-sib progeny group size of 8.

- Genotype concordance rates increased considerably as the number of genotyped half-sib progeny increased from four (mean animal allele concordance rate of 0.94 for the medium-density genotype panel) to five (mean animal allele concordance rate of 0.96 for the medium-density genotype panel) after which it was relatively stable up to a half-sib progeny group size of eight.
- Where biological material or genotypes are not available on individual animals, at least five progeny can be genotyped (on either a medium or high-density genotyping platform) and the parental alleles imputed with, on average, $\geqslant$96% accuracy.
- Genomic predictions using a Bayesian approach were more accurate than statistical approaches that applied equal variance to each genetic marker
- Genomic predictions were improved by placing weight on the pedigree information as well as genomic information
- Genomic predictions were improved for some traits when higher density genotypes were used
- Results from the National Surveys on Consumer and Public Interest in Genomics in Beef and Dairy clearly showed that Willingness To Pay (WTP) for beef that is more feed efficient is positive and significant as well as a positive but smaller premium for more feed efficient beef produced using genomic information in selective breeding (as compared to feed efficient beef produced through other technologies

5. Opportunity/Benefit:

- Feed intake and efficiency data on growing animals from multiple locations are genetically similar and can be merged which is consistent with results in dairy
- That not animals have to be sequenced but instead sequence data can be imputed from lower density (and thus lower cost) genotypes – this can save considerable costs

6. Dissemination:

Main publications:

Stothard P, Choi J, Basu U, Sumner-Thomson JM, Meng Y, Liao X, Moore SS (2011) Whole genome resequencing of Black Angus and Holstein cattle for SNP and CNV discovery. BMC Genomics 12:559.

Grant JR, Arantes AS, Xiaoping L, Stothard P (2011) In-depth annotation of SNPs arising from resequencing projects using NGS-SNP. Bioinformatics 27:2300-2301.

Berry, D.P., McParland, S., Kearney, J. F., Sargolzaei, M. and Mullen, M. P. (2014). Imputation of ungenotyped parental genotypes in dairy and beef cattle from progeny genotypes. Animal 8:6, pp 895-903

International conferences

Presented at the several international scientific conferences including the World Congress on Genetic Applied to Livestock Production, Joint Animal Science Meetings as well as breeder conferences particularly in Canada

National conferences and seminars

Results presented at the Agricultural Research Forum as well as ICBF industry consultation days

7. Compiled by: Dr. Donagh Berry