

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

*F. Miglior<sup>1,2</sup>, R. Finocchiaro<sup>3</sup>, F. Malchiodi<sup>2</sup>, A. Fleming<sup>2</sup>, L. F. Brito<sup>2</sup>, C. F. Baes<sup>2</sup>, P. Martin<sup>2</sup>, J. Chesnais<sup>4</sup>, and F. S. Schenkel<sup>2</sup>*

<sup>1</sup> *Canadian Dairy Network, 660 Speedvale Avenue West, Suite 102, Guelph, ON, Canada*  
[Miglior@cdn.ca](mailto:Miglior@cdn.ca) (Corresponding Author)

<sup>2</sup> *CGIL, Dept. of Animal Biosciences, University of Guelph, Guelph, ON, Canada*

<sup>3</sup> *Italian Holstein Association, ANAFI, Cremona, Italy*

<sup>4</sup> *Semex Alliance, Guelph, Ontario, Canada*

## Abstract

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 genomic 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); Digital Dermatitis (December 2017); feed efficiency and methane emissions (by end of 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.

*Keywords: Genomic selection, novel traits, data pipelines*

## **Introduction**

The collection of new traits, the ability to have a single, central, collection and exchange of data among all stakeholders involved in the genetic improvement (producers, milk recording organizations, AI organizations, evaluation units and research centers) is very important in this new context, aimed to enhancing animal health and welfare. In Canada herds are distributed throughout the national territory, with a greater concentration in Quebec (49%) and in Ontario (33%) and to a lesser density in the western provinces (13%) and eastern Canada (6%). The national dairy herd includes approximately one million dairy cows from 11,200 herds (Agriculture and Agri-Food Canada, 2017). The milk recording is managed exclusively by Canadian DHI (Canwest DHI, Guelph, ON for Ontario and western Canada; Valacta, Sainte-Anne-de-Bellevue, QC, for Quebec and eastern Canada), whereas the genetic evaluation is performed by Canadian Dairy Network (CDN, Guelph, ON). The centralization of data collection includes various stakeholders: milk recording agencies, breed associations, AI organizations, CDN and to a lesser extent veterinarians and hoof trimmers, who currently still share limited data.

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.

## Health Traits

In Canada, a national dairy cattle health and disease data management system was started in 2007 for 8 diseases: mastitis, displaced abomasum, ketosis, milk fever, retained placenta, metritis, ovarian cysts and lameness (CDN, 2006). Recording of disease cases is done by producers and partly by veterinarians. About 40% of milk recorded farms has voluntarily adhered to this system; farmers are provided with all relevant information in order to identify and record correctly the pathology, then the data is collected by DHI technicians and transmitted to Canadian DHI and then to CDN (Koeck et al., 2012b; Neuenschwander et al., 2012). The result of this data collection has been the development of mastitis resistance index, index resistance to metabolic diseases, and in the near future the index resistance to fertility disorders.

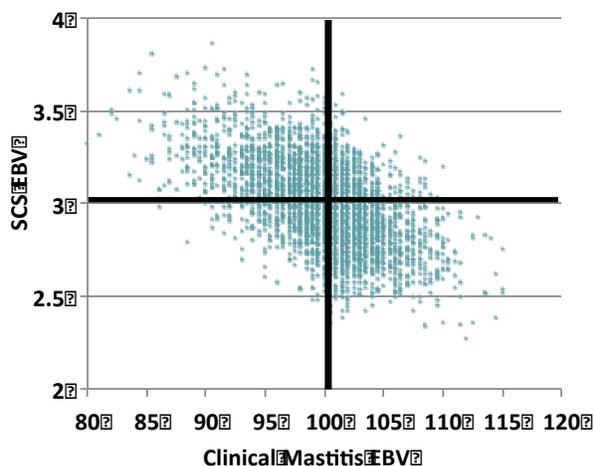
### Mastitis Resistance (MR)

In August 2014, a routine evaluation for Mastitis Resistance was officially implemented for Holstein, Ayrshire and Jersey breeds in Canada. The multiple-trait evaluation includes mastitis data and related traits with higher heritability as, “somatic cells”, “fore udder attachment”, “udder depth” and body condition score (BCS) (Koeck et al., 2012a; Koeck et al., 2012c). The newly developed index Mastitis Resistance (MR) gives equal weights to clinical mastitis (CM) of first and second parities and somatic cell score (CDN, 2013; Jamrozik et al., 2013):

$$\text{MR} = 1/3 \text{ CM-primiparous} + 1/3 \text{ CM-multiparous} + 1/3 \text{ somatic cell score}$$

Thanks to genomics, over 2,600 bulls are used as a reference for this trait, bringing an increase in the accuracy of the index, +8 points for proven bulls but especially for the young bulls with +16 points. It was decided to consider both the clinical mastitis traits (presence or absence of mastitis), and SCS because the first is an indicator of clinical mastitis, whereas the second is an indicator for subclinical mastitis. The two traits are well correlated, but as shown in Figure 1, the relationship is not equal to 1 and the best bulls are those with high index for mastitis resistance and low index for somatic cell score (Miglior et al., 2014).

*Figure 1. Distribution of proven sires for Mastitis Resistance (SCS vs. clinical mastitis EBV).*



## **Metabolic Diseases Resistance (MDR)**

In December 2016, a genetic evaluation for clinical (CK) and subclinical ketosis (SCK) and displaced abomasum (DA) was officially implemented for Holstein, Ayrshire and Jersey breeds in Canada (Koeck et al., 2015; Jamrozik et al., 2016b; Koeck et al., 2016). Clinical ketosis and DA are producer recorded traits, whereas SCK is based on the laboratory testing of milk samples for levels of BHB (beta-hydroxybutyrate), a service provided by DHI. Each of these traits is evaluated separately for cows in first lactation compared to later lactations. The overall resulting index for metabolic diseases resistance, combines the genetic evaluations of these traits into a single value. To improve the accuracy of the index, the genetic evaluation system also includes two related traits, fat and protein in early lactation and the BCS in first lactation. The relative weight of each trait is 50% for subclinical ketosis, 25% for clinical ketosis and 25% for displaced abomasum:

$$\text{MDR} = 0.50 \text{ SCK} + 0.25 \text{ CK} + 0.25 \text{ DA}$$

## **Fertility Disorders Resistance**

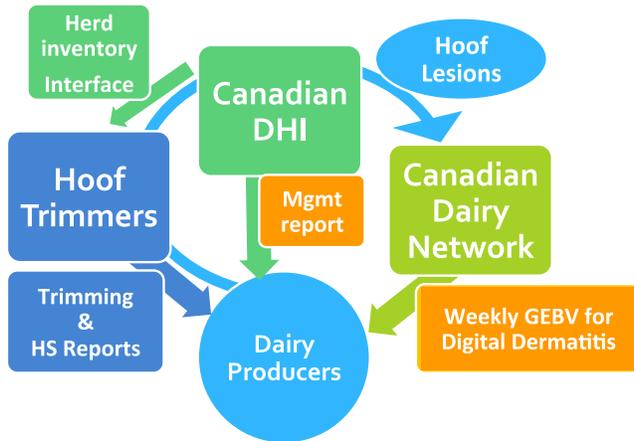
Stemming from the same source of national on-farm data collection of health events mentioned above, genetic evaluations are under development for resistance to fertility disorders, including Metritis and Retained Placenta (Jamrozik et al., 2016a). Work is in progress to identify indicator traits that are well related to fertility disorders and with moderate heritability. In the meanwhile, the Daughter Fertility (DF) index was modified in December 2016 to include a total of six traits, namely Age at First Service (11%), Non-Return Rate in heifers (16%), First Service to Conception in heifers (8%), Calving to First Service (15%), Non-Return Rate in cows (34%) and First Service to Conception in cows (16%).

## **Hoof Health**

Forty to 70% of the animals in North America and Europe has at least one hoof lesion. In addition to compromising the welfare of animals this type of disease can cause economic losses related to costs incurred for the treatment and reduction of animal performance (production and fertility). Correlations between hoof lesions and conformation feet and legs traits are low and for this reason it is essential to have a direct data collection from trimmers in order to improve hoof health both at the management and genetic level (Chapinal et al., 2013). In Canada, a new data exchange system has been set up for this type of data collection. Over 40 trimmers have been recruited. Via a tablet, they are sending data related to hoof lesions to DHI (data are then transferred to CDN), while at same time the trimmers download herd inventories before visiting the next herd. Figure 2 includes a flow chart that illustrates the new data pipeline.

Table 1 shows the hoof lesions and their frequency and heritability (Malchiodi et al., 2017). Digital Dermatitis has a highest prevalence in Canada compared to other pathologies and will be the first trait to be evaluated. The goal is to have an official genetic evaluation by December 2017 using a Single Step Genomic evaluation.

**Figure 2. New data pipeline for hoof lesions from trimmer to DHI and CDN.**

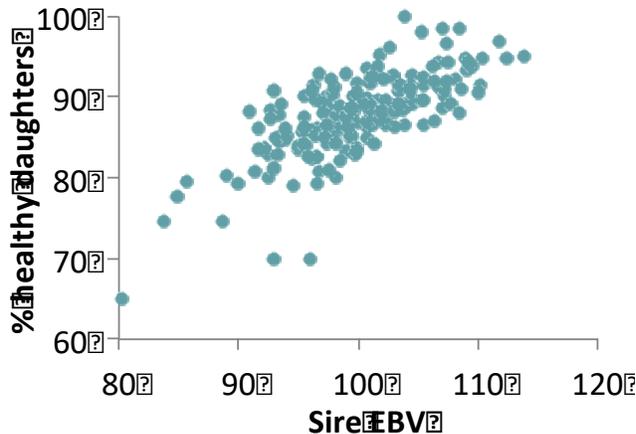


**Table 1. Hoof lesion frequencies and estimated heritabilities.**

Traits	%	h <sup>2</sup>
Digital Dermatitis	14.0	0.05
Interdigital Dermatitis	1.3	0.01
Interdigital Hyperplasia	1.4	0.03
Sole Hemorrhage	3.4	0.01
Sole Ulcer	4.9	0.03
Toe Ulcer	1.1	0.01
White Line	3.8	0.01

Heritability is low, however, the genetic variance among sires is large. Figure 3 illustrates clearly how best sires for resistance to Digital Dermatitis have most healthy daughters when compared to bottom sires.

**Figure 3. Distribution of sire EBV by percentage of healthy daughters for Digital Dermatitis.**



## **Additional Novel Traits**

A large project currently under way in Canada has the overall objective to improve the feed efficiency (FE) and reduce methane emissions (ME) in dairy cows using genomics (<http://genomedairy.ualberta.ca>). This project involves the creation of a worldwide database and an exchange with partners involved (Canada, UK, USA, Australia and Switzerland, with three additional European countries interested in participating: Denmark, France, and the Netherlands). By selecting animals more feed efficient and with lower methane emissions we will have an opportunity decrease the environmental impact from dairying, as well as having benefits for producers, and reducing costs related to feed. The same project also has the objective of saving all milk spectral records from DHI laboratories. Currently only two of twelve FOSS lines can store the milk spectral records. By receiving data from the remaining 10 spectrometers in DHI laboratories in the Canadian territory, we will get information for all cows under the milking recording national system. Milk spectral data can then be used as predictors for energy balance, methane emissions, and other fine milk components.

## **Conclusions**

In order to remain competitive, it is key to include novel traits in breeding programs. To achieve this goal it is necessary to have a female reference population based on genotyped animals already phenotyped for the novel traits. The priority is to maximize the genetic progress for new traits expanding female population reference, identifying the best farms to be sampled and genotyped large number of cows.

Genomics offers benefits quickly, high efficiency, fast uptake and doubled genetic progress doubled. However, it is essential to have a data collection system and an efficient data exchange for the recording of new phenotypes and genotypes. Include the cows in the reference population provides reliable genomic evaluations for costly new phenotypes.

## **Acknowledgements**

All dairy producers participating to research projects are gratefully acknowledged. We acknowledge partial funding by the DairyGen Council of Canadian Dairy Network (Guelph, Ontario, Canada) and the Natural Sciences and Engineering Research Council of Canada (Ottawa, Ontario, Canada). We acknowledge partial funding by a contribution from the Dairy Research Cluster Initiative (Dairy Farmers of Canada, Agriculture and Agri-Food Canada, the Canadian Dairy Network, and the Canadian Dairy Commission). We acknowledge partial funding by the Efficient Dairy Genome Project, funded by Genome Canada (Ottawa, Canada), Genome Alberta (Calgary, Canada), Ontario Genomics (Toronto, Canada), Alberta Livestock and Meat Agency (Edmonton, Canada), Ontario Ministry of Research and Innovation (Toronto, Canada), Ontario Ministry of Agriculture, Food and Rural Affairs (Guelph, Canada), Canadian Dairy Network (Guelph, Canada), GrowSafe Systems (Airdrie, Canada), Alberta Milk (Edmonton, Canada), Department of Development, Jobs, Transport and Resources (Australia), Scotland's Rural College (Edinburgh, UK), USDA Agricultural Research Service (United States), Qualitas AG (Switzerland).

## References

- Agriculture and Agri-Food Canada. 2017. Canadian Dairy Information Centre (CDIC). Dairy Facts and Figures. [http://www.dairyinfo.gc.ca/index\\_e.php?s1=dff-fcil](http://www.dairyinfo.gc.ca/index_e.php?s1=dff-fcil). Accessed on April 2017.
- Canadian Dairy Network. 2006. CDN Articles: Is Animal Health Important to You? <https://www.cdn.ca/document.php?id=124> Accessed on April 2017.
- Canadian Dairy Network. 2013. CDN Articles: Official Genetic Evaluation for Mastitis Resistance <https://www.cdn.ca/document.php?id=333> Accessed on April 2017.
- Chapinal, N., A. Koeck, A. Sewalem, D. F. Kelton, S. Mason, G. Cramer, F. Miglior. 2013. Genetic parameters for hoof lesions in Canadian Holstein cows estimated from hoof trimming records. *J. Dairy Sci.* 96:2596–2604.
- Jamrozik, J., A. Koeck, F. Miglior, G. Kistemaker, F. Schenkel, D. Kelton, B. Van Doormaal. 2013. Genetic and Genomic Evaluation of Mastitis Resistance in Canada. *Interbull Bulletin* No. 47: 43-51
- Jamrozik, J., A. Koeck, G.J. Kistemaker, and F. Miglior. 2016a. Genetic Parameters for Metabolic Disease Traits and Fertility Disorders in Canadian Holsteins. *J. Dairy Sci.* 99:1990–1998.
- Jamrozik, J., G. Kistemaker, B. Van Doormaal, A. Fleming, A. Koeck, and F. Miglior. 2016n. Genetic Evaluation for Resistance to Metabolic Diseases in Canadian Dairy Breeds. *Interbull Bulletin* No. 50:9-16
- Koeck, A., F. Miglior, D. F. Kelton, and F. S. Schenkel. 2012a. Alternative somatic cell count traits to improve mastitis resistance in Canadian Holsteins. *J. Dairy Sci.* 95:432–439.
- Koeck, A., F. Miglior, D. F. Kelton, and F. S. Schenkel. 2012b. Health recording in Canadian Holsteins: Data and genetic parameters. *J. Dairy Sci.* 95:4099–4108.
- Koeck, A., F. Miglior, D. F. Kelton, and F. S. Schenkel. 2012c. Short communication: Genetic parameters for mastitis and its predictors in Canadian Holsteins. *J. Dairy Sci.* 95:7363–7366.
- Koeck, A., J. Jamrozik, G.J. Kistemaker, F.S. Schenkel, R.K. Moore, D.M. Lefebvre, D.F. Kelton, and F. Miglior. 2015. Development of Genetic Evaluations for Metabolic Disease Traits for Canadian Dairy Cattle. *Interbull Bulletin* No. 49:76-79
- Koeck, A., J. Jamrozik, G. J. Kistemaker, F. S. Schenkel, R. K. Moore, D. M. Lefebvre, D. F. Kelton, and F. Miglior. 2016. Genetic and phenotypic associations of milk  $\beta$ -hydroxybutyrate with ketosis in Canadian Holsteins. *Can. J. Anim. Sci.* Vol. 96, No. 3 : 302-305
- Malchiodi, F., A. Koeck, S. Mason, A.M. Christen, D.F. Kelton, F.S. Schenkel, F. Miglior. 2017. Genetic parameters for hoof health traits estimated with linear and threshold models using alternative cohorts. *J Dairy Sci.* 100: 2828-2836
- Miglior, F., A. Koeck, J. Jamrozik, F. Schenkel, D. Kelton, G. Kistemaker, B. Van Doormaal. 2014. Index for Mastitis Resistance and Use of BHBA for Evaluation of Health Traits in Canadian Holsteins. *Interbull Bulletin* No. 48: 73-78.
- Neuenschwander, T. F.-O., F. Miglior, J. Jamrozik, O. Berke, D. F. Kelton, and L. R. Schaeffer. 2012. Genetic parameters for producer-recorded health data in Canadian Holstein cattle. *Animal* 6:571–578.