# Matthew McClure, PhD Analyzes Genotypic Data to Determine Cattle's Breeding Potential

For lead geneticist Dr. Matthew McClure, of the Irish Cattle Breeding Federation (ICBF, *www.icbf.com*), finding a software program that could handle massive amounts of data as well as more than the standard twenty-three pairs of chromosomes found in humans was a challenge. McClure says, "Agriculture often feels like the black sheep of the genomics industry. The majority of current genetic analysis programs are written for humans and do not know what to do with those extra chromosomes found in animals and plants."

McClure grew up around the farming and agricultural industries. Before he attended University of Missouri – Columbia for a bachelor's degree in Biochemistry, he and his family manned a 400-acre Missouri farm. Throughout his academic career, McClure continued to surround himself with the agricultural industry. He focused on beef cattle while completing a PhD in Genetics, also at the University of Missouri – Columbia. Prior to his joining the Irish Cattle Breeding Federation in early 2013, McClure completed his postdoctoral work with the U.S. Department of Agriculture's principal in-house research agency, the USDA Agricultural Research Service (ARS) in its Bovine Functional Genomics Laboratory, which is where he first heard about Golden Helix.

"We were looking for easier ways to import, manage, share, and recode data as many programs make it difficult to work with agrigenomic data." McClure was impressed with Golden Helix's software, SNP & Variation Suite (SVS), and its ability to store hundreds of thousands of samples, load massive data quickly and easily, automatically record all workflow steps in a log, and create a myriad of visualizations applicable to his research.

When he transitioned to the ICBF, he said that purchasing SVS was a must.

In the last five years, the Irish Cattle Breeding Federation has worked toward its mission of achieving the greatest possible genetic improvement in the national herd for the wider benefit of all cattle farmers, the dairy and beef industries, and its member organizations. Dr. McClure estimates that by the end of 2014, the ICBF will have every bull in the Irish dairy and beef industry genotyped, as well as fifteen percent of the beef cattle population.

With that amount of data, efficiency is incredibly important. By using SVS, the time it would take to analyze data has been reduced to a fraction of what it would take without the program. "SVS has improved our workflow substantially. We are able to view sequence data, zoom in on variants, and examine alignments to see if they agree. Then we're able to use SNP data to preform genetic analysis without

## QUICK FACTS

**Name:** Dr. Matt McClure

**Title:** Lead Geneticist

**Institution:** Irish Cattle Breeding Federation (ICBF)

**Location:** Bandon, County Cork, Ireland

**Studying:** Cattle breeding traits

**Data:** Illumina IDB custom chip, LD, and 50K, Affymetrix BOS1

**Number of Samples:** 42,000+

**Number of SNPs:** 17,000+ on the IDB

**Analysis:** SNP and DNA-Seq



*Work at the ICBF helps farmers determine breeding potential before using a bull or sire.*

ever leaving the program," stated McClure. "Additionally, having a log that tracks all of the steps of an analysis makes it easy to show exactly what we did and repeat the same process on different datasets."

Another efficiency improvement was found as SVS eliminates the need for researchers to convert and code data into a binary format as other programs do. SVS isn't just a product for human DNA that has been jerry-rigged for animal data, but a fully functional program tailored to the needs of agrigenomic researchers.

And when McClure needs something not currently in the program? In his own words: "Golden Helix's excellent customer service is one of the positive points I always mention when other agrigenomic researchers ask me about SVS. Golden Helix is great about adding new scripts as we needed them for agrigenomic-specific data, such as when we began working with microsatellite data and imputation programs."

*Dr. Matt McClure*

In recent years, SNP data collected on cattle is used for parentage verification, genomic breeding value determination, and disease mapping. However, genetic researchers used to rely on microsatellites for parentage verification. If a calf is SNP genotyped and there is no SNP data for a parent, microsatellite information traditionally had to be collected for the calf to verify the pedigree. To address this issue, McClure developed a method to impute microsatellite alleles using SNP haplotype data, which has been implemented at ICBF since 2013. This means farmers don't have to collect both SNP and microsatellite data on new calves, saving them money.

Much of McClure's and other ICBF research work will directly benefit the Irish agricultural industry as well as the industry as a whole. Before genomics, to get a bull's breeding value, one had to breed him to a large set of cows and then collect the data off their offspring. For young dairy bulls, this represented a large risk for

Artificial Insemination companies—over the course of the 4-5 years it takes to collect the offspring phenotypes, they would invest about $40,000-$50,000 on an animal whose breeding value was largely based on its parents. With SNP genotypes, a young bull can have its genomic breeding value determined soon after birth and at a much lower cost. ICBF currently generates genomic breeding values for Irish dairy cattle and is busily collecting data so they will soon be able to provide this service for Irish beef cattle.

The Irish Cattle Breeding Federation has been keeping McClure quite busy since he joined, and the work won't slow down any time soon. McClure states: "We are really seeing how genetics can improve agriculture. SNP data allows us to determine an animal's genetic disease status along with its breeding value. With this information farmers are able to use the best sire for their farm and avoid breeding two carriers, thus maximizing genetic gain while minimizing genetic disease risk. It's an exciting time, and I feel like the work myself and the ICBF are doing is really making a difference."

*About Golden Helix*

*Founded in 1998, Golden Helix is known for helping genetic research groups working with large-scale DNA-sequencing or microarray data overcome the frustration and challenges of bioinformatic roadblocks: delayed projects, lack of quality findings, and low productivity. By empowering researchers with highly effective software tools, world-class support, and an array of complementary analytic services, we refute the notion that analysis has to be difficult or time consuming. Golden Helix's flagship software product, SNP & Variation Suite (SVS), is an integrated collection of powerful data management, quality assurance, visualization, and tertiary analysis tools for genetic data. SVS is delivered in a user-friendly, scalable platform and is supported by a team of highly trained bioinformaticians, statistical geneticists, and computer scientists that together make advanced statistical and bioinformatic methods accessible to scientists of all levels.*