

Enhancing the data pipeline for novel traits in the genomic era: From farms to DHI to evaluation centres

<1 line>

F. Miglior^{1,2}, and J. Chesnais³

<1 line>

¹ *Canadian Dairy Network, 660 Speedvale Avenue West, Suite 102, Guelph, Ontario, N1K 1E5*

² *CGIL, Dept. of Animal Biosciences, University of Guelph, Guelph, ON, Canada N1G 2W1*

³ *Semex Alliance, Guelph, Ontario, Canada*

<2 lines>

Abstract

<1 line>

Sophisticated statistical methods are used to predict the genetic merit of individuals. By utilizing millions of performance records from daughters, parents and other relatives, genetic and environmental effects are well separated to predict the genetic worth of each animal. Their application has been very successful in dairy cattle for the improvement of easily measured traits. The success of this approach is mainly due to the feasibility of recording traits from thousands of herds across the country. However, this strategy is not feasible for economically important traits that are costly to measure. The advent of genomics provides an exceptional opportunity to genetically improve the national herd for those key novel traits that are more costly to record: health, animal welfare, enhanced product quality, feed efficiency and methane emissions.

In order to provide routine evaluation services for novel traits, a cost effective data pipeline needs to be developed to transfer new data from farm to DHI and from DHI to genetic evaluation centres. In Canada, several large research projects have focused on a series of novel traits that affect profitability of dairy producers and industry: eight diseases, hoof health, nutritional value of milk, feed efficiency and methane emissions. Generally, the projects include various steps: a) targeting of those novel traits that affect profitability; b) identification of predictors that have the potential to increase genomic EBV reliability of targeted traits; c) development of a cost effective pipeline to move new phenotypes from herds to DHI, from DHI to CDN data base, from CDN to R&D; d) development of genetic and genomic evaluations; e) development of sub-indexes for targeted selection; f) genome wide association and functional studies to identify SNP of interest to add to revised SNP panels in order to increase accuracy of genomic selection.

Up to now, 40% of dairy producers collect health events and their data is collected by DHI and analyzed at CDN. More recently, a new data pipeline has been developed to retrieve hoof lesions data from tablets of participating hoof trimmers. Trimmers receive DHI herd inventory, whereas their detailed data for each trimming session is sent to Canadian DHI, which is then transferred to Canadian Dairy Network. Another key pipeline has been developed to store milk mid-infrared spectra from two FOSS lines in two DHI labs. The accumulated MIR spectra records are then routinely transferred to Canadian DHI and then to CDN. By the end of 2016, the remaining ten FOSS lines will be enabled to store MIR data as well, so that we will have a MIR spectra record for every analyzed DHI milk vial in Canada. Furthermore, through a large international genome project a data base is under development to retrieve individual animal feed intake and methane emissions for participating partners. Data will then routinely collated and exchanged among all international partners.

On-going research is set to release a wealth of new genetic evaluation services as a consequence of new data pipelines. CDN has implemented genetic and genomic evaluations for mastitis resistance

in August 2014. Development is in progress to implement genetic and genetic evaluations for a series of novel traits within the next 5 years: ketosis and subclinical ketosis (BHB) together with displaced abomasum in the new Metabolic Disease Resistance index (December 2016); metritis and retained placenta in the new Fertility Disorder Resistance index (2017); hoof infectious lesions and horn lesions together with lameness in the new Hoof Health index (2018); milk fatty acid profile (2018); feed efficiency and methane emissions (2019). Research projects have been key to develop new data pipelines in order to implement new genetic and genomic evaluation systems for these novel traits in the coming years.

<1 line>

Keywords: Genomic selection, novel traits, data pipelines