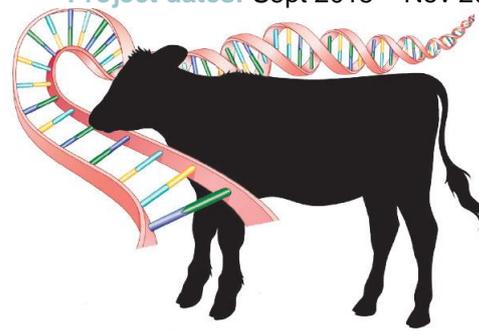**Project number:** 6787
Funding source: Science Foundation Ireland

Date: Nov 2019
Project dates: Sept 2015 – Nov 2019

Precision cattle breeding using precision genomics

Key external stakeholders:
The Irish Cattle Breeding Federation
Scientists
Stakeholders in the breeding industry
Genotyping laboratories
Farmers

Practical implications for stakeholders:
This study revealed that

- Accurate imputation to whole genome sequence is possible, but should be done stepwise from low density to medium density to high density and then to whole genome sequence; imputation from earlier to later generations, one generation at a time, also improves the accuracy of imputation.
- Routinely available genotype data can be used to generate some karyotype characteristics of animals.
- Regions of the genome associated with a whole range of performance traits were detected although, in most cases, only a small proportion of the variance was explained and thus the benefit of incorporating into genomic evaluations was limited.
- Most of the detected genomic variants associated with performance traits were unique to individual breeds thus implying minimal benefit in aiding across-breed genomic evaluations.
- More accurate relationships among individual can be derived using genomic information compared to pedigree information.
- The accuracy of the estimated genetic merit of a team of bulls is conditional on their relationships but also the evenness of their usage on the farm.

Main results:
- All performance traits evaluated are highly polygenic in nature.
- Genetic variances and correlations among morphological traits in cattle differ by breed especially when comparing Continental versus British breeds.
- Imputation of genotypes carried out in multiple steps (across genotype panel density and generations) is more accurate.

Opportunity / Benefit:
- A new and improved sire mating advice system for dairy cattle was generated and deployed; the algorithm is far more efficient than used previously and also incorporates genomic information thus making it more precise.
- Additional benefits from routinely available genotypes (i.e., parentage discovery, breed prediction, karyotyping, major gene monitoring) were documented contributing to a greater return-on-investment in genotyping.
- A decision support tool to quantify the return-on-investment in genotyping or the price that can be paid for genotyping was developed.

Collaborating Institutions:
Irish Cattle Breeding Federation, University College Dublin, Waterford Institute of Technology

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

Teagasc project team:	Donagh Berry (PI),
	Tara Carthy,
	Jennifer Doyle,
	Pierce Rafter
External collaborators:	Andrew Cromie, John McCarthy, Ross Evans and Francis Kearney (ICBF),
	Claire Gormley and Andrew Parnell (UCD),
	Siobhan Walsh (WIT)

1. Project background:

Animal breeding accounts for approximately half the gains in performance internationally. Breeding, therefore, will be a major contributor to the targeted 50% increase in milk production and 40% increase in meat value output by 2020 set down in the Irish government strategy document, Food Harvest 2020. Genomic selection in Ireland is an unsupervised approach, taking no cognisance of the underlying biology. This is simply due to a lack of precise information on the genomic variants affecting performance traits. Furthermore, cattle breeding programmes have to-date focused primarily on additive genetic effects because they are transmissible across generations. Non-additive genetic effects, nonetheless, account for a sizeable proportion of the phenotypic variance but are generally ignored in breeding programmes because they are not directly transmissible. Exploitation of non-additive genetic effects, however, can be extremely beneficial, especially if the progeny are not retained for breeding (i.e., generation of elite animals for slaughter). To-date single nucleotide polymorphisms (SNPs) are the primary genomic variants examined despite other non-SNP variants existing. Studies relating non-SNP structural variants to performance are few, or entirely lacking.

Most genome-wide association studies fail to detect many significant genomic regions (which subsequently validate in other populations) due primarily to low sample size; the addition of more animals may alleviate this. All international cattle populations implementing genomic selection exploit imputation to generate more dense genotypes. The ramifications for the precision of genomic predictions from imputed genotypes on animals with back-pedigree with imputed genotypes are not known. Moreover, the impact on accuracy of genomic predictions from a very large reference population with lower accuracy (imputed) high density genotypes/sequence (i.e., imputed from a lower density) versus a smaller reference population with more accurate (i.e., not imputed) high density genotypes/sequence is unknown. Additionally, the interaction between reference population size and genotype density on the accuracy of genomic predictions has not been established. The answers to these questions may have serious consequences for modern-day genomic selection breeding strategies. They also have implications for on-going international genomics research which also exploit imputed genotypes.

2. Questions addressed by the project:

- What is the accuracy of imputation to whole genome sequence and what is the best imputation strategy to maximise the accuracy?
- Are regions of the genome associated with performance traits common across breeds?
- What is the extent of variability in structural variants in the cattle genome and what is their association with performance traits?

- **The experimental studies:**
- This was a desktop study
- Whole genome sequence was generated on a total of 153 animals which were included in the 1000 bull's genome project providing access to the whole genome sequence of >2,000 cattle.
- Full sequence data were imputed on 621,211 animals.
- Allele intensity genotype values were available on 5,564 animals from which structural variants could be called.
- Phenotype and pedigree information was available from the ICBF.

3. Main results:

- Imputation accuracy of whole genome sequence data was 96.65% with an allele concordance rate of 98.16%.
- Allele intensity values could be used to discover karyotype abnormalities; a Turner syndrome female was detected.
- Differences in heritability of morphological traits exist between cattle breeds but not between sexes.
- Genomic regions associated with a range of performance and health traits were discovered but a) they generally only explained a very small proportion of the genetic variance and, b) were generally novel to only one breed.

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

- For copy number variants, a non-SNP variant in the genome, the mean length of duplications was 202kb and the mean length of deletions was 87kb. There was a significant inter-breed difference in the length of duplications. The mean number of deletions per animal was 87 and the mean number of duplications was 140. The majority of copy number variants were rare; 65% of deletions and 78% of duplications were present in no more than one animal in the population.
- A linear programming algorithm that can incorporate genomic information was derived to aid in mating decisions on farm.

4. Opportunity/Benefit:

- A large number of genomic variants were included on the bespoke Irish cattle genotyping platform.
- A complete revision of both the front-end and back-end of the national dairy cow sire advice system was undertaken. The linear programming algorithm was revised and made more efficient; the pedigree-based relationships were replaced with genome-based relationships.
- The algorithms to derive the mean accuracy of the estimated genetic merit of a team of bull was revised which now considers both the relationships among the bulls as well as the proportional use of each bull in the herd. This was deployed in the revised national sire advice system run by the ICBF.
- Significant genomic regions associated with performance generally do not exist across breeds thus limiting their usefulness in across-breed genomic evaluations

5. Dissemination:

International conferences

Presented at many international conferences, invited and contributed, such as the European Association of Animal Production, International Society of Animal Science, and the World Congress on Genetic Applied to Livestock Production

National Conferences and seminars

Presented at the Teagasc Dairy Open Day as well as many industry days

Main publications:

Rafter, P, D.C. Purfield, D.P. Berry, A.C. Parnell, I.C. Gormley, J.F. Kearney, M.P. Coffey, and T.R. Carthy. 2018. Characterization of copy number variants in a large multibreed population of beef and dairy cattle using high-density single nucleotide polymorphism genotype data. J. Anim. Sci. 2018.96:4112–4124
Carthy, T.R., J. McCarthy, and D. P. Berry. 2019. A mating advice system in dairy cattle incorporating genomic information J. Dairy Sci. 102:8210–8220
Bouwman A.C., …., D.C. Purfield, D.P. Berry, … and B.J. Hayes. 2018. Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. Nature Genetics. 50: 362-367.

6. Compiled by: Dr. Donagh Berry

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