

**Department of Agriculture, Trade and Consumer Protection
Division of Agricultural Development
Agricultural Development & Diversification Program (ADD)
Grant Project Final Report**

Contract Number: 22003

Grant Project Title: **Application of Genomics to Improve Dairy Cow Fertility**

Amount of Funding Awarded: **\$43,000**

Name of Principal Contact Person: **Dr. Abdullah Kaya**

Organization: **Alta Genetics, Inc.**

Email Address: **akaya@altagenetics.com**

WEB Address: **www.altagenetics.com**

Report Submitted on: **December 31, 2008**

Please use the following questions as a guide for writing your grant project final report. In your final report, please answer each question as it relates to your grant project.

1) What was the original intent of the grant?

- What did you want to accomplish with the grant?
- How was it expected to benefit Wisconsin Agriculture?
- What makes this project work important or significant?

Our overall goal of this study was to improve dairy cattle fertility using genomics approaches for DNA marker assisted sire selection. To achieve our objective, we used Single Nucleotide Polymorphisms (SNP) DNA microarrays using DNA samples from bulls with various levels of fertility. Our project has been a synergistic integration of Alta Genetics' innovative progeny testing program (Alta Advantage[®]) with MSU's proven cutting-edge research in genomics. Now that we have identified molecular markers associated with fertility, once validated and further tested, these will make significant improvement in cattle genetics and fertility-positively impacting, the AI industry, dairy producers, and consumers of dairy products. Research performed in this project is quite important for Wisconsin to maintain and further develop a leading edge in agricultural knowledge and technologies and secondly when lower fertility animals are selected and removed from progeny testing programs through marker assisted selection, the Wisconsin cattle industry could save millions of dollar in the cost of low fertility, which is a major concern and challenge for Wisconsin producers.

2) What steps did you take to reach your goal?

- What worked?
- What challenges did you face?
- What would you do differently?

We accomplished our overall goal through the following specific objectives:

Objective 1: To identify bulls with high and low fertility: The most important element for this study was to select correct materials for the molecular marker selection procedure. In order to associate fertility with molecular markers, the most important aspect is the accuracy of the fertility evaluation. In this study, we used the Alta Advantage[®] database, which is the most

accurate and first of its kind database in the industry to accurately analyze both male (bull) and female fertility. We have used most recent fertility data from collaborating 200 herds across United States. The updated data was edited by the staff to make sure any errors which could mislead fertility ranking of the sires. Then, statistical analysis was carried out to predict individual sire fertility by using threshold models to eliminate the environmental and herd management factors that influence fertility performance of sires. The generated data was consisting of 688,095 breeding records from 876 Holstein bulls with an average of 788 breedings. Because of Alta Genetics' expertise in this area, we did not face any significant challenges. However, although we originally proposed to use 200 bulls, we had to decrease the sample size due to the unexpected high cost of newly developed Bovine SNP microarray containing 54 000 SNPs. Therefore, we removed 60 bulls from original selection. Finally, a total of 140 bulls (70 low and 70 high fertility bulls) were selected for this study from the established data sources available at Alta Genetics.

Objective 2: Genotype 140 bulls with varying fertility with a dense SNP marker panel: Currently molecular genetic markers based on single nucleotide polymorphism (SNP) information are the genetic marker system of choice in cattle for research and application by industry. There was a timely advance in technology relevant to this project: The Illumina 50K Bovine SNP microarray containing 54 000 SNPs has just become available. This array has been developed by a consortium of researchers including USDA and University of Missouri, and it is the densest SNP microarray available for the cattle.

We first isolated genomic DNA from 140 animals in the research laboratory at Mississippi State University. The genomic DNA samples were of high quality. Total of two microgram genomic DNA per bull has been submitted to GeneSeek Company. DNA microarray (Illumina 50K Bovine SNP DNA microarrays) experiments were performed by the scientists at the GeneSeek Company according to their established procedures.

Objective 3: Association of SNP variation with fertility traits: The final step was to detect statistically significant associations between fertility traits and individual SNP variants within the 54 000 SNPs. Our collaborator, Professor Grier Page of University of Alabama-Birmingham (now at the RTI International in Atlanta, GA) has used modern statistical genetic tools to investigate differences in the frequency of specific SNP genotypes between the animals with the highest and the lowest fertility traits in the Holstein bulls. The challenge was that the statistical analyses of the microarray data has taken more time that we originally anticipated. Thus, in the future we will devote more time for data analyses.

3) What were you able to accomplish?

- What are the results from this project?
- Include any analysis of data collected or materials developed through project work.

Objective 1: To identify bulls with high and low fertility: After ranking sires by their fertility, we were able to generate low and high fertility groups by putting a distance of 1 SD (standard deviation) from the tail of each group. By doing this, we wanted to make sure that the selected samples are optimum fit for the genomic association analysis. The bulls scoring highest and lowest fertility deviation from average with higher reliability, average 980 breeding/bull, were selected for this study. The differences in the average fertility indexes between high and low

fertility groups were 7.8% which was obtained from bulls having adequate records for higher reliability. While 70 bulls which were scored 3.4 % above the average were considered high fertile, 70 bulls which were scored – 4.4 % below the average were defined as low fertility.

Objective 2: Genotype 140 bulls with varying fertility with a dense SNP marker panel:

The qualities of the extracted DNA samples were good and the amount was enough for the experiments. The results of the SNP genotyping experiments showed that the SNP DNA microarrays worked well, i.e., there was no technical errors associated with the microarrays. There were no problems in microarray experiments, i.e., labeling, hybridization, washing, scanning of the microarrays and analyses of the data. The results of the microarray experiments were delivered by GeneSeek to us in a DVD.

Objective 3: Association of SNP variation with fertility traits: The 140 bulls of varied fertility were typed for 54,001 SNPs using DNA microarray (Illumina 50K Bovine SNP DNA microarray). While 44,485 SNP DNA markers remain in the data set, a total of 9,516 SNP markers were excluded from further analysis due to the potential issues such as monomorphic for A allele, monomorphic for A allele and low call rate < 95%. Additionally, 4 sires were excluded due to low call (<0.95) rate. Markers were tested for Hardy-Weinberg Equilibrium using a 1 degree of freedom chi-square test. Markers with a p value < 0.05 were flagged as potentially out of HWE. A total of 1,345 markers failed HWE. Markers with a p < 0.05 were flagged as being potentially out of HWE, but were used for association analysis.

Taken together, 2,154 out of 44,485 SNP markers resulted in a significant association between fertility at $p < 0.05$. When we narrow down the data, 421 SNPs were significant level at $p < 0.01$. The results also showed that 34 SNPs were highly significant at $p < 0.001$. There were only 2 SNPs were very strongly associated with fertility at $p < 0.0001$.

4) What conclusions can you make based on project work the analysis of collected data?

- There was a significant variation of fertility among the dairy bulls, despite the normal motility, morphology and number of the spermatozoa from these bulls.
- Genomic DNA samples isolated from the spermatozoa were of high quality and suitable for SNP DNA microarray experiments.
- The newly developed Bovine SNP microarray (Bovine 50K SNP microarray from Illumina, the highest density SNP array available for bovine) worked well for detection of SNPs in the bovine genome.
- The research efforts were able to identify SNPs significantly associated with the fertility traits.

5) What do you plan to do in the future as a result of this project?

This project produced new knowledge, specific genomic markers associated with fertility which is an economically important trait. The knowledge produced is an essential part, however, more research needs to be done for definitively identifying the molecular markers and mechanism(s) by which these markers play important roles in fertility. Specifically, we plan to pursue further research through the following specific objectives:

- 1. Validate marker associations in an independent sample of animals.** We will validate the identified markers using DNA samples of additional bulls with various fertility phenotypes. Our *working hypothesis* for this objective is that validation of SNPs associated with bull fertility is necessary to definitively validate significant association of the mutation with fertility phenotype.
- 2. Examine candidate genes in validated genomic regions and identify and test potential functional polymorphisms.** Our *working hypothesis* for this objective is that location of SNPs in the genome provides significant information about the importance of the region for fertility. We will examine the validated regions to identify any plausible candidate genes in the region. We will then resequence the most significant validated regions based upon the size of the candidate gene or genes. We will test our hypothesis by analyzing large lengths of DNA sequences of the most significant validated regions associated with fertility.

6) What information or additional resources are needed to commercially develop this enterprise?

In order to accomplish the above specific objectives under question#5, and to commercially develop this enterprise, more research resources and funding is needed. Additional efforts are necessary to validate the markers (genotyping an independent population of animals for the top most significant markers, i.e., the 34 markers associated with fertility at $p < 0.001$), data analyses, and further studies on QTL and bioinformatics.

7) How should the agricultural industry use the results from your grant project?

As an *outcome* of the proposed investigations, we expect that using the identified markers linked with fertility QTL in selection programs, private industry will rapidly adapt these genome-enabled technologies to increase cattle production efficiency which will provide economic savings. The producers can do marker assisted selection (MAS) by using the semen from the high fertility bulls (high fertility bulls that are carrying the SNPs strongly associated with the identified SNPs) for their breeding programs. The use of high fertility bulls will increase the production efficiency of and product quality, ultimately enhancing the economic viability of the dairy industry in the state and the country. These will also benefit the consumers as well. Using the markers identified in this project, the producers will be able genotype their bulls to predict the fertility of the bulls and this will save the costs associated with keeping the low fertility bulls.