

Réseau laitier canadien



Canadian Dairy Network

The Magic of “Imputation”

The technology of genomics continues to evolve rapidly. While the methods used by Canadian Dairy Network (CDN) to compute genomic evaluations have reached a level of consistency, significant development has occurred in terms of the technology available for genotyping animals. For this reason, there are two new genotyping panels to be released by Illumina in the near future. This article briefly describes the coming technology and how it will be used for genetic evaluations in Canada.

Different Panels for Genotyping

In Canada, North America and most of the rest of the world, genotyping of dairy animals has been done using a panel produced by Illumina that has roughly 58,000 markers equally spaced across the 30 chromosomes that make up the bovine genome. For simplicity, this panel has been termed the “50K panel” and, so far, over 45,000 Holsteins have been genotyped in North America using this panel.

In July or August 2010, Illumina is expected to release a lower density panel for commercial use to genotype dairy animals. In Canada, Holstein Canada will add this level of genotyping to the services already provided using the 50K panel, likely in conjunction with industry partners. The major difference to breeders will be the reduced cost of under \$50 for genotyping using this panel consisting of 3072 markers (or SNPs). Known as the “3K panel”, genotyping of heifers and cows is expected to rapidly expand to include a broader group within the Canadian Holstein population rather than just the elite. In developing this 3K panel, researchers identified the optimal subset of the markers already included on the 50K panel, therefore maintaining a consistent group of about 3,000 markers for all genotyped animals.

In the coming months, Illumina is also expected to release a new panel to be termed the “Bovine HD” panel, where HD refers to high density. Details presently available indicate this panel will include roughly 850,000 markers, which is 17 times more than the current standard 50K panel. With 850,000 equally spaced markers spread across the entire bovine genome the distance between each one is significantly reduced compared to using only 50,000 markers. In terms of genotyping cost, the 850K panel is expected to cost more than the current price for the 50K panel but exact rates are yet to be finalized by Illumina.

Why the Three Different Panels?

What are the benefits of having the 3K, 50K and 850K (Bovine HD) panels all available for genotyping dairy animals? For the 3K panel, the most obvious answer to this question is price since genotyping with it will be far more economical for large scale genotyping compared to the costs for either the 50K or the Bovine HD panels. Although the Bovine HD panel will surely be priced higher than the 50K panel, the huge jump in number of markers is expected to give greater accuracy in the resulting genomic

evaluations. For this reason, the planned strategy for using these three different panels includes genotyping very large numbers of heifer and bull calves with the 3K panel as a tool for basic population screening; genotyping A.I. sires and cows with the 50K panel; and genotyping key reference proven sires in the population with the Bovine HD panel to ensure the maximum accuracy possible when estimating the effects of the different markers on each trait of interest.

Genotypes and Haplotypes

The DNA in each cell of an animal is made up of a long series of molecules, known as nucleotides. In fact, it is structured with two parallel strings that are connected at each molecule, forming the double helix structure well known to represent DNA. Across the entire bovine genome, the string of paired molecules counts more than 3 billion in total and together they form the estimated 30,000 genes in dairy cattle. There are only four nucleotides that make up DNA, which are labelled as A, C, G or T. In the double helix structure of DNA, these molecules are paired in combinations such as AA, AC, AG, AT, CA, CC, CG, CT, etc... An animal's genotype refers to the pairs of molecules inherited by the animal from its parents at specific points within its DNA. One molecule is inherited from the sire and the other from the dam but simply knowing the genotype does not identify the parental source of each DNA molecule.

However, each parent transmits its DNA to its progeny in very large blocks, which gradually get broken up over many generations. Examining the genotypes of many related animals, as is done with the 50K panel currently, identifies the blocks of DNA that remain constant across the population and transmit from parent to progeny. These blocks are called haplotypes and once most have been identified the genotype of an animal can be used to determine which haplotypes were inherited from each parent.

Imputing Genotypes

Given the availability of tens of thousands of genotypes based on the 50K panel, researchers have been able to determine the most frequent haplotypes in a given breed population. Genotypes from the 3K panel can be used to identify the DNA blocks an animal received from its sire and dam, which helps rebuild the animal's 50K genotype. This process is referred to as "imputation". The major advantage of imputation is that many animals can be genotyped using the lower cost 3K panel but if there are sufficient other animals, especially family members, already genotyped with the 50K panel, the 3K genotypes can be imputed to become 50K genotypes with relatively high accuracy.

Using Imputed Genotypes in Genetic Evaluation

Once Holstein Canada is able to launch its service to allow Canadian breeders to genotype their heifers and cows with either the 3K or 50K panel, CDN will subsequently start receiving the resulting 3K genotypes. Using the developed systems, the 3K genotypes will be used to impute a 50K genotype for each animal along with an associated level of imputation accuracy. It is intended that these imputed 50K genotypes will be included in the official genetic evaluations to be released by CDN in December 2010.

Summary

Following extensive use around the world of the standard 50K panel available since the end of 2007, Illumina will soon be releasing a new 3K panel and the Bovine HD panel with 850,000 markers is expected later this year. This portfolio of genotyping panels allows for a strategic use of each panel to maximize the possible returns while minimizing genotyping costs. Of major benefit has been the development of methods for imputing 50K genotypes from 3K genotypes, which basically means filling in the genotypes for the other $\approx 47,000$ markers not evaluated in the 3K panel. The magic of imputation is expected to provide nearly the same increases in accuracy of genetic evaluation with the 3K panel compared to what would have been achieved if all animals had been genotyped with the 50K panel. Imputed genotypes from the 3K panel are scheduled to be included in official genetic evaluations starting December 2010.

Author: Brian Van Doormaal
Date: June 2010