

Project number: 6458
Funding source: H2020

Date: May 2012
Project dates: May 2012 – May 2016

Global cooperation to develop next generation whole genome SEQUENCE SELECTION tools for novel traits



Key external stakeholders:

Dairy, beef and sheep farmers, scientists, genotyping laboratories, breed associations, Irish Cattle Breeding Federation, breeding companies

Practical implications for stakeholders:

- The first international database for daily feed intake and associated genotypic information in lactating dairy cows was developed and used to generate the first ever international genetic and genomic evaluations for feed intake.
- The extent of genotype-by-environment interaction for feed intake in dairy cows across 9 countries was quantified.
- The ability of mid-infrared spectroscopic (MIR) analysis of milk to predict feed efficiency was quantified.
- The usefulness of a national genetic evaluation for MIR-predicted feed intake and efficiency in segregating animals divergent for the gold standard equivalent of these traits was proven.
- Pipelines were developed for the efficient and effective translation of raw sequence reads to genotype data that can be used in down-stream genomic analyses.
- The generated sequence data ensures that Ireland became a member of the 1000 genomes project, providing full access to whole genome sequence data on >2500 animals to increase the accuracy of across-breed and trans-generational genomic predictions

Main results:

- Combining harmonised feed intake and genomic data for lactating dairy cows and heifers from 10 different populations in 9 different countries contributes to an increase in accuracy of genetic and genomic evaluations.
- Genotype-by-environment interactions exist for feed intake between countries with intensive management systems versus grazing systems such as in Ireland.
- It is possible to predict, with reasonable accuracy, feed intake in lactating dairy cows from routinely available mid-infrared spectral data from individual cow and bulk milk samples.
- Based on the whole genome sequence data of 2,333 influential animals of multiple *Bos Taurus* breeds, there are over 55 million mutations.

Opportunity / Benefit:

Algorithms and knowledge to accurately impute whole genome sequence data on >0.5 million animals of multiple breeds which may be useful in increasing the accuracy of pan-breed and a trans-generational genomic prediction.

Collaborating Institutions:

Wageningen University, SRUC, Department of Primary Industries and Resources South Australia, Massey University, DairyNZ

Teagasc project team:	Prof. Donagh Berry (PI) Dr. Michelle Judge Dr. Sinead McParland
External collaborators:	Prof. Roel Veerkamp, Wageningen University Dr. Yvette de Hass Wageningen University Prof Eileen Wall, SRUC Prof Mike Coffey, SRUC Dr Jennie Pryce, DEPI Prof John Roche, DairyNZ Prof Nicolas Lopez-Villalobos, Massey University

1. Project background:

The world's human population is expected to grow to 9.1 billion by 2050. To satisfy the 70% increase in food demand from a constant or, likely, shrinking agricultural land base, is a real and significant global challenge. This challenge can only be met through greater efficiencies of production in both plants and animal production with no or minimal adverse environmental consequences. Dairy products are, and will remain, a significant proportion of the staple human diet as a dietary source of essential amino acids, minerals and vitamins. To-date, most research on feed intake and utilisation in cattle has been undertaken in beef cattle. Factors hindering research on feed intake and efficiency in dairy cattle include: 1) lack of sufficient data within country and 2) knowledge of the most relevant trait to reflect feed efficiency in dairy cattle. The objective of the SEQSEL project was to build upon, solidify and formalise relationships between several world-leading research partners in the EU (Ireland, The UK, and the Netherlands) and Oceania (i.e., Australia and New Zealand) in order to conduct joint research and participate in knowledge exchange and to 1) develop the logistics for sharing data, 2) discuss the merits of alternative definitions of feed efficiency, and 3) identify regions of the bovine genome associated with feed intake and efficiency.

2. Questions addressed by the project:

- Is it possible to combine feed intake data on lactating dairy cows from different sources?
- What is the usefulness of combining data sources to increase the accuracy of selection for difficult to measure traits, using feed intake as a proxy?
- What are the phenotypic and genetic correlations between feed intake, efficiency and production traits in lactating dairy cows?
- How many total mutations exist in the bovine genome and what type are they?
- Is the routinely available milk MIR data useful to predict the feed intake complex in lactating dairy cows?

3. The experimental studies:

This project was an International Research Staff Exchange Scheme facilitating the funding of staff secondments to learn new skillsets and develop collaborative links. All research was desktop based. The largest database in the world on feed intake in dairy cows consisting of 224,174 feed intake records from 10,068 lactations on 6,957 cows was collated; the data originated from 10 populations in 9 countries; all data originated from previously collected data from multiple experiments in the different countries. Influential Holstein-Friesian AI sires in Ireland were also identified and sequenced and added to the 1000 bulls' genome database which now contains 2,333 animals. Over 50 million sequence variants have been identified in these animals.

4. Main results:

International genetic evaluation for feed intake in dairy cows

The heritability of feed intake across all countries was 0.34 indicating that 34% of the differences in feed intake of dairy cows was due to genetics. The genetic correlation between feed intake in cows and growing heifers was 0.67 suggesting that feed intake in growing heifers is largely under similar genetic control to feed intake in cows offering a mechanism to measure feed intake in heifers as a proxy for cows and also at an earlier age. Sometimes the effect of genes on the expression of a trait can vary by the environment the animal is producing in – this phenomenon is called genotype-by-environment interactions. The presence of genotype-by-environment interactions for feed intake across the 10 different populations was examined and did not exist with the exception of the genes underpinning feed intake in confinement cows being different to feed intake in grazing cows. This limits the potential benefit of sharing feed intake data on cows on TMR versus grazing diets. The optimal scenario of genomic prediction based on shared data, resulted in a mean prediction accuracy of 0.44 across all countries, ranging from 0.37 (Denmark) to 0.54 (the Netherlands). If no feed intake records were available in a country, the accuracy based on the data from the other populations

ranged from 0.23 to 0.53. The conclusions from the present study are that there was clearly a benefit in collaboration, because phenotypic information for feed intake from other countries can be used to augment the accuracy of genomic evaluations of individual countries

Alternative measures of feed efficiency in dairy cows and their inter-correlations

A total of 13 alternative definitions for feed efficiency were derived. Absolute correlations between the residual efficiency traits defined with energy balance were strong and varied from 0.70 to 0.81. The absolute correlations between the feed efficiency traits defined as ratio traits with energy balance were also strong, varying from 0.82 to 0.99. The strengths and weaknesses of each feed efficiency trait were identified but the lack of unity genetic correlations among traits suggest that each trait is describing a different characteristic of feed efficiency.

Predicting the feed intake complex from milk MIR

The accuracy of predicting (i.e., correlation between predicted and actual values) energy intake, energy balance and residual energy intake from milk mid infrared spectroscopy in dairy cows was up to 0.88, 0.78 and 0.63, respectively, based on cross-validation. Genetic evaluations for the predicted trait was able to segregate animals when the gold standard feed intake of these animals was eventually available.

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) 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 several invited talks at international conferences (with associated invited reviews); outcomes were also presented at the Moorepark Open Days

Main publications:

Hurley, A.M., Lopez-Villalobos, N, McParland, S., Kennedy, E., Lewis, E., O'Donovan, M., Burke, J.L. and Berry, D.P. (2016). Inter-relationships among alternative definitions of feed efficiency in grazing lactating dairy cows. *Journal of Dairy Science* 99:468-479

Judge, M.M Purfield, D.C., Sleator R.D. and Berry, D.P. 2017. The impact of multi-generational genotype imputation strategies on imputation accuracy and subsequent genomic predictions. *J. Anim. Sci.* .95:1489–1501

McParland, S. and Berry, D.P. (2016). The potential of fourier transform infrared spectroscopy of milk samples to predict energy intake and efficiency in dairy cows. *Journal of Dairy Science* 99:4056-4070

7. Compiled by: Prof Donagh Berry