Project number: 6589
Funding source: FIRM DAFM
Date: Dec 2016
Project dates: Dec 2013 – Dec 2015

Genomic strategies for animal and meat provenance, authenticity and traceability

Key external stakeholders:

Dairy, beef and sheep farmers, meat processors, Bord Bia, genotyping laboratories, breed associations, ICBF, breeding companies

Practical implications for stakeholders:

- A genotype panel of 800 carefully selected genetic markers, generated using either traditional "SNP chips" or genotype-by-sequencing approaches, is sufficient for detecting and, where possible, rectifying parentage errors, thereby, reducing the associated cost and computational requirement as well as opening up opportunities for using a range of different genotyping technologies
- A genotype panel of only 300 carefully selected genetic markers is required for breed assignment of biological material facilitating the use of rapid turnaround technologies
- Between 3000-6000 genetic markers are required for accurate imputation to higher density for genomic evaluations facilitating a potential reduction in cost of genotyping

Main results:

- Parentage errors for sire to offspring were approximately 17% nationally in beef cattle; where a missing or incorrect sire was identified, 15% of the time a sire could be confidently assigned to the animal in commercial herds and this increased to 60% in pedigree herds.
- Only 800 single nucleotide polymorphisms (SNPs) are required for accurate parentage assignment
- Excellent concordance (98%) was achieved between genotypes generated using traditional "SNP-chips" and genotype-by-sequencing approaches
- Using a method of selection of genetic markers developed, between 3000 and 6000 genetic markers are required to ensure accurate imputation to higher density genotypes for use in national genomic evaluations
- Using a novel developed approach for the selection of genetic markers, approximately 300 SNP are required to accurately determine the breed composition of a biological sample

Opportunity / Benefit:

Algorithms and knowledge to develop a low-cost genotyping platform for the generation of a large quantity of (value-added) information including traceability, breed assignment and genomic evaluations

Collaborating Institutions:

Irish Cattle Breeding Federation

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

Teagasc project team: Prof. Donagh Berry (PI)
Michelle Judge

External collaborators: Dr. Francis Kearney, ICBF

1. Project background:

Genomic technologies, if used properly, are highly accurate entities as testified by their use in the forensic sciences. Genomic technologies are, therefore, very suitable as enhanced analytical tools for food provenance, authenticity and traceability and in some instances are the only such tools. For example, only genomic technologies can be used to accurately ascertain the breed composition of a piece of whole meat; this is particularly important when authenticating the breed proportion of breed-specific niche markets. Although ad hoc genomic testing is done on selected samples, this sampling it is not undertaken within a highly integrated information system which facilitates linking back to the contemporaries and pedigree of the animal. Routine use of accurate low-cost genomic technologies, combined with phenotypic (e.g., number of herd movements, contemporaries) and pedigree information from the national database can be used to achieve excellent quality assurance. This may be extremely useful for export into higher value markets as well as providing extremely useful information to advance genetic gain. Parentage errors hinder genetic gain and are thought to be approximately 7.5% in Irish dairy herds.

2. Questions addressed by the project:

- The extent of parentage errors in the national beef herd and what proportion of these errors can be rectified using genomic technologies
- The usefulness and cost of genomic information in traceability and meat provenance
- The minimum number of genetic markers required for accurate parentage testing and assignment and how these genetic markers can be best selected
- The minimum number of genetic markers required for accurate breed assignment of biological materials and how these can be best selected
- The minimum number of genetic markers required for accurate imputation to higher density for genomic prediction and how these genetic markers can be best selected

3. The experimental studies:

This study was a desktop small-scale pilot study for proof of principle: 1) data on the quality of the animal recording can be made available to the processor when the animal arrives at the abattoir, and 2) a meat sample can be used to trace exactly back to a single animal. Genotypes from several hundred animals and their parents from several farms were generated as well as having access to over 0.5million genotypes from the national database. Quality measures were derived for each animal and conflicts resolved. The meat samples from animals already genotyped were also compared to ensure traceability was achieved. Two lower density genotyping panels were developed 1) to facilitate imputation to higher density genotypes with minimal loss in accuracy, and 2) for breed prediction. A third genotype panel developed, selected the minimal number of SNPs for use in parentage verification and assignment to reduce the computational resources required for routine implementation. Genotype-by-sequencing using the Illumina Tru-Seq was also used and genotypes were compared to those generated from SNP-chips. This is the first such use of this technology in cattle traceability.

4. Main results:

Parentage analysis and genotype-by-sequencing

Genotype information can be used to identify parentage errors – parentage errors were approximately 17% for sire to offspring. Where a missing or incorrect sire was identified, 15% of the time a sire could be confidently assigned to the animal in commercial herds and this can be increased to 60% in pedigree herds where a greater proportion of the candidate sires are genotyped. Rather than using the entire complement of SNPs to (in)validate pedigree, only 800 SNPs are required which reduces the computational time considerably. Genotype-by-sequencing, although requiring more research, could be a useful tool to achieve rapid genotype turnaround for a small number of SNPs at very low cost. Excellent concordance (98%) was achieved between genotypes generated using SNP chips and genotype.

Low density panels for imputation

Imputation from low to high density genotype improved at a diminishing rate as the number of SNPs included in the lower density genotype panel increased from 384 to 12,000 SNPs. Additionally, the variability in mean imputation accuracy per individual decreased as panel density increased. The method of selecting the SNPs had a major impact on mean allele concordance rate, although its impact diminished as panel density

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

increased. Imputation accuracy for SNPs selected using a combination of high SNP minor allele frequency, linkage disequilibrium structure, and relatively equal genomic distance between SNPs, outperformed all other selection methods in densities <12,000 SNPs. Using this method of SNP selection, the correlation between the imputed and actual genotypes for a 3,000 SNP panel was 0.90 and 0.96 when applied to beef and dairy populations, respectively; the respective correlations for the 6,000 SNP panel were 0.95 and 0.98. It is necessary to include between 3,000 and 6,000 SNPs in a low density panel to achieve adequate imputation accuracy to either medium density (c.a. 50,000 SNPs in the dairy population) or high density (c.a. 700,000 SNPs in the beef population) across diverse and independent populations.

Breed prediction

Panel density, SNP selection method, and the breed under investigation all had a significant effect on the correlation of actual and predicted breed proportion from the different genotype panels. Regardless of breed, an index combining the Fst and Delta methods of SNP selection numerically (but not significantly) outperformed all other selection methods in accuracy (i.e., correlation and root mean square of prediction) when panel density was ≥300 SNPs. The correlation between actual and predicted breed proportion increased as panel density increased. Using 300 SNPs (selected using the Index method), the correlation between predicted and actual breed proportion was 0.993 and 0.995 in the Angus and Hereford validation populations, respectively. When SNP panels optimised for breed prediction in one population were used to predict the breed proportion of a separate population, the correlation between predicted and actual breed proportion was 0.034 and 0.044 weaker in the Hereford and Angus populations, respectively (using the 300 SNP panel). It is necessary to include at least 300 to 400 SNPs (per breed) on genotype panels to accurately predict breed proportion from biological samples.

5. Opportunity/Benefit:

The systems and knowledge now exist for the development of low density genotype panels for a) imputation to higher density genotypes panels with minimal loss in genotype accuracy, b) for rapid parentage validation and assignment, and c) prediction of breed composition. Knowledge and experience also exists in genotype-by-sequencing approaches and interrogation of the subsequent data. The greatest impact of all approaches is the ability to reduce the cost of genotype procurement; the logistics and pipelines are currently being applied to other species.

6. Dissemination:

Results were presented at several ICBF national cattle industry consultation meetings as well as sheep consultation meetings. Results were also presented at Moorepark and Grange Open Days.

In addition to scientific, popular press articles and Open Day events, the results were also presented at meetings for the development of a national low-cost genotype platform and the genetic markers identified are now included in the panel as well as having used the algorithms developed to identify informative genetic markers.

Main publications:

Judge M. M, J. F. Kearney, M. C. McClure, R. D. Sleator, and D. P. Berry. (2016). Evaluation of developed low-density genotype panels for imputation to higher density in independent dairy and beef cattle populations. J. Anim. Sci. 94:949–962

Judge, M.M, M.M. Kelleher, J.F. Kearney, R. D. Sleator, D.P. Berry. (2016). Ultra-low density genotype panels for breed assignment of Angus and Hereford cattle. Animal. (In Press)

Judge, M.M, D. Purfield, R. D. Sleator, D.P. Berry. (2016). Impact of multi-generational genotype imputation strategies on imputation accuracy and genomic predictions. J. Animal Science (In Press)

Popular publications:

Judge, M.M., Kearney, F., McClure, M.C. and Berry, D.P. (2014). Development of low density genotype panels for dairy and beef cattle. In: 10th World Congress Genetics Applied to Livestock Production, Vancouver, Canada, 17-Aug-2014, p. 491-495

Judge, M., Berry, D.P. and McParland, S. (2015). Genomic selection in dairy cattle. TResearch 10 (2): 8-9

Berry, D.P. and Judge, M.M. (2015). High accuracy of imputation from the Irish custom bovine genotype panel to higher density genotypes. In: Agricultural Research Forum, Tullamore Court Hotel, 10-Mar-2015,

Judge, M.M., Kearney, J.F., McClure, M.C. and Berry, D.P. (2015). Low density genotype panels for dairy and beef cattle. In: Agricultural Research Forum, Tullamore Court Hotel, 10-Mar-2015, p. 146

7. Compiled by: Prof Donagh Berry

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