Project Plan
NP 101 – Food Animal Production
March–June 2017

Old ARS Research Project Number
8042-31000-101-00D

Research Management Unit
Animal Genomics and Improvement Laboratory (AGIL)

Location
Beltsville, Maryland

Project Title
Improving Dairy Animals by Increasing Accuracy of Genomic Prediction, Evaluating New Traits, and Redefining Selection Goals

Investigators
Paul M. VanRaden, Lead Scientist 1.00
John B. Cole.....1.00
Curtis P. Van Tassell.....0.25
Research Geneticist (vacant) 1.00
Research Geneticist (vacant) 1.00

Scientific Staff Years
4.25

Planned Duration
60 months

Pre-Peer Review Signature Page

Paul M. VanRaden
8042-31000-101-00D
Improving Dairy Animals by Increasing Accuracy of Genomic Prediction, Evaluating New Traits, and Redefining Selection Goals

This project plan was revised, as appropriate, according to the peer review recommendations and/or other insights developed while considering the peer review recommendations. A response to each peer review recommendation is attached. If recommendations were not adopted, a rationale is provided.

_____ _____
Research Leader Date

This final version of the project plan reflects the best efforts of the research team to consider the recommendations provided by peer reviewers. The responses to the peer review recommendations are satisfactory.

_____ _____
Center, Institute or Lab Director Date

The attached plan for the project identified above was created by a team of credible researchers and externally reviewed and recognized by the team's management and National Program Leader to establish the project's relevance and dedication to the Agricultural Research Service's mission and Congressional mandates. It reflects the best efforts of the research team to consider the recommendations provided by peer reviewers. The responses to the peer review recommendations are satisfactory. The project plan has completed a scientific merit peer review in accordance with the Research Title of the 1998 Farm Bill (PL105-185) and was deemed feasible for implementation. Reasonable consideration was given to each recommendation for improvement provided by the peer reviewers.

_____ _____
Area Director (original signature required) Date

PrePlan Signature Page for ONP Validation(s)

Pre-Peer Review

Paul M. VanRaden
8042-31000-101-00D, NP 101
Improving Dairy Animals by Increasing Accuracy of Genomic Prediction , Evaluating New Traits, and Redefining Selection Goals

❑ Signature Page Completed for Research Leader through Area Director

X The objectives in this PrePlan are those provided in the PDRAM or subsequently approved by the Office of National Programs and the approaches are suitable for achieving the objectives.

2/3/2017

National Program Leader Date

Comments:

Table of Contents

Project Summary

Genetic progress of dairy cattle for traits of economic importance can be advanced more rapidly through improved genomic prediction methods, thereby leading to immediate benefits to producers and consumers worldwide. This project will use whole-genome or targeted DNA sequence data to discover **naturally occurring variants** that cause trait differences between animals or genetic **markers** closely associated with those differences to improve genotyping arrays and the potential for gene editing. The rapid growth in number and size of international databases and larger variant sets require deriving more accurate imputation methods, advanced statistical models, and efficient computer programs for processing the big data associated with dairy cattle records. New models will allow monitoring and removing potential biases caused by genomic pre-selection. Additional traits will be evaluated if the estimated economic value and heritability are sufficiently high to justify selection and use. Updated genetic-economic indexes for combining all traits will guide breeders on selection goals, and producer profits from alternative breeding programs and potential investments in data will be compared. Development of more cost-effective genotyping tools will be optimized through collaboration with other scientists in ARS, universities, and industry. Genomic predictions for crossbreds will be developed from phenotypic data using an all-breed scale instead of separate within-breed scales. Phenotypic effects of management practices and interactions of genotype with environment will also be documented and predicted. Collecting and combining information from phenotypes, genotypes, and pedigrees into more accurate predictions will allow breeders to greatly improve the production efficiency of future dairy cattle. Other species will also be improved by using the genomic selection methods and programs developed in this research.

Objectives

The primary objective is to improve dairy cattle productive efficiency, health, fertility, and other traits of economic importance by evaluating genetic merit and comparing management practices so that the United States and other countries can meet the dietary needs of their populations. Specific objectives include:

Objective 1. Expand genomic data used in prediction by selecting new variants that more precisely track the true gene mutations that cause phenotypic differences.

Objective 2. Evaluate new traits that can all be predicted at birth from the same inexpensive DNA sample.

Objective 3. Improve efficiency of genomic prediction and computation by developing faster algorithms, testing new adjustments and models, and accounting for genomic pre-selection in evaluation.

For Objective 1, research will focus on a) obtaining whole-genome sequence data for additional bulls and dairy breeds, b) using the sequence data to identify and select causal **variants** to replace the linked markers currently used, c) assisting with design of new genotyping arrays used by dairy cattle breeders, d) choosing optimal variant lists for routine prediction to improve imputation accuracy and prediction accuracy across breeds, e) cooperating in international sequencing research, and f) investigating potential uses for gene editing. For Objective 2, research will focus on a) traits such as age at first calving, gestation length, and lactation persistency that can be computed from data already available at no cost nationally, b) traits such as feed efficiency, lameness, milk infra-red spectral data, milking speed, clinical mastitis, or other health conditions that require additional investment to obtain a national data set, c) economic analysis to estimate the value of each trait in an index as well as the investment justified to collect additional data, and d) genetic-by-environmental interactions, non-additive effects, mating programs, or other data reports that could improve the management of genotyped dairy cattle with little added cost. For Objective 3, research will focus on a) deriving, programming, and providing new software to process the rapidly growing data with reasonable processing times and computing costs, b) developing and comparing statistical models for predictive ability, c) monitoring accuracy of evaluations for all traits, and d) correcting biases that may occur as breeders change selection programs and use new reproductive or management technologies.

Need for Research

Genomic predictions computed from the national database shared by the Council on Dairy Cattle Breeding (CDCB) and the Animal Genomics and Improvement Laboratory (AGIL) are used by breeders all over the world to improve their dairy cattle. Research by AGIL to increase the prediction accuracy and the traits included thus can lead to immediate global benefits. Benefits to U.S. producers and consumers are many times larger than the $75 million yearly cost of data collection (5 million cows times $15 per year) for cows enrolled in Dairy Herd Improvement. Extremely large numbers of both phenotypes and genotypes are needed for accurate genomic selection, and the U.S. database is the world's largest for cattle with over 1.5 million genotyped and over 30 million phenotyped animals. Genomic selection has already

nearly doubled the rate of genetic progress in the dairy industry, and further acceleration is very likely given the advances in genomic tools and decreasing costs of reading and analyzing DNA. Tools include lower-density arrays costing about $40 for 10,000 markers to higher-density arrays costing about $120 for 150,000 markers. Most animals are genotyped with low-cost arrays and their missing genotypes are imputed (estimated). International genotype exchanges have also increased the data available, including exchange of all animal genotypes with Canada and bull genotypes with Italy, the United Kingdom, Denmark, Germany, Switzerland, Japan, and others. In 2015, AGIL joined the global exchange of whole genome sequence genotypes (1000 Bull Genomes Project), which provided over 30 million additional variants to select from to build future arrays.

Statistical methods to combine genotypic, phenotypic, and pedigree information need constant revision to account for rapid data growth, new data types, and changing breeding programs. When phenotypes are added primarily from animals that were pre-selected based on genomic merit, traditional genetic evaluations that use only phenotypes and pedigrees may become biased. Current genomic evaluations are a post-processing step that uses traditional evaluations as input data. Those programs should be revised to account for all three data types (pedigree, phenotype, and genotype) simultaneously, but available algorithms are either too computationally costly or less accurate with very large data sets. Research is needed to 1) further investigate single-step methods that account for genomic pre-selection, 2) develop all-breed instead of single-breed genomic equations to improve marker effect estimates and improve evaluations on crossbred animals, 3) improve genotype imputation methods, and 4) discover the location of causative variants as well as improve detection methods for lethal recessive alleles.

Data for several traits affecting profit have not been available historically because of cost but are now collected by on-farm management software. The problem is to define traits uniformly and provide incentives for transferring data to the national database. Industry partners need much better estimates of the value of potential traits to the accuracy of the overall system. Income and cost factors continue to change, making economic analysis and selection goals ongoing needs. To support the ultimate impact of research, an economic analysis is needed to optimize genetic progress and maximize financial benefits from collected data. Analyses conducted should characterize effects of herd management practices on profitability, determine optimal systems for genetic improvement, quantify economic values for potential new traits such as feed efficiency, monitor and update economic values of individual traits, and design methods to select healthy, fertile animals with high lifetime production of affordable milk. Expanded research is also needed to include additional economically valuable traits for health, lactation persistency, and adaptation to differing climates and production systems.

Genomic predictions have already nearly doubled the rate of genetic progress in dairy cattle (Norman et al., 2014; García-Ruiz et al., 2016) and will be applied to many more animals as part of this project. The trend toward evaluating more traits is expected to accelerate because each genotype can provide predictions for many phenotypes. Initial efforts to standardize trait definitions and organize data collection will provide the data needed to develop future evaluations, but may not result in routine evaluations for some traits during this project. Maintenance of evaluations for previous traits requires much research to accommodate database growth and avoid new biases that may occur with changes in incoming data. The CDCB recently became responsible for routine processing of dairy data but relies on AGIL staff for major upgrades to computer algorithms and scientific methods. Improved programs for processing DNA data can benefit dairy cattle globally and also many other species because the methods transfer very easily.

Improved genotyping arrays will result from research to identify new variants from sequence data that more closely track the genetic **variants** that affect economic traits of animals. Genetic evaluations for new traits will result in more balanced progress from selection and improved

animal health and efficiency. Faster algorithms and efficient computation will allow more animals and data per animal to be included while improving the accuracy and reducing any biases in genetic rankings. The computer programs will be made available for general use in other countries or for other species. Economic research will give dairy producers direct guidance on breeding goals that provide optimum long-term genetic progress.

Owners of dairy animals have improved their herds for over a century by sending records of individual animals to USDA and using the genetic rankings provided by USDA to select the best bulls and cows. The U.S. national database now collects records from over half of all U.S. cows each year, and percentages are steadily growing. Breeders in other countries began using USDA evaluations several decades ago to import U.S. breeding stock, embryos, and semen (over $150 million annually) to improve their foreign herds. The Interbull Centre (Uppsala, Sweden) began formal exchange of genetic evaluations in 1995 and now exchanges national evaluations from more than 30 countries. Thus, U.S. national evaluations are used indirectly for selection by customers across the globe. Genomic evaluations have transformed international breeders into direct customers of U.S. evaluations. Genotypes from DNA of both foreign and domestic animals allow dairy cattle breeders everywhere to receive predictions and rankings that are directly and easily comparable across the world. The U.S. database now includes genotypes sent from more than 50 countries, including 120,000 from Europe, 5,000 from Asia, 400 from Africa, 9,000 from Oceania, and 13,000 from Latin America, in addition to 1.2 million genotyped dairy cattle from North America. The USDA/CDCB database was the first in the world of any species to reach 1 million genotypes, slightly before the human genomic companies 23andMe or Ancestry.com reached 1 million genotypes, primarily because customers can directly use the genomic predictions computed by AGIL programs.

The research addresses the following research components in the 2018–22 Food Animal Production National Program (NP 101) Action Plan: Component 1: Increasing Production and Production Efficiencies while Enhancing Animal Well-Being across Diverse Food Animal Production Systems, Problem Statement 1A, Improving the Efficiency of Growth and Nutrient Utilization; Problem Statement 1B, Improving Reproductive Efficiency; and Problem Statement 1C, Enhancing Animal Well-Being and Reducing Stress; and Component 2: Understanding, Improving, and Effectively using Animal Genetic and Genomic Resources, Problem Statement 2A, Develop Bioinformatic and other Required Capacities for Research in Genomics and Metagenomics; Problem statement 2B, Characterize Functional Genomic Pathways and Their Interactions; and Problem Statement 2D, Develop and Implement Genetic Improvement Programs Using Genomic Tools.

Scientific Background

For over 100 years, USDA has collaborated with the U.S. dairy industry to collect data on economically important traits of dairy cattle and used those data for genetic improvement (VanRaden and Miller, 2008). The national database of phenotypic and pedigree information that began in 1908 was converted to computer processing around 1960. Bull evaluations for milk and fat yields have been calculated and provided to breeders since 1926. Since then, data have been collected and genetic evaluations developed and released to the industry for additional traits: protein (1977), conformation (1978), somatic cell score and productive life (1994), calving ease (2002), daughter pregnancy rate (2003), stillbirth (2006), and cow and heifer conception rates (2010), and cow livability (2016). Several additional traits also affect cow profit (e.g., feed efficiency, health disorders, and gestation length), and selection indexes will be revised as data for those traits become available.

Methods to test DNA were developed and applied to dairy cattle in 2007 by researchers from USDA's Animal Genomics and Improvement Laboratory (AGIL) and an international consortium of university and industry cooperators (Matukumalli et al., 2009). The genomic predictions developed by AGIL quickly became official in 2009 (Wiggans et al., 2011) and superseded the traditional evaluations that used only phenotype and pedigree data. The U.S.-designed chip and the national genomic evaluation became international because standard genotypes are much easier to transfer, combine, and use than previously possible with national phenotype files. This led to formal genotype exchanges with several major countries and rapid growth of customers from 50 countries all over the world, which has greatly increased worldwide demand for genomic predictions.

The flow of both traditional and genomic data from the dairy industry to AGIL is currently sustained through a Nonfunded Cooperative Agreement (NFCA, see Appendix 1) with CDCB and supported by Dairy Herd Information (DHI) groups and their data processing centers, artificial-insemination (AI) organizations, and breed associations. After fully implementing the NFCA in 2015, CDCB took responsibility for collecting the genotype, phenotype, and pedigree data and computing genetic evaluations for production, reproduction, type, calving, and fitness traits, while AGIL researchers continued to derive the methods and program much of the software. Industry control over collection of and access to the data allows additional traits to be collected because CDCB collects fees to provide incentives for data contributions and because the data are no longer subject to freedom-of-information queries, a concern that slowed or halted collection of data for some traits in the past. The NFCA allows expansion of research on genomic evaluations that was not previously possible when AGIL was responsible for providing free service functions for the national genetic evaluation system.

Most benefits from recording phenotypes on cows or daughters of bulls previously went directly to the animal's owner, but genomic selection has changed incentives for data collection away from individual breeders and towards breeders in general. Thus, genomic selection makes experimental design and economic analysis at the population level much more important because additional data on reference animals benefits all breeders nearly equally. Investments to obtain more traits or reference genotypes require cooperation across companies or formal international agreements instead of within-company progeny-test programs. Scientific analyses can directly guide those industry business decisions. Examples are research to estimate reliability gains from international genotype trades (Olson et al., 2011), from higher density genotyping arrays (VanRaden et al., 2011a), and from additional phenotypes for males or females (Cooper et al., 2015). Early emphasis was on adding genotypes, but once those are available, new traits can be added without extra genotyping cost (Gonzalez-Recio et al., 2014a) based on economic and exploratory research.

Expanded genomic data

Genomic selection began in 2008 with one genotyping array (chip) containing 50,000 genetic markers, but has quickly evolved to include 23 different chips containing either fewer markers to reduce costs or more markers to improve accuracy. The first chips purposely excluded known causal variants and quantitative trait loci (QTLs) because only a few were known, and most of those were patented. The U.S. Supreme Court (2013) ruled that patents on DNA testing are no longer valid, thereby allowing both markers and QTLs to be included on chips without royalty, at least in U.S. genomic predictions. Many other countries still honor gene test patents and still use only the 50,000 original markers, whereas U.S. predictions include both markers and QTLs (Wiggans et al., 2013, 2016a). Recent rapid growth in QTL discovery and numbers of chips now makes the process of deciding which markers and QTLs to use in predictions much more challenging.

Whole-genome sequencing has become more affordable every year but still costs about $1,000 per animal for 10X coverage compared with about $40 to $200 per animal for chip

genotypes that range from 4,000 to 777,000 markers. The chips designed recently nearly always include some QTLs, but only GeneSeek (Neogen Genomics, Lincoln, NE) has been sending the QTL genotypes to include in U.S. predictions. The QTL genotypes can also be obtained from sequence data, but because of the expense, cattle researchers from much of the world have joined data sets in the 1000 Bull Genomes Project (Daetwyler et al., 2014; Hayes et al., 2014) or made data freely available (Stouthard et al., 2015) to reduce cost and increase power of detection. Researchers at AGIL joined this exchange to obtain global sequence data for more than 1,500 bulls in 2015. Many of those were beef bulls, and several dairy breeds still have too few animals sequenced (or even genotyped) to obtain accurate predictions. Initial results (O'Connell et al., 2016) with Holstein data look promising, and additional research on the current sequence data and on future data updates should lead to further discoveries of QTLs that are useful in prediction.

The sequence data include 39 million discovered variants, many times more than available on any of the chips. However, about 10% of variants from the chips are not present in the sequence data, and the variants include only single-nucleotide polymorphisms (SNPs) and small insertions and deletions (indels) because more complex **variants** are more difficult to call reliably (Bickhart and Liu, 2014). Very high linkage among several neighboring variants makes isolating QTLs from markers very difficult and most gains from further variant discovery small (van Binsbergen et al., 2014b; Brøndum et al., 2015). Exceptions are if new mutations occur within common haplotypes, which makes the new QTLs difficult to detect or impute. To use sequence data in routine genomic evaluation, the variant list must be reduced to a manageable size, the higher error rate (compared to chip genotypes) must be accounted for, and additional variants and variant classes not in the current data must be discovered.

Many new lethal mutations have been discovered first from haplotypes (VanRaden et al., 2011b; Fritz et al., 2013; Pausch et al., 2015; Kipp et al., 2016) and then by sequencing carrier or homozygous animals. An alternative is to discover the mutations directly from whole-genome sequence data at the population level (Charlier et al., 2016). Similar projects are a major focus of human sequencing projects (Xue et al., 2012; Gao et al., 2015), whereas genomic prediction is just beginning for human traits (Dhurandhar et al., 2015). For U.S. cattle, the genetic defects are then imputed to predict carrier status for all genotyped animals of the same breed. As new defects are identified, they should be reported to breeders, and the accuracy of detection or imputation must be monitored (van Binsbergen et al., 2014a).

New traits

Traits with low heritability, such as health conditions, or expensive to collect, such as feed efficiency, were not evaluated until recently because even progeny-tested bulls had low reliability. Genomic evaluations now provide much higher reliability for many traits that were previously difficult to predict. As a result, phenotypes become the limiting factor if most animals are genotyped (Gonzalez-Recio et al., 2014a; Wientjes et al., 2016). Most traits are selected because of direct economic benefit to the animal's owner, but society may expect some trait improvement for the animal's benefit as well (Ellen et al., 2014). Some traits of interest, such as feed intake and immune function, are expensive or difficult to measure directly, so the use of low-cost, correlated traits often is necessary to produce high-reliability genetic evaluations (Egger-Danner et al., 2014). To keep selection indexes relevant, economic values of new traits must be estimated, and the benefits from all potential traits must be balanced.

Feed efficiency has received much attention and investment in recent years (Vallimont et al., 2012; Lu et al., 2015; VandeHaar et al., 2016), and genomic predictions are computed routinely in The Netherlands (Veerkamp et al., 2014) and Australia (Pryce et al., 2015). Most research uses feed intake from cows, and a few also from heifers (Davis et al., 2014; Gonzalez-Recio et al., 2014b). International data sets have been combined to obtain more data for research (Berry et al., 2014a; de Haas et al., 2014) but have not been implemented routinely yet. Scientists at

AGIL have contributed feed intake data from the Beltsville research herd (Connor et al., 2013) to these projects and have provided genotype editing for all U.S. project cows. Strategies for ongoing data collection are now being developed so that routine genetic predictions can be implemented on a national basis.

Gestation length can be useful in maternity pen management, mating programs to group all birth dates together in seasonal calving, or as a correlated trait to improve calving ease (Dhakal et al., 2013). Norman et al. (2011), Eaglen et al. (2013), and Jenkins et al. (2015) estimated correlations with other traits, and Maltecca et al. (2011) computed genomic predictions for gestation length. Norman et al. (2009) computed genetic evaluations from U.S. data for research, and routine genomic evaluations for gestation length are currently being developed by AGIL scientists. Age at first calving is another fertility-related trait that may have sufficient economic value to justify routine selection (Cole and Null, 2010).

Producer-recorded health events are routinely stored in on-farm computer systems and represent a valuable source of data for use in genetic improvement of dairy cow health (Parker Gaddis et al., 2012; Jamrozik et al., 2016). Genetic evaluations have been calculated using those data (Zwald et al., 2004; Parker Gaddis et al., 2014) on a research basis, and genomic evaluations for some health traits are now commercially available in the United States (Vukasinovic et al., 2017). Threshold animal models have been used to calculate breeding values using small data sets, but computational limitations may require the use of sire-maternal grandsire threshold models or linear animal models for national data sets. The use of multivariate models that combine limited health data with commonly recorded data, such as longevity, may result in higher accuracies. In addition to their use in genetic improvement programs, health data also may be used to develop benchmarks in support of herd management decisions (Parker Gaddis et al., 2016).

Scientists at AGIL currently are collaborating with CDCB research staff to develop routine genomic evaluations for six health traits (clinical mastitis, displaced abomasum, ketosis, metritis, milk fever/hypocalcemia, and retained placenta). Additional traits regarding lameness and hoof health (Weber et al., 2013; Koeck et al., 2014; Dhakal et al., 2015) have been evaluated in other countries, but a U.S. database must be developed before evaluation research can begin at AGIL. Calf mortality can be evaluated but has low heritability (Hansen et al., 2003; Fuerst-Waltl and Sorensen, 2010), and preliminary evaluations have been computed from U.S. data (Henderson et al., 2011; VanRaden et al., 2016). The best models to use for routine genetic evaluation are still to be determined as well as how health traits should be included in economic indexes.

Selection indexes must be periodically revised to ensure that economic assumptions are consistent with current industry conditions as well as to incorporate new traits (VanRaden, 2004; Shook, 2006). The first USDA national index implemented in 1971 included only milk and fat yields and was expanded to include protein yield in 1977. To address the needs of dairy producers that market milk based on fluid milk and cheese pricing, fluid and cheese indexes also were implemented in 1983. The net merit index introduced in 1994 included productive life to measure longevity and somatic cell score to indicate mastitis resistance. Additional traits added to the USDA indexes include conformation traits in 2000, daughter pregnancy rate and calving ease in 2003, stillbirth rate in 2006, heifer and cow conception rates in 2014, and cow mortality will be added in April 2017. Relative emphasis on each trait in the current net merit index is 22% for fat yield, 20% for protein yield, 19% for productive life, −7% for somatic cell score, 8% for udder conformation, 3% for feet and legs conformation, −5% for body size traits, 10% for fertility traits, and 5% for calving traits (VanRaden and Cole, 2014). Economic values of all traits are updated with each index revision.

Efficiency of prediction and computation

Most genomic prediction models treat all marker effects as having equal prior expected variance. Evaluations in the United States use a method analogous to Bayes A (VanRaden, 2008), with some marker effects assumed to be much larger than from a normal distribution. With markers and QTLs now both included in the variant list, the prior could adjust for variant type to assign more prior variance to QTLs than to markers (Kichaev et al., 2014; Pérez-Enciso et al., 2015; MacLeod et al., 2016). Also, variance of effects can be adjusted based on size of adjacent effects (Yang and Tempelman, 2012) to improve reliability of predictions. Finally, gene expression levels can act as an independent data source to separate causal from linked markers, but data sets must be large and the specific tissues most associated with particular traits may not be clear (Zhu et al., 2016). As variant selection strategies become more complex, Bayesian priors could also be revised to capture more of the known prior information.

Processing of sequence data generates very large data files and requires lengthy computation per animal. Existing programs were nearly all developed for use in human genetics, where budgets are much larger and computer resources more available than for agricultural research (Womack, 2005). Similarly, many programs to impute missing genotypes were available in human genetics, but animal breeders developed much faster and more accurate programs for use with livestock populations (VanRaden et al., 2013; Sargolzaei et al., 2014). Faster programs and more efficient methods to store sequence data are also needed to make larger scale processing feasible (De Donato et al., 2013; Hickey, 2013; Gorjanc et al., 2015). Community members must agree on new forms of exchange because the data files are shared (Li et al., 2009). New methods of sequence alignment and variant calling have been developed by AGIL and tested on simulated data but must also be applied to large actual data sets. Updates to software and formats help keep processing costs reasonable as more animals are sequenced.

Traditional evaluation models can become biased when the phenotyped animals are a highly selected subset of the genotyped animals. Animal models adjust for pedigree merit of parents, progeny, mates, and herdmates but not for their improved merit due to genomic pre-selection of the best sibs, potentially biasing both genetic trend and individual animal estimates. Such biases were demonstrated in simulations (Patry and Ducrocq, 2011), but little bias has yet been demonstrated in actual data sets (VanRaden, 2016). However, rapid changes in breeding programs such as use of only young bulls or selected mates instead of random mates can make unbiased estimation of genetic trend and genetic merit very difficult. Single-step genomic predictions can eliminate the biases by accounting for the pre-selection in the model (Vitezica et al., 2011; Ma et al., 2015b). Algorithms for single-step prediction did not scale to large numbers of genotypes until recently (Misztal, 2016), and even then the observed reliability declined when hundreds of thousands of young-animal genotypes were included in the equations (Masuda et al., 2016a). Proper accounting for unknown parents also helps prevent bias (Misztal et al., 2013; Tsuruta et al., 2014; Matilainen et al., 2016). The U.S. evaluations now include over 200,000 cows in the reference population along with over 30,000 bulls, but the cow data are not always helpful for traits with low heritability (Cooper et al., 2015). Bias must be monitored; if detected, algorithms must be developed to correct the bias.

Multitrait evaluations are needed when new traits are recorded and previous data from correlated traits are available. Most national genomic evaluations are single-trait, but multitrait models can improve reliability (Calus and Veerkamp, 2011; Jia and Jannink, 2012). However, computational methods must be efficient as data sets grow (VanRaden et al., 2014; Calus et al., 2015). Early generations might not add to prediction reliability because the trait definitions or linkage patterns change across time (Lourenco et al., 2014), but more complex models to deal with this are difficult to manage. Foreign phenotypic data available from multitrait across-country evaluation (MACE) via Interbull should also be incorporated into domestic evaluations (VanRaden et al., 2014; Přibyl et al., 2015; Vandenplas et al., 2015). More complex models can

be tested and compared using the flexible software packages and algorithms that have been developed recently (Misztal et al., 2014; VanRaden et al., 2014).

Crossbred animals have been included in U.S. traditional evaluations since 2007 but excluded from genomic evaluations because marker effects are estimated separately on within-breed scales. Genomic evaluations for crossbreds can be obtained on an all-breed scale by combining purebred marker effects weighted by breed composition (Makgahlela et al., 2013). Imputation of the crossbred genotypes should be accurate if genotypes from both pure breeds are included (Berry et al., 2014b; Bouwman and Veerkamp, 2014; Brøndum et al., 2014; Ventura et al., 2014). Much previous genomic evaluation research in dairy cattle has focused on improving purebred predictions using genotypes from multiple breeds to estimate effects (Karoui et al., 2012; Hozé et al., 2014; Lund et al., 2014; Zhou et al., 2014; Kemper et al., 2015; van den Berg et al., 2016), but most countries have had few crossbreds to evaluate and have not implemented multibreed genomics except for a few such as New Zealand (Harris and Johnson, 2010; Winkelman et al., 2015). Multibreed evaluation has become a higher priority because over 10,000 U.S. crossbred animals have been genotyped but have not received genomic predictions (Olson et al., 2012), and many of the top Jersey animals contain some Holstein contribution (VanRaden and Cooper, 2015).

Related Research

This project is closely coordinated with two other CRIS projects within AGIL: “Enhancing Genetic Merit of Ruminants through Genome Selection and Analysis” (8042-31000-104-00D), and “Improving Feed Efficiency and Environmental Sustainability of Dairy Cattle through Genomics and Novel Technologies” (8042-31320-077-00D). The first AGIL project focuses on developing new tools across ruminant species; the lead scientist (G.E. Liu) has collaborated with the proposed project’s scientists on past projects (Hou et al., 2011, 2012; McClure et al., 2013a, 2014; Xu et al., 2014; Bickhart and Liu, 2014; Bickhart et al., 2015, Ma et al., 2015a). The second AGIL project focuses on understanding the genetics and physiology of feed efficiency; the lead scientist (E.E. Connor) has also collaborated on past projects (Connor et al., 2012, 2013; VanRaden et al., 2013). The proposed project focuses on applying new tools and data directly to ongoing selection programs for dairy cattle.

A NIFA-CRIS search of research on November 7, 2016, revealed 72,998 projects generally associated with genetic evaluation or the dairy industry. After narrowing the search to exclude plants and species not of interest, 208 projects remained, of which 32 appeared to have a possible connection to this project. Examination of those projects revealed 3 projects that were ongoing and of direct relevance:

- **Development of Animal Genetic Resources Information Network (Animal-GRIN) for U.S. and Brazilian Use** (3012-31000-005-16N)
  EMBRAPA (H.D. Blackburn, A. Mariante, and S. Paiva)
  Terminates February 15, 2020
  Note: The focus of this project is on the preservation of viable animal germplasm, not genetic improvement. Drs. H.D. Blackburn of GRIN and J.B. Cole of AGIL have previously discussed strategies for comparing GRIN collections with pedigree populations to identify lineages that are not adequately represented in the GRIN collection, but no active collaboration between the groups exists at this time.

- **Improvement of Genetic Evaluation of Dairy Cattle Through Expanded Genomic Data and Improved Computing Procedures** (8042-31000-101-09T; subordinate project of prior project period)
  ARS (G.R. Wiggans, retired; currently P.M. VanRaden)
  Terminates July 23, 2018
  Note: This Trust Fund Cooperative Agreement with the Council on Dairy Cattle Breeding (see Appendix 2) will automatically become a subordinate project of the proposed project and will have the same lead scientist (P.M. VanRaden).

- **Pooled Genotyping and Sequencing for Development of Molecular Breeding Values for Resistance to Bovine Respiratory Disease** (3040-31320-012-06R)
  ARS (L.A. Kuehn, M. Enns, T.P. Smith, T.G. McDaneld, and J.W. Keele
  Terminates August 31, 2018
  Note: This project focuses on beef cattle production. As of Aug. 31, 2016, sample collection was still ongoing, and genotyping and genetic analysis of those samples is expected to be completed during 2017; development of SNP markers with consistent predictive merit for resistance/susceptibility will begin in 2017. The identified markers could be investigated for their usefulness in predicting resistance/susceptibility to bovine respiratory disease in dairy cattle. Several of this project's scientists (L.A. Kuehn, T.P. Smith, T.G. McDaneld, and J.W. Keele) have collaborated with AGIL scientists on past projects (McClure et al., 2013b; McDaneld et al., 2014).

Approach and Research Procedures

Objective 1: Expand genomic data used in prediction by selecting new variants that more precisely track the true gene mutations that cause phenotypic differences.

Hypothesis 1: Addition of new variants from sequence data will improve reliability of genomic predictions compared to the single nucleotide polymorphism markers currently used from genotyping arrays.

Experimental Design: Analyses of the 1000 Bull Genomes data that began in 2015 with run 5 data will continue with further testing of **variant selection** strategies and run 6 data when that arrives, possibly in 2018. The run 5 data included information for 1,500 bulls of all breeds, and the run 6 data are expected to include information for 2,700 bulls. Our initial tests used only the 440 sequenced Holsteins and imputed the sequence genotypes for 27,000 Holstein bulls in the national reference population. Additional imputed genotypes for reference cows could help isolate causal variants because of more crossovers within families. However, imputation to sequence is more difficult for cows because most are genotyped at lower densities. Two-stage imputation will be used, first to high density and then to sequence, and accuracy of imputation will be compared by simulation.

To supplement the global sequence data, **local sequence data** will be generated for families exhibiting new fertility defects or other health conditions of interest and for bulls that are homozygous for less frequent haplotypes. Initial sequencing will focus on 200 additional Holstein bulls and 50 Jersey bulls at high coverage (about 30X) to allow discovery of copy number variants that are not available from the 1000 Bull Genomes project. Animals will be selected for sequencing using an algorithm that maximizes coverage of rare haplotypes and minimizes re-sequencing of common haplotypes (Bickhart et al., 2016). **These sequences will serve as reference "sub-genomes," where particular haplotypes are targeted across the**

portfolio of bulls, and help to identify variants with a high level of certainty. Because of systematic biases in Illumina library preparation, the higher coverage sequence data provide some measure of protection from those biases. The DNA library preparation and sequencing both may be contracted out or done within AGIL depending on availability of machine and staff time. Alignment, variant calling, and imputation will be done within AGIL. In later years as costs drop, additional sequencing at lower coverage will proceed with Jersey, Brown Swiss, Ayrshire, or Guernsey breeds if sufficient array genotypes are available to indicate families or regions deserving further investigation. Breed associations often ask if unusual new phenotypes are inherited, and growing data sets allow discovery of additional lethal recessive defects. Public sequence data will be downloaded, and local data will be traded with other groups for animals not already included in the 1000 Bull Genomes Project. This will provide raw data to potentially identify new variants not already discovered. Data from the different sources of sequence data will be blended together using imputation because some variants identified in the new data will not be in the 1000 Bull Genomes data and some variants previously identified in the 1000 Bull Genomes data may have limited or no information in the local data.

All previous data should be realigned to a new **reference map** expected to replace UMD3.1 in the near future. However, much of the original sequence data may not be available. In that case, benefits from the new map can still be obtained from the new corrected locations using liftover tools for previous data, such as from genotyping arrays. Candidate variants will be re-selected using improved annotation, better bioinformatics, and additional information from discoveries across species. These methods will help to distinguish causal variants from linked markers. The new map will be checked against the large database of array genotypes before release to ensure that no regions are still mapped incorrectly as evidenced by excessive numbers of haplotypes, low correlations with neighboring variants, or double crossovers.

New **genotyping arrays** are being designed every year by several companies. Scientists at AGIL, with CDCB approval, will contribute to chip design by supplying lists of candidate variants estimated to have the largest effects. Most customized chips include known QTLs identified in previous literature and added by the genotyping laboratories. However, many of those QTLs have not yet been included in the list of variants routinely used in national genomic evaluations. After allowing sufficient time for data to accumulate and quality to be examined, AGIL scientists will incorporate the QTLs and candidate variants into larger or more optimal variant lists and estimate reliability gains from using the new lists (e.g., Wiggans et al., 2016a,b). Some reference bulls in the smaller breeds may need to be re-genotyped with the new arrays to avoid waiting to use the added variants for all breeds.

Genetic merit of animals could be changed directly by **gene editing** (Jenko et al., 2015). To make this technology profitable, more of the major genetic effects must be discovered as well as their true location rather than a linked marker. Separation of true effects from markers requires large sets of sequence data plus knowledge of gene function. Although AGIL does not intend to create edited animals in this project, breeders and breeding companies may need guidance on best strategies to include those animals in breeding programs, their potential value, or confirmation of the phenotypic effects of any gene edits. Simulation will be the main early tool because few edited animals are yet available; existing AGIL software (Cole, 2015) will be modified to support simulation of gene editing. Similarly, AGIL researchers have estimated selection limits given the haplotypes currently available (Cole and VanRaden, 2011) and have forecast additional long-term gains from selecting for favorable alleles with low-frequency alleles (Sun and VanRaden, 2014). Further simulation will reveal optimum strategies for combining favorable haplotypes to obtain higher progress in less time but without excessive inbreeding. The best strategies may be similar whether the favorable variants are newly introduced mutations, alleles already present in the population at low frequency, or alleles introgressed from another population.

Contingencies: If additional computational resources are needed for special tasks, such as realignment of a large quantity of sequence data, the ARS SCINet high-performance computing system, Ceres, may be used. Resources available on that system include 69 nodes with 1,560 cores, 15 TB of memory, and 1.4 PB of storage. Many journals now require that all sequence data used in the research be made public, but the Cooperative Dairy DNA Repository (CDDR) has not supported making data public. The options then are to publish only in the few remaining journals that do not require data upload, to not use DNA from the CDDR in research, or to not do the research; USDA favors making data freely available when possible.

Collaborations: The 1000 Bull Genomes Project regularly provides updated sequence genotypes to participating research groups that provide data (see Appendix 3). Other projects within AGIL also may generate raw sequence data to allow discovery of additional variants. Dr. Christian Maltecca from North Carolina State University will collaborate on research to identify genomic regions that may contain undesirable haplotypes (see Appendix 4).

Objective 2: Evaluate new traits that can all be predicted at birth from the same inexpensive DNA sample.

Hypothesis 2 (Goal): Determine if the national increase in economic progress is more valuable than the combined expenses of collecting data and computing genetic evaluations for several individual traits such as feed efficiency, milking speed, health conditions, age at first calving, gestation length, etc.

Experimental Design: Genetic evaluations will be easy to develop for traits that are **already measured** such as age at first calving, gestation length, and persistency, but some traits also may have low heritability or moderate economic value (Egger-Danner et al., 2014). Previous research evaluated gestation length as a trait of the service sire or the cow's sire (Norman et al., 2009), but reliability should be higher if evaluated by an animal model as a trait of the calf. Heat stress effects from preliminary research were significant for yield traits but small for most other traits (Wright and VanRaden, 2015). Careful choice of scale for reporting the interactions and extra education are needed because predicted regressions on temperature may be difficult for breeders to understand as rankings differ with each degree of temperature (or temperature-humidity index). Guidance will be needed for extrapolating to foreign climates because those may be more extreme than any U.S. climates used to derive the predictions.

Investment in data collection is needed before several other traits can be evaluated routinely. Industry funding is needed for ongoing trait measurement, and research to estimate economic values and determine the reliability expected for those new traits is the first step before committing funds. **Many decisions are determined by competition and marketability, and the total national profit deriving from a new trait may set an upper limit on the cost of recording. Options for choosing the most profitable set of animals to phenotype and genotype will be explored.** The CDCB budget for 2017 data acquisition was increased by $1.2 million annually compared with the 2016 budget and now includes feed intake and health trait data. Scientists at AGIL will develop data analysis methods for much of these new data and consult with CDCB on priorities for additional trait data to collect. Feed efficiency has the greatest potential benefit but also probably the greatest cost for continued collection after the initial 5-year, $5-million USDA investment (VandeHaar et al., 2016). Inclusion of foreign measurements of feed intake have improved reliability in research studies (de Haas et al., 2015), but data edits and statistical methods then become more complex because raw phenotypes of cows usually have not been exchanged across countries in the past, only genetic evaluations.

Lameness, health conditions, milking speed, and spectral image data from milk samples are some other **additional traits** likely to be more valuable than the costs of data collection and processing. Coordination and standardization are major obstacles to obtaining accurate predictions from the dispersed data that already exists. Although AGIL does not control the data for any of those traits, AGIL scientists can do much of the model development and editing needed with any new trait. Several of those traits already have local, regional, breed-specific, or foreign research studies that will guide the analysis of national data. Choosing national trait definitions similar to international definitions will also increase the reliability of combined genomic evaluations if global data eventually become available via MACE. Computer simulation will be used to determine the best combination of direct (e.g., feed intake) and indirect (e.g., mid-infrared milk spectral data) phenotypes to use for genetic improvement of traits that are difficult or expensive to measure.

On-farm sensors have the potential to provide large amounts of data in the future, but substantial problems preclude their routine use in genetic evaluations at this time (most notably a lack of certification and quality control guidelines). The International Committee for Animal Recording (ICAR) has established a Sensor Device Task Force to develop such guidelines, and that task force is interacting with the ICAR Functional Traits Working Group (of which Cole is a member) as well as industry groups such as the National DHI Association. The proposed work focuses on information that is currently available through certified laboratories and data recording centers because they represent the best opportunities for moving forward in the short-term. If new data become available during the 5-year period of the project, then appropriate statistical models will be used to compute evaluations for those traits (including test-day models) if needed.

Selection indexes require **economic values** calculated as the value of each trait while holding all other measured traits in the index constant and can include correlated effects on non-measured traits but not correlated effects on traits already in the index. Historical rather than current prices must be used because milk, feed, beef, and replacement prices vary widely from year to year, whereas the benefits from selection are often realized 5 years in the future. Averages over the past 4 years plus forecast prices for the next year are used, if available. Sensitivity of the index to each price will be calculated by adding or subtracting 10% of the forecast economic value. Some economic values such as extra hours and prices for on-farm labor associated with trait differences are not well estimated. Single-gene conditions can be included in merit indexes by estimating probabilities that each non-genotyped animal carries each lethal allele (e.g., Cole et al., 2015). However, some additive effects of those genotypes may already be included in quantitative traits. Mating programs will also be improved if carrier probabilities are routinely available. Gene content will be predicted for Mendelian recessives (see Appendix 5) using the methodology of Gengler et al. (2008), and cross-validation will be used to assess prediction accuracy for non-genotyped animals. The effect on annual rates of genetic gain from adding recessives to the index will be studied using selection index methodology (Cameron, 1997).

The proposed research on gene editing will attempt to answer a number of questions, including the optimal proportion of bulls and elite cows to edit. The focus will be on genetic defects rather than QTLs because many of the former have been identified but few of the latter. Polledness is not a genetic defect but is of growing importance, and research also will address the best way to increase the frequency of polled animals in the population using gene editing. This will avoid overlap with the work of Jenko et al. (2015) because their focus was on hypothetical technologies that permit editing of many known QTLs in parallel. Gene editing may have little effect on estimation of breeding values but could perhaps affect evaluation reliability if many loci were edited and the resulting animal's true genetic merit was predicted with less uncertainty.

Studies to recommend **management practices** and document phenotypic trends and effects are not the main focus of this project. However, many such questions can be answered more precisely using AGIL's direct access to CDCB national data rather than by designed, small-scale experiments at other research locations. New reports will document and benchmark the incidence, correlations, and effects of more traits as data sources are added. For example, the increased use of gender-selected semen to produce dairy heifers has allowed more dairy cows to be bred to beef bulls. This allows the opportunity to rank beef bulls for fertility using data from dairy herds because no evaluations of fertility are yet available from beef herds. Separation of breed differences from heterosis is not possible from the beef crosses because purebred beef bull fertility is not available; however, breed differences and heterosis for dairy bull fertility will be estimated using crossbred matings (primarily Jersey by Holstein) and an all-breed model instead of current within-breed models. Constant monitoring of input data in cooperation with data providers is needed to ensure continued high quality of evaluations. For example, increased use of embryo transfer and in-vitro fertilization may cause biases if not properly accounted for in existing evaluations for calving traits.

Contingencies: New investments in data collection depend on industry funding and priorities. Decisions regarding which traits to collect and exact timing of any new evaluations produced are under CDCB rather than USDA control. The new NFCA does not allow AGIL to directly release any trait evaluations for individual animals from the shared database **or make public the data received from CDCB**. Scientists within AGIL are generally free to publish research from the shared database provided that CDCB is given an opportunity to review draft publications at least 30 days prior to submission (see Appendix 1, Article 3). Health trait data may be regarded as too sensitive and private to be copied onto U.S. Government computers. **Access to future data is not guaranteed but is critical to continuing the close ties between research and application. The NFCA with CDCB expires in 2017, and a new 5-year agreement with CDCB has been drafted. The relationship between AGIL and CDCB staff is very positive and productive. If dramatic changes in leadership at CDCB or AGIL result in a deterioration of the close relationship, other partnerships could be developed. For example, dairy records processing centers and AI companies can offer evaluations that other companies do not, and CRADAs could be developed to foster the uptake of AGIL research by the industry.** Ongoing interaction with **industry representatives** and revised estimates of trait value may shift research priorities from the listed traits to others with greater economic importance to the dairy industry.

Collaborations: Routine genomic evaluations for feed efficiency will be developed in close cooperation with CDCB under a nonfunded cooperative agreement (see Appendix 1); CDCB is negotiating with Drs. Mike VandeHaar of Michigan State University, Kent Weigel of the University of Wisconsin, Erin Connor of AGIL, and other research groups for continued feed intake collection after the $5 million initial project. This feed efficiency will be shared with the proposed project to test evaluation methods and to estimate the economic value. Mid-infrared milk spectral data for cows from the Beltsville research herd will be collected by Dr. Erin Connor of AGIL. Dr. John Cole of AGIL is a participant in SCC84 (Genetic Selection and Mating Strategies to Improve the Well-Being and Efficiency of Dairy Cattle), a multistate research project, southern region (https://www.nimss.org/projects/15456). As in the past, other participants are collecting health incidence data and treatment costs from their research herds to help in estimating economic values of traits and updating merit indexes. **Many of the traits listed under Objective 2 were requested specifically by stakeholders at the National Dairy Genetics Workshop in 2014 and during the listening process held prior to writing this proposal in 2016. Input will continue to be collected from stakeholders such as DHI association members, purebred cattle associations, AI companies, and CDCB through**

correspondence and face-to-face meetings. Traits will be prioritized in order of data availability so that new tools can be delivered to the industry as quickly as possible. The proposed project's lead scientist participates in monthly meetings and conference calls with CDCB staff and the Board of Directors to determine data collection and analysis strategies. All scientists in the project participate in annual or semi-annual meetings with representatives from AI companies, breed associations, and DHI associations to determine research priorities and to set policies for implementing new trait evaluations.

Objective 3: Improve efficiency of genomic prediction and computation by developing faster algorithms, testing new adjustments and models, and accounting for genomic pre-selection in evaluation.

Hypothesis 3: Developing, testing, and implementing new computational methods can improve accuracy with little extra cost or provide the same or similar accuracy with reduced cost.

Experimental Design: Algorithms will be developed to improve efficiency of **aligning sequence** segments to a reference genome and simultaneously calling the variants (differences) of each animal's DNA as compared to the reference. Preliminary tests have shown large improvements in speed (over 30 times faster) and also small improvements in alignment accuracy when applied to simulated DNA reads (VanRaden and Bickhart, 2016). Further tests will apply the new algorithms to data from hundreds of actual animals or thousands of simulated animals and set optimal limits for separating true variants from read errors. The new programs do not yet have the full features to allow compatibility with many existing software tools, and some of those features will be added to allow wider adoption of the methods.

Genomic models including more **informative priors** could obtain higher accuracy from the same data (Gianola, 2013). Candidate variants near genes or predicted to directly affect protein structure should get more prior weight than the random markers previously selected with equal spacing and high minor allele frequency. The previous markers still need to be included for imputation because most reference animals will not be re-genotyped for the candidate variants, but when estimating SNP effects the previous markers can be assigned less variance and the candidate variants or known QTLs more variance. Tests will compare predictive ability for future data within the same breed or to predict animals of a different breed. Multibreed estimation of marker effects as correlated traits will be attempted after sufficient numbers of genotypes from each breed are available to allow accurate imputation of the new variants. Gene expression levels can also indicate which variants may be causal if sufficiently large data sets become available.

Potential biases from **genomic pre-selection** will be monitored using differences across time in percentages of genotyped mates or daughters for each bull (VanRaden and Wright, 2013). Use of single-step models to correct the bias will be further explored using recent algorithms to approximate the inverse of genomic relationships or to model SNP effects directly. Because the evaluation already includes genotypes for 1.5 million Holsteins, the methods must scale to at least several million to justify implementation. Single-step algorithms continue to be developed at the University of Georgia and Iowa State University (Fernando et al., 2014), in Finland (Koivula et al., 2015) and France, and by other groups. Dr. Yutaka Masuda from the University of Georgia has tested their new programs directly on AGIL computers applied to CDCB data (Masuda et al., 2016b), but researchers have not yet obtained better results for single-step than the current multistep programs when applied to the full national data. **Single-step methods originally were described as simpler than multistep; however, the coding has become much more complex in recently derived algorithms, and code sharing will become more important. Coordination of code use may also involve CDCB because**

some groups provide free code only for research use but not for routine evaluations. Potential biases in international evaluations will also be investigated in cooperation with an Interbull working group.

Genomic evaluations of **crossbred animals** have not been available but will be developed by weighting the marker effects from each breed by the animal's genomic breed composition. A preliminary test was conducted only for one trait (milk yield) and with no foreign data included. Routine evaluations will require moving all genomic calculations to the all-breed scale instead of the current calculations on separate within-breed scales. Net merit should also be computed on the all-breed scale, but conformation and calving traits are only evaluated separately by breed. Approximate conversion formulas will be developed to rank crossbreds for those traits using math similar to that of VanRaden and Sanders (2003). Foreign data from MACE and from breeding values of dams will be converted from the within-breed to all-breed scale. Publishing evaluations directly on the all-breed scale may be a simpler option, and online tools to predict merit of crossbreds by combining breed differences and heterosis are needed to guide producers on mating programs. Research will also continue on predicting nonadditive effects, for example specific heterosis for each major breed combination instead of general heterosis across breeds, **and perhaps also recombination loss**. **Genomic future inbreeding will be improved by computing each animal's average genomic relationship to a more recent group of potential mates instead of to the breed's reference population.**

Different models may be needed for **new traits** as compared to models applied to previous traits. **Test-day models will be considered when they are the appropriate evaluation methodology, such as for longitudinal data.** Adjustments for additional factors will be tested using truncated data to predict more recent data. Correlations between domestic and foreign evaluations will also be used to test reliability and consistency of trait definitions. Most new traits are not yet exchanged in MACE, but some foreign evaluations may be available for testing accuracy of national models. Multitrait processing of national data will be used to obtain greater benefits from new traits without losing information from previous, correlated traits (VanRaden et al., 2014). In some cases, potential gains in reliability from multitrait modeling will be estimated from theory or from simulated data if accurate parameter estimates are available from the literature or from exploratory data sets.

Contingencies: Growth of genomic data and speed of changing reproductive or selection practices may affect priorities for algorithm development. Availability of computing resources and ability to transfer data also impact the algorithms needed. Computer hardware for routine processing may differ from that used for research projects, especially if remote computing systems are used. **Moderate-scale sequence analyses can be conducted with current in-house servers plus additional disk space. Large-scale projects (hundreds of animals) will require the use of additional servers already available via other projects within AGIL or remote access to the high-performance computing resources SCINet and Ceres, which are available to ARS researchers via a dedicated high-speed Internet2 network. Analyses of CDCB phenotype and genotype files require routine (weekly) updating using a database that is much more difficult to transfer to a remote site. Currently the project's hardware matches or exceeds CDCB hardware, but servers and disk space will both need to double about every 3 years to keep up with expected data growth.**

Collaborations: Research on use of additional prior information will be conducted in cooperation with Drs. Jeff O'Connell and Li Ma of the University of Maryland, who have direct access to AGIL computers as part of a non-assistance cooperative agreement (see Appendix 6) and a reimbursable cooperative agreement (RCA; see Appendix 7). Dr. Derek Bickhart of the Dairy Forage Research Center, ARS, USDA, will continue to provide advice on alignment and variant calling strategies as an investigator in the CDCB NFCA (see Appendix 1), the University

of Maryland NCA (see Appendix 6) and RCA (see Appendix 7). **A number of university groups (e.g., North Carolina State University, University of Guelph, and University of Minnesota) are doing research on inbreeding concerns, and AGIL scientists are collaborating to determine where homozygosity in the genome is harmful.**

Physical and Human Resources

Physical Resources

The facilities utilized in Building 005, BARC-West, are satisfactory to meet most needs of the project. Current rented space averages about 200 square feet per employee. More offices are available if needed. The occupied space has suitable environmental control, electrical capacity, and network connectivity. Additional AGIL computer hardware and office space is available in Building 306, BARC-East, primarily for processing sequence data.

Primary computer support has been obtained from an IBM xSeries 3850 server. The server has four 64-bit Intel(R) Xeon(R) X7560 dies, providing a total of 64 computing threads running at 2.27 GHz. This machine has a total of 640 GB of memory and approximately 20 TB of directly attached storage. A second server, an HP 580DL Gen8, is used as a database (DB2) and SAS server. This workstation has two 64-bit Intel(R) Xeon(R) CPU E7-8893 v2 dies, providing a total of 24 computing threads running at 3.40 GHz. This machine has a total of 256 GB of memory and approximately 30 TB of directly attached storage. Most of the storage for these machines is managed by an 8-GB fiber-connected IBM v3700 storage area network. Also available is a slightly slower, but inexpensive, mass-storage array based on designs by Backblaze, an online storage company. That device currently provides AGIL with 22 TB of usable disk space and easy future expansion. An IBM xSeries 3550 M4 server with dual Intel(R) Xeon(R) E5-2609 2.4-GHz processors is used to support a Tivoli Storage Manager (TSM) backup system.

These servers share a closed 10-GB fiber/ethernet network for protected communication and filesystem sharing. A separate SuperMicro server runs web and file services. A Linear Tape-Open technology library with two generation-five drives and a capacity of 42 tapes is used for data backup and archive. The library has a storage capacity of approximately 126 TB. Fifteen personal computers are available for employee use; 20% of those computers are upgraded each year. Laptops have been provided to employees who telecommute on a regular basis. A 1-GB local-area network is used to communicate among personal computers, and between those computers and the workstations. Additional computing resources include a Pogo Linux Atlas 1205 workstation with two 8-core AMD Opteron 6328 processors and 128 GB of RAM, and two Thinkmate RAX QS6-4210 computer servers with four 12-core AMD Opteron 6344 processors and 256 GB of RAM each. Key software includes SAS for UNIX version 9.1 and Intel Fortran and C compilers with high-performance numerical libraries.

Scientists at AGIL have access to the national database from CDCB with over1.5 million genotyped animals, 30 million phenotyped animals, and 70 million animals in the pedigree file as well as next-generation sequence data for many animals and the variant calls from run5 of the 1,000 Bull Genomes Project.

Human Resources

The project is conducted by a highly coordinated team of AGIL researchers, research support personnel, and data-processing experts. Two vacancies were created in 2016 with the retirement of Dr. George Wiggans and the departure of Dr. Derek Bickhart:

Personnel[1]	Scientist Years	Full-time Equivalents	Service years with Laboratory
Scientists			
VanRaden, P.M.	1.00	. . .	28
Cole, J.B.	1.00	. . .	13
Van Tassell, C.P.	0.25	. . .	20
Research Geneticist (vacant)	1.00	. . .	0
Research Geneticist (vacant)	1.00	. . .	0
Total	**4.25**		
Support scientists	. . .	5.0	19 (average)
Information technology support	. . .	3.0	25 (average)
Administrative	. . .	1.0	20
Total		**9.0**	

[1]As of January 12, 2017.

Project Management and Evaluation

This project involves collaborative research among four AGIL research geneticists. The project leader is responsible for coordinating research in each sub-objective to ensure that milestones are met. Weekly information-exchange meetings are held with all project employees and are followed by a brief research update and planning session. Each research geneticist has authority to identify research goals and direction and is responsible for planning, designing, and executing research related to assigned objectives, including analyzing, interpreting, and reporting results. The Research Leader is kept informed of general plans and results and reviews all manuscripts before submission for Center approval. Manuscripts are submitted for approval under prescribed ARS procedures and are typically accepted as technically sound. Review of overall project research results is primarily through annual progress reports, performance evaluations, and periodic project reviews. Contingency plans will be implemented if necessary after a full discussion has been held among impacted scientists and consensus is reached. No management problems are anticipated.

Milestones

Project Title		Improving Dairy Animals by Increasing Accuracy of Genomic Prediction, Evaluating New Traits, and Redefining Selection Goals		
Project No.		8042-31000-101-00D		
National Program		101: Food Animal Production		
Objective		1 - Expand genomic data used in prediction by selecting new variants that more precisely track the true gene mutations that cause phenotypic differences.		
NP Action Plan Component		2: Understanding, improving, and effectively using animal genetic and genomic resources		
NP Action Plan Problem Statement		2A: Develop bioinformatic and other required capacities for research in genomics and metagenomics; 2B: Characterize functional genomic pathways and their interactions; 2D: Develop and implement genetic improvement programs using genomic tools		
Goal/Hypothesis		Hypothesis: Addition of new variants from sequence data will improve reliability of genomic predictions compared with that from single-nucleotide polymorphism markers currently used in genotyping arrays.		
SY Team	Months	Milestone	Anticipated Product	Progress/Changes
Vacant, CPV, PMV	12	Obtain whole genome sequence data for 200 additional Holstein and 50 Jersey bulls	New variants in U.S. dairy cattle not already discovered by global researchers	*This column for plan management after peer review.*
JBC	24	Routinely calculate gene content for nongenotyped cows for known Mendelian traits	Peer-reviewed publication; new information for distribution to animal owners	
PMV, Vacant	36	Impute and select new variants from run 6 of the 1000 Bull Genomes Project	Revised genotyping arrays with additional variants that more closely track genetic effects	
JBC, Vacant	48	Investigate potential to use edited genes or combine the best chromosomes into one animal by simulation	Peer-reviewed publication(s) on simulation study results	
Vacant, PMV	60	Obtain and evaluate whole-genome sequence data for Brown Swiss, Ayrshires, Guernseys, and additional Jerseys	Breed-specific variants and improved prediction reliability for all breeds	

Objective		2 - Evaluate new traits that can all be predicted at birth from the same inexpensive DNA sample.		
NP Action Plan Component		1: Increasing Production and Production Efficiencies while Enhancing Animal Well-Being across Diverse Food Animal Production Systems; 2: Understanding, improving, and effectively using animal genetic and genomic resources		
NP Action Plan Problem Statement		1A: Improving the efficiency of growth and nutrient utilization; 1B: Improving reproductive efficiency; 1C: Enhancing animal well-being and reducing stress; 2A: Develop bioinformatic and other required capacities for research in genomics and metagenomics; 2B: Characterize functional genomic pathways and their interactions; 2D: Develop and implement genetic improvement programs using genomic tools		
Goal/Hypothesis		Goal: Determine if the national increase in economic progress is more valuable than the combined expenses of collecting data and computing genetic evaluations for several individual new traits.		
SY Team	Months	Milestone	Anticipated Product	Progress/Changes
PMV	12	Evaluate gestation length as new trait	Genetic rankings distributed worldwide by the Council on Dairy Cattle Breeding	*This column for plan management after peer review.*
JBC, Vacant	24	Incorporate recessive traits into economic indexes	Genetic rankings for economic merit which account for Mendelian traits (e.g., polled) for industry use	
JBC, PMV, Vacant	36	Estimate potential economic values and develop genetic rankings for additional new traits	Peer-reviewed publication on economics and genetic rankings of new traits; industry guidance on investment in data collection	
	48			
JBC, PMV	60	Update net merit formula with new traits and changing prices	Improved selection index for industry use	
Objective		3 - Improve efficiency of genomic prediction and computation by developing faster algorithms, testing new adjustments and models, and accounting for genomic pre-selection in evaluation.		
NP Action Plan Component		2: Understanding, improving, and effectively using animal genetic and genomic resources		
NP Action Plan Problem Statement		2A: Develop bioinformatic and other required capacities for research in genomics and metagenomics; 2B: Characterize functional genomic pathways and their interactions; 2D: Develop and implement genetic improvement programs using genomic tools.		
Goal/Hypothesis		Hypothesis: Developing, testing, and implementing new computational methods can improve accuracy with little extra cost or provide the same or similar accuracy with reduced cost.		
SY Team	Months	Milestone	Anticipated Product	Progress/Changes
PMV, DMB	12	Apply new algorithms in Findmap and Findvar software to large-scale cattle sequence data	New software to manage rapidly growing data efficiently	*This column for plan management after peer review.*
JBC, PMV	24	Evaluate crossbred animals by combining purebred marker effects weighted by breed composition	Peer-reviewed publication on genomic evaluation of crossbred cattle; new technology to transfer to Council on Dairy Cattle Breeding	
JBC, Vacant	36	Determine optimal combinations of directly measured and correlated phenotypes for novel traits	Peer-reviewed publication	
PMV, DMB	48	Develop and compare prior information for weighting different types of variants (markers vs. quantitative trait loci)	Software; improved genomic evaluations	
PMV, Vacant	60	Test potential bias in trait evaluations caused by genomic preselection of mates, progeny, and herdmates	Peer-reviewed publication on genomic evaluation bias caused by preselection; possible revisions in genetic evaluation software	

Accomplishments from Prior Project Period

Terminating ARS research project number: 8042-31000-101-00D

Title: Improving genetic predictions for dairy animals using phenotypic and genomic information

Project period: July 24, 2012 – July 23, 2017

Investigators and FTE:	Project start	Project end
Paul M. VanRaden, Lead Scientist	1.00	1.00
John B. Cole	1.00	1.00
Derek M. Bickhart (hired in 2012; left AGIL in 2016)	—	—
George R. Wiggans (retired in 2016)	1.00	—
Curtis P. Van Tassell	0.00	0.00
Tad S. Sonstegard (left ARS in 2015)	0.00	—
Research Geneticist (vacant)	1.00	—
Research Geneticist (vacant)	—	1.00
Research Geneticist (vacant)	—	1.00

Prior accomplishments and impacts as related to proposal objectives

Proposal Objective 1: Expand genomic data used in prediction by selecting new variants that more precisely track the true gene mutations that cause phenotypic differences.

- **Introduction of free genetic tests for inherited defects of dairy cattle.** A method to identify exact locations of loss-of-function mutations and DNA sequences associated with lethal or undesirable conditions of dairy cattle was developed and automated. However, results from that method could not be made available to the dairy industry for DNA sequences associated with patented genes. Although genetic tests were available for several lethal mutations, most females were not tested because individual gene tests were expensive and had not been included on genotyping chips. The June 2013 U.S. Supreme Court unanimous decision that biotechnology companies cannot patent genes that occur naturally has made possible the release of information from genetic tests for bovine leukocyte adhesion deficiency (BLAD), deficiency of uridine monophosphate synthase (DUMPS), and mulefoot in Holsteins and Weaver Syndrome, spinal dysmyelination (SDM), and spinal muscular atrophy (SMA) in Brown Swiss; for Holsteins, the method also can be applied to identify DNA markers associated with complex vertebral malformation (CVM) and brachyspina as well as for desired traits such as red coat color and polledness (no horns). In addition, four new deleterious DNA sequences were identified for dairy cattle fertility, and those sequences were incorporated into new genotyping chips. The first release of genomic status information for the inherited defects was in August 2013, and the list will continue to be updated, potentially with new defects discovered directly from sequence data. Dairy producers now can reduce or eliminate costs for genetic testing, decrease the frequency of undesired traits, and increase the rate of genetic progress for desired traits.
  Publications:
  Cole, J.B., P.M. VanRaden, D.J. Null, J.L. Hutchison, T.A. Cooper, and S.M. Hubbard. 2013. Haplotype tests for recessive disorders that affect fertility and other traits. AIP Res. Rep. Genomic3 (09-13). https://aipl.arsusda.gov/reference/recessive_haplotypes_ARR-G3.html. [Last update in May 2016]

Lawlor, T.J., P.M. VanRaden, D. Null, J. Levisee, and B. Dorhorst. 2014. Using haplotypes to unravel the inheritance of Holstein coat color. Proc. World Congr. Genet. Appl. Livest. Prod., Commun. 289.

McClure, M.C., D. Bickhart, D. Null, P. VanRaden, L. Xu, G. Wiggans, G. Liu, S. Schroeder, J. Glasscock, J. Armstrong, J.B. Cole, C.P. Van Tassell, and T.S. Sonstegard. 2014. Bovine exome sequence analysis and targeted SNP genotyping of recessive fertility defects BH1, HH2, and HH3 reveal causative mutation in *SMC2* for HH3. PLoS ONE 9:e92769.

McClure, M., E. Kim, D. Bickhart, D. Null, T. Cooper, J. Cole, G. Wiggans, P. Ajmone-Marsan, L. Colli, E. Santus, G.E. Liu, S. Schroeder, L. Matukumalli, C. Van Tassell, and T. Sonstegard. 2013. Fine mapping for Weaver Syndrome in Brown Swiss cattle and the identification of 41 concordant mutations across NRCAM, PNPLA8 and CTTNBP2. PLoS ONE 8:e59251.

Sonstegard, T.S., J.B. Cole, P.M. VanRaden, C.P. Van Tassell, D.J., Null, S.G. Schroeder, D. Bickhart, and M.C. McClure. 2013. Identification of a nonsense mutation in CWC15 associated with decreased reproductive efficiency in Jersey cattle. PLoS ONE 8:e54872.

VanRaden, P.M., C. Sun, T.A. Cooper, D.J. Null, and J.B. Cole. 2014. *Keynote presentation III:* Genotypes are useful for more than genomic evaluation. Proc. 39th Int. Comm. Anim. Recording Sess., 4 pp.

- **Identification of specific chromosomal regions with significant effects on economically important traits.** A granddaughter design is an analysis of genetic linkage of quantitative loci to DNA markers in which the markers are identified in grandsires and sons but the quantitative analysis is carried out on granddaughter performance. The granddaughter design has been applied to nearly all major commercial dairy cattle populations, but relatively little practical use had been made of the results. A granddaughter design was applied to the entire Holstein genome to determine specific chromosomal regions (haplotypes) that had significant effects on 33 economically important traits, including production, disease resistance, longevity, fertility, calving, conformation, and overall net merit. Each trait was found to have at least one significant haplotype within family. This identification method has been beneficial in providing information on gene function and the architecture of quantitative trait loci as well as determining the causative genetic variant for desired traits. However, the causative variant for a major genetic effect on cow size and calving traits was difficult to discover because the current reference map may not reflect the true chromosome structure.

  Publications:

Cole, J.B., J.L. Hutchison, D.J. Null, P.M. VanRaden, G.E. Liu, S.G. Schroeder, T.P. Smith, T.S. Sonstegard, C.P. Van Tassell, and D.M. Bickhart. 2014. The hunt for a functional mutation affecting conformation and calving traits on chromosome 18 in Holstein cattle. Proc. 10th World Congr. Genet. Appl. Livest. Prod., Commun. 304.

Weller, J.I., J.B. Cole, P.M. VanRaden, and G.R. Wiggans. 2014. Application of the a posteriori granddaughter design to the Holstein genome. Animal 8:511–519.

Weller, J.I., P.M. VanRaden, and G.R. Wiggans. 2013. Application of a posteriori granddaughter and modified granddaughter designs to determine Holstein haplotype effects. J. Dairy Sci. 96:5376–5387.

Wiggans, G.R., and J.I. Weller. 2015. Revisiting the "a posteriori" granddaughter design. Interbull Bull. 49:36–42.

- **Improved accuracy of genomic evaluations for dairy cattle through use of more DNA markers**. Accuracy of genomic evaluation is expected to increase when more DNA markers are used because of better tracking of causative genetic variants. However, high-density genotypes based on almost 800,000 markers had not been used for U.S. genomic evaluations because the small accuracy gain achieved did not justify the genotyping cost. To investigate the use of more than 50,000 markers in genomic evaluation, two genotyping chips were developed in cooperation with Neogen Corporation containing approximately 77,000 markers and later with more than 140,000 markers specifically chosen to be highly informative with largest effects. Two new lower-density chips also included some of the selected markers. Use of additional markers resulted in 1 to 2 percentage points more accurate genomic prediction for most traits. Further increases in numbers of markers should also consider accuracy of imputation, which will improve as more animals are genotyped with these new chips. The Council of Dairy Cattle Breeding implemented 61,000 markers in December 2013 and plans to implement 77,000 markers as the standard for national evaluations in 2017. Use of a larger set of markers in genomic evaluations allows dairy producers to make more accurate breeding selections for economically important traits.

  Publications:

  Wiggans, G.R., T.A. Cooper, D.J. Null, and P.M. VanRaden. 2014. Increasing the number of single nucleotide polymorphisms used in genomic evaluations of dairy cattle. Proc. 10th World Congr. Genet. Appl. Livest. Prod., Commun. 301.

  Wiggans, G.R., T.A. Cooper, and P.M. VanRaden. 2013. Using 90,113 single nucleotide polymorphisms in genomic evaluation of dairy cattle. J. Dairy Sci. 96(E-Suppl. 1):617. (Abstr.).

  Wiggans, G.R., T.A. Cooper, P.M. VanRaden, C.P. Van Tassell, D.M. Bickhart, and T.S. Sonstegard. 2015. Effect of increasing the number of single nucleotide polymorphisms from 60,000 to 85,000 in genomic evaluation of Holsteins. J. Dairy Sci. 98(Suppl. 2):578. (Abstr.)

  Wiggans, G.R., T.A. Cooper, P.M. VanRaden, C.P. Van Tassell, D.M. Bickhart, and T.S. Sonstegard. 2016. Increasing the number of single nucleotide polymorphisms used in genomic evaluation of dairy cattle. J. Dairy Sci. 99:4504–4511.

- **National genomic evaluations for Ayrshire and Guernsey dairy cattle**. Although genetic evaluations of the Holstein, Jersey, and Brown Swiss dairy breeds in the United States have included genomic information since 2009, too few Ayrshires and Guernseys had been genotyped to allow genomic evaluation. In 2013, data from over 1,100 genotyped Ayrshires with performance and pedigree records in the North American database made possible the development of genomic evaluations for Ayrshires. Compared with traditional parent averages, those evaluations improved accuracy of prediction of genetic merit by 8 percentage points over all traits. In addition, a DNA segment (haplotype) that affects fertility was discovered on chromosome 17; sire conception rate was 3 percentage points lower for carriers of the haplotype, and the carrier frequency for genotyped Ayrshires is 23%. Genomic predictions for Guernseys were implemented in 2016 in cooperation with the United Kingdom and the Isle of Guernsey, and provided an average reliability gain of 17 percentage points. Multibreed estimation of marker effects could improve these gains in the future. Ayrshire and Guernsey breeders now can make better selection decisions and increase the rate of genetic gain for economically important traits as a result of improved knowledge of the genomic makeup and merit of their animals.

  Publications:

  Cooper, T.A., S.A.E. Eaglen, G.R. Wiggans, J. Jenko, H.J. Huson, D.M. Morrice, M. Bichard, W.G. de L. Luff, and J.A. Woolliams. 2016. Genomic evaluation, breed identification, and

population structure of Guernsey cattle in North America, Great Britain, and the Isle of Guernsey. J. Dairy Sci. 99:5508–5515.
Cooper, T.A., G.R. Wiggans, D.J. Null, J.L. Hutchison, and J.B. Cole. 2014. Genomic evaluation, breed identification, and discovery of a haplotype affecting fertility for Ayrshire dairy cattle. J. Dairy Sci. 97:3878–3882.

- **More accurate genetic evaluation of Jersey cattle through international exchange of bull genotypes.** Within a single country, the number of progeny-tested bulls is limited, especially for small dairy cattle populations, and achieving an adequately sized reference population for genomic selection is a challenge. Accuracy of genomic prediction was improved by exchanging genotypes of 1,168 evaluated Jersey bulls with Viking Genetics, which serves Denmark, Sweden, and Finland. The average increase in evaluation accuracy was 1.8 percentage points for young U.S. animals as compared with previous genomic predictions from North American data; routine use of these genotypes began in January 2014. Including U.S. Jersey bulls in the Danish Jersey reference population increased evaluation accuracy for young Danish animals by an average of 4.0 percentage points (range from 1.6 to 12.5) for milk, fat, and protein yields; fertility; mastitis; longevity; body conformation; and feet and legs compared with using the Danish reference population alone. The larger international reference population increased the accuracy of genomic prediction, especially for the Danish population, which had fewer reference bulls, and is allowing genetic progress in the Jersey breed to remain competitive with progress made by the much larger Holstein population.

  Publications:

  Su, G., P. Ma, U.S. Nielsen, G.P. Aamand, G.R. Wiggans, B. Guldbrandtsen, and M.S. Lund. 2016. Sharing reference data and including cows in reference population improve genomic predictions in Danish Jersey. Anim. 10:1067–1075.
  Su, G., U.S. Nielsen, G. Wiggans, G.P. Aamand, B. Guldbrandtsen and M.S. Lund. 2014. Improving genomic prediction for Danish Jersey using a joint Danish-US reference population. Proc. 10th World Congr. Genet. Appl. Livest. Prod., Commun. 060.
  Wiggans, G.R., G. Su, T.A. Cooper, U.S. Nielsen, G.P. Aamand, B. Guldbrandtsen, M.S. Lund, and P.M. VanRaden. 2015. *Short communication:* Improving accuracy of Jersey genomic evaluations in the United States and Denmark by sharing reference population bulls. J. Dairy Sci. 98:3508–3513.

- **Improved detection of DNA mutations in whole genome sequencing.** Several types of structural DNA variations remain difficult to detect within sequenced genomes. Mutations that affect traits of interest often are not simple differences in single-nucleotide polymorphisms but instead are deletions or insertions that can result in false positive detection. Software was developed for the detection of such variants using the orientation and distance of paired-end and split-read mappings in whole genome DNA sequence data. In simulations and with real data, the method was 27.5 times more precise than two competing programs in detecting tandem duplications and also was able to detect twice the number of duplications. This high degree of precision enables better functional prediction of structural DNA variants from short-read sequence data and allows discovery of many more of the actual mutations that affect traits. Investigators from the University of Missouri, Texas A&M University, and Australia's La Trobe University already are using this software for research projects. To detect such mutations more efficiently, improved methods were also developed for deciding which animals to sequence. Those methods use the frequency of haplotypes (groups of alleles that are inherited together from a single parent) to determine the most cost-effective set of bulls to sequence and will be used for future projects.

Publications:

Bickhart, D.M., J.B. Cole, J.L. Hutchison, L. Xu, and G.E. Liu. 2014. Using the whole read: Structural variant detection using NGS data. Proc. 10th World Congr. Genet. Appl. Livest. Prod., Commun. 164.

Bickhart, D.M., J.L. Hutchison, D.J. Null, P.M. VanRaden, and J.B. Cole. 2016. Reducing animal sequencing redundancy by preferentially selecting animals with low-frequency haplotypes. J. Dairy Sci. 99:5526–5534.

Bickhart, D.M., J.L. Hutchison, L. Xu, R.D. Schnabel, J.F. Taylor, J.M. Reecy, S.G. Schroeder, C.P. Van Tassell, T.S. Sonstegard, and G. Liu. 2015. RAPTR-SV: a hybrid method for the detection of structural variants. Bioinformatics. 31:2084–2090.

- **Selection of DNA sequence variants to improve reliability of genomic predictions for dairy cattle.** Effective genomic selection depends on the availability of genotypic information for DNA variants that affect traits of economic importance. The national database maintained by the Council on Dairy Cattle Breeding was updated by AGIL to include genotypes from several new arrays, for a total of 23 arrays now. A cooperative effort was initiated in Australia to exchange sequence data of cattle internationally. Millions of additional DNA sequence variants from the 1,000 Bull Genomes Project were examined by AGIL, and 5,000 variants with largest estimated effects were added to customized genotyping arrays of 3 different companies. A data file for 440 Holstein bulls was merged with phenotypes and array genotypes for 27,000 other Holstein bulls to estimate effects of the newly discovered variants for 33 dairy cow traits. The new genotype arrays that include these variants will allow the dairy industry to select directly for causative mutations instead of using indirect selection via genetic markers.

  Publications:

  Tooker, M.E., P.M. VanRaden, D.M. Bickhart, and J.R. O'Connell. 2016. Selection of sequence variants to improve dairy cattle genomic predictions. J. Dairy Sci. 99(E–Suppl. 1):138. (Abstr.)

  VanRaden, P.M., and J.R. O'Connell. 2015. Strategies to choose from millions of imputed sequence variants. Interbull Bull. 49:10–13.

  Wiggans, G.R., P.M. VanRaden, D.M. Bickhart, and M.E. Tooker. 2016. Strategy for incorporating newly discovered causative genetic variants into genomic evaluations. J. Dairy Sci. 99(E–Suppl. 1):137. (Abstr.)

Proposal Objective 2. Evaluate new traits that can all be predicted at birth from the same inexpensive DNA sample.

- **Multitrait fertility evaluation for dairy cattle.** Fertility traits of dairy cattle can benefit from multitrait processing because of high genetic correlations among traits and many missing observations. Instead of the previous single-trait, single-breed models, heifer and cow conception rates (HCR and CCR, respectively) were evaluated with multitrait, multibreed models using 4 million HCR and 14 million CCR lactation records stored since 2003 and 66 million daughter pregnancy rate (DPR) records collected since 1960. Conception rates are now pre-adjusted for environmental effects and combined into lactation records for simpler multitrait analysis with DPR. Crossbred cows are now included and get the combined effects of heterosis and no inbreeding compared with purebreds, which may average 6% inbreeding. Evaluations from the new and previous models were correlated by more than 95% for both HCR and CCR for recent Holstein bulls that had a reliability of more than 50% but were less correlated for other breeds because of additional crossbred daughters and contemporaries. The new model that combines data from all breeds and from correlated traits to improve genetic evaluations for fertility was implemented in December 2013. More accurate

predictions of genetic merit for fertility traits and for other traits in the future will allow breeders to make faster progress.

Publications:

VanRaden, P.M., M.E. Tooker, J.R. Wright, C. Sun, and J.L. Hutchison. 2014. Comparison of single-trait to multi-trait national evaluations for yield, health, and fertility. J. Dairy Sci. 97:7952–7962.

VanRaden, P.M., J.R. Wright, C. Sun, J.L. Hutchison, and M.E. Tooker. 2014. Multi-trait, multi-breed conception rate evaluations. J. Dairy Sci. 97(E-Suppl. 1):472. (Abstr.)

- **Introduction of genetic-economic selection index for pasture-based dairy cattle.** Pasture-based dairy producers have costs, revenue streams, and management challenges that often differ from those associated with conventional dairy production systems and consequently need a selection index designed specifically for breeding grazing herds. In collaboration with Purdue University (West Lafayette, IN), an index was developed to rank animals based on "grazing merit" in conjunction with economic updates to the traditional lifetime merit indexes; all of the indexes also included heifer and cow conception rates as traits of economic importance for the first time. The new and updated indexes were implemented by the Council of Dairy Cattle Breeding in December 2014 along with a genetic base change. The updated indexes promote balanced selection to ensure maximum economic progress for herds with differing milk markets.

  Publications:

  Gay, K.D., N.J.O. Widmar, T.D. Nennich, A.P. Schinckel, J.B. Cole, and M.M. Schutz. 2014. Development of a lifetime merit-based selection index for US dairy grazing systems. J. Dairy Sci. 97:4568–4578.

  VanRaden, P.M., and J.B. Cole. 2014. Net merit as a measure of lifetime profit: 2014 revision. AIP Res. Rep. NM$5 (10-14). http://aipl.arsusda.gov/reference/nmcalc-2014.htm.

  VanRaden, P.M., M.E. Tooker, J.R. Wright, J.B. Cole, D.J. Null, and T.J. Lawlor. 2014. Genetic base changes for December 2014. AIP Res. Rep. Base3 (10-14). http://aipl.arsusda.gov/reference/base2014.htm.

- **Development of genetic and genomic evaluations for dairy cow livability.** About 17% of cows die instead of being sold, with death losses averaging 6% per lactation; producer income is about $1,200 less for cows that die than those sold for beef. A new trait "cow livability" measures the genetic ability of a cow to stay alive while on the farm, whereas the previous trait "productive life" measured a cow's ability to avoid either dying on the farm or being culled. Selection tools for cow livability were developed to improve cow health, welfare, and profitability with no additional cost for data collection; the national database includes reports of death loss for 69,710,392 lactations of 25,514,760 cows since 1970. Cow livability could receive 7% emphasis in national selection indexes for net economic merit. Genetic and genomic evaluations for cow livability were implemented by the Council on Dairy Cattle Breeding and released officially to the dairy industry in August 2016, and livability could be included in the national indexes in 2017.

  Publications:

  Norman, H.D., J. Wright, and P. VanRaden. 2016. Genetic evaluation for cow livability. Council on Dairy Cattle Breeding, What's New, Aug. 12. https://www.cdcb.us/News/GENETIC%20EVALUATION%20FOR%20COW%20LIVABILIITY.pdf.

  VanRaden, P.M., J.R. Wright, M.E. Tooker, and H.D. Norman. 2016. Value of selecting for cow and calf livability. Interbull Bull. (In press)

  Wright, J.R., and P.M. VanRaden. 2016. Genetic evaluation of dairy cow livability. J. Dairy Sci. 99(E–Suppl. 1)):174. (Abstr.)

- **Determination of phenotypic effects of lethal recessives.** Mutations that cause embryo loss, stillbirth, or calf death when homozygous (having the same allele from both parents for a gene) are fairly easy to discover and track using genomic methods developed in recent years by ARS researchers in Beltsville, Maryland. However, the effects of heterozygosity (having two different alleles for a gene) for such mutations on animal performance and health had not been studied. In 2016, the first large study that examined phenotypic effects on cows that are heterozygous for many different recessive defects was completed. This information can help to predict the future frequency of the defects and to explain why the defects became frequent in the past. Most recessive lethal defects had very little effect on other traits when heterozygous. Selection to reduce defect frequencies by the dairy industry will be independent of progress for other traits of economic importance.
  Publications:
  Cole, J.B., D.J. Null, and P.M. VanRaden, P.M. 2016. Phenotypic and genetic effects of recessive haplotypes on yield, longevity, and fertility. J. Dairy Sci. 99:7274–7288.

Proposal Objective 3: Improve efficiency of genomic prediction and computation by developing faster algorithms, testing new adjustments and models, and accounting for genomic pre-selection in evaluation.

- **Development of a genomic mating program for dairy cattle.** Breed associations, AI organizations, and on-farm software providers needed new computerized mating programs for genomic selection so that genomic inbreeding could be minimized by comparing genotypes of potential mates. Effective methods for transferring genomic relationships from a central database to customers were developed, and more efficient formats were introduced to reduce file sizes. Methods also were developed and tested to consider dominant effects of individual markers when assigning mates to improve the merit of offspring further. Mating programs that included genomic relationships were more effective than those using pedigree relationships because they improved expected value of offspring as well as decreased expected offspring inbreeding. The expected decrease in inbreeding in 2013 was worth over $3 million annually for U.S. Holsteins, and that economic value has increased as more cows were genotyped.
  Publications:
  Sun, C., P.M. VanRaden, J.R. O'Connell, K.A. Weigel, and D. Gianola. 2013. Mating programs including genomic relationships and dominance effects. J. Dairy Sci. 96:8014–8023.

- **Weekly national genomic evaluation of dairy cattle.** Dairy producers and genotyping laboratories both wanted to reduce the time between collecting DNA samples and receiving genomic evaluations, which were calculated monthly. Methods were developed to calculate preliminary genomic evaluations daily or weekly before the release of official monthly evaluations by processing only newly genotyped animals using estimated marker effects from the previous official evaluation. The Council on Dairy Cattle Breeding implemented this method in November 2014 for national genomic evaluations. Additional methods were developed that allow weekly computation of individual reliabilities and inbreeding statistics, and CDCB implemented these in February 2016. Further research will reduce the time needed for monthly evaluations by storing the imputed genotypes and adding new animals from the weekly system. The CDCB provides the evaluations to nominators, dairy records processing centers, and breed associations to facilitate transfer to owners of more than 1.5 million animals in 50 countries. Earlier access to genomic evaluations benefits producers by enabling earlier sale or culling of animals (or embryos) not needed for breeding purposes to minimize the expense and environmental impact of raising newborn calves.

Publications:

Wiggans, G.R., P.M. VanRaden, and T.A. Cooper. 2015. *Technical note:* Rapid calculation of genomic evaluations for new animals. J. Dairy Sci. 98:2039–2042.

- **Improved software for more robust genetic evaluation of dairy cattle.** National genetic evaluations for milk, fat, protein, somatic cell score, productive life, and daughter pregnancy rate were being computed with software that had been developed in 1989. New software was developed that has many additional features and allows a wider variety of statistical models to be compared quickly. The new software was implemented by the Council on Dairy Cattle Breeding in December 2014 for evaluation of yield and health traits, and the genetic evaluation for daughter pregnancy rate was revised to match more closely the methods implemented in 2013 for conception rate traits. With the new software, inbreeding and heterosis effects for each trait are now updated automatically whenever new records are added. Several research projects already are using the revised software to test new models with factors such as interactions between genotype and environment, which was not possible with the previous software.

  Publications:

  VanRaden, P.M., M.E. Tooker, J.R. Wright, C. Sun, and J.L. Hutchison. 2014. Comparison of single-trait to multi-trait national evaluations for yield, health, and fertility. J. Dairy Sci. 97:7952–7962.

  Wright, J.R., and P.M. VanRaden. 2015. Genetic interactions for heat stress and herd yield level: Predicting foreign genetic merit from domestic data. J. Dairy Sci. 98(Suppl. 2):350. (Abstr.)

- **Fast imputation of DNA sequence variants.** Individual DNA sequences help to accelerate understanding of biology and apply to many practical fields; however, deeply sequencing a large number of individuals is still not affordable. An alternative strategy is to use imputation. Some individuals are sequenced and others are genotyped using high-density single-nucleotide polymorphism arrays. Then whole-genome sequences can be imputed (predicted) from array data using computer programs with accurate and efficient algorithms. A new algorithm was developed and implemented in the Findhap software program that finds haplotypes and imputes genotypes from multiple marker sets; the new algorithm was tested using simulated sequences for 10,000 bulls and actual sequences for 1,000 humans. Compared with Beagle, a popular imputation software, the accuracy of the enhanced Findhap software was much better with low-coverage DNA sequences and slightly less with high coverage; however, processing time was about 400 times faster than with Beagle. This allows sequence data to be combined with the vast database of array genotypes. More efficient processing and imputation of sequence data for many individuals in dairy cattle and in other species should increase the rate of genetic improvement when selection is based on genomic evaluations. The new software is available to the public through the ARS web site and is being used internationally.

  Publications:

  VanRaden, P.M. 2016. Findhap.f90. Find haplotypes and impute genotypes using multiple chip sets and sequence data. http://aipl.arsusda.gov/software/findhap/.

  VanRaden, P.M., C. Sun, and J.R. O'Connell. 2015. Fast imputation using medium or low-coverage sequence data. BMC Genetics 16:82.

- **Improved software to align DNA sequence data to a reference genome.** Many computer programs for processing DNA sequence data are available from researchers in human genetics, but those programs often require more computing resources than are available to agricultural researchers. Computation is becoming a more limiting factor each year as sequence data sets grow. Findmap was developed to align short DNA segments to a reference genome (representing the species' standard set of genes), identify mutations, and determine alleles, about 30 times faster than most previous algorithms. Accuracy was improved by storing all known differences (variants) in memory so that alignment could compare to DNA of all previously sequenced animals and not just the reference animal. Speed was improved by allowing multiple processors to share the same memory and by a fast hashing algorithm. The new methods and programs in the Findmap software were documented and released in January 2016 for use by scientific researchers. Availability of more efficient software will allow agricultural researchers to identify mutations at a faster rate and will lead to more accurate genomic selection for livestock species.
  Publications:
  VanRaden, P.M. 2016. Findmap.f90. Align sequence reads to reference map, call previous variants, and identify new variants. http://aipl.arsusda.gov/software/findmap/.
  VanRaden, P.M., and D.M. Bickhart. 2016. Fast single-pass alignment and variant calling using sequencing data. Plant Anim. Genome XXIV Conf., Abstr. W161. https://pag.confex.com/pag/xxiv/meetingapp.cgi/Paper/19292.
  VanRaden, P.M., D.M. Bickhart, and J.R. O'Connell. 2016. Identifying and calling insertions, deletions, and single-base mutations efficiently from sequence data. J. Dairy Sci. 99(E–Suppl. 1):140. (Abstr.)

- **Expanded national genomic evaluation service for dairy cattle.** The rapidly growing service activities related to genetic and genomic evaluation of U.S. dairy cattle, control of the national database, and responsibility for routine delivery of evaluations was transferred from ARS in Beltsville, Maryland, to the Council on Dairy Cattle Breeding (CDCB) in Bowie, Maryland due to insufficient ARS funding; CDCB had no employees in 2012 but has 12 employees in 2016, who work closely with ARS researchers on a daily basis. A copy of the CDCB database is maintained at ARS to allow expanded research on evaluation development and methodology, and CDCB routinely executes hundreds of ARS computer programs with hundreds of thousands of lines of computer code to update the database and provide evaluations. This new arrangement allows CDCB to continue expanding data collection and service to the dairy industry while ARS staff focus on research. Projects completed jointly by ARS and CDCB include exchange of Holstein bull genotypes with Switzerland (March 2016), Japan (May 2016), and Germany (August 2016), genomic prediction of breed composition (June 2016), and expansion of genomic evaluations to Guernseys as a fifth dairy cattle breed (April 2016). The collaboration between ARS and the dairy industry has resulted in a world-leading genomic prediction system and vast database that producers in about 50 countries now use routinely to produce healthier, more productive dairy cattle.
  Publications:
  Wiggans, G.R. 2012. Background on the development of a nonfunded cooperative agreement between USDA's Agricultural Research Service (ARS) and the Council on Dairy Cattle Breeding (CDCB). AIPL. Res. Rep. NFCA-CDCB1 (06-12). http://aipl.arsusda.gov/publish/other/2012/NFCA-CDCB1.pdf.
  Wiggans, G.R., T.A. Cooper, P.M. VanRaden, D.J. Null, J.L. Hutchison, O.M. Meland, and H.D. Norman. 2014. Calculation and delivery of US genomic evaluations of dairy cattle. J. Dairy Sci. 97(E-Suppl. 1):77–78. (Abstr.)

Literature Cited

Berry, D.P., M.P. Coffey, J.E. Pryce, Y. de Haas, P. Løvendahl, N. Krattenmacher, J.J. Crowley, Z. Wang, D. Spurlock, K. Weigel, K. Macdonald, and R.F. Veerkamp. 2014a. International genetic evaluations for feed intake in dairy cattle through the collation of data from multiple sources. J. Dairy Sci. 97:3894–3905.

Berry, D.P., M.C. McClure, and M.P. Mullen. 2014b. Within- and across-breed imputation of high-density genotypes in dairy and beef cattle from medium- and low-density genotypes. J. Anim. Breed. Genet. 131:165–172.

Bickhart, D.M., J.L. Hutchison, D.J. Null, P.M. VanRaden, and J.B. Cole. 2016. Reducing animal sequencing redundancy by preferentially selecting animals with low-frequency haplotypes. J. Dairy Sci. 99:5526–5534.

Bickhart, D.M., J.L.Hutchison, L. Xu, R.D. Schnabel, J.F. Taylor, J.M. Reecy, S.G. Schroeder, C.P. Van Tassell, T.S. Sonstegard, and G. Liu. 2015. G. RAPTR-SV: a hybrid method for the detection of structural variants. Bioinformatics. 31:2084–2090.

Bickhart, D.M., and G.E. Liu. 2014. The challenges and importance of structural variation detection in livestock. Front Genet. 5:37.

Bouwman, A.C., and R. Veerkamp. 2014. Consequences of splitting whole-genome sequencing effort over multiple breeds on imputation accuracy. BMC Genetics 15:105.

Brøndum, R.F., B. Guldbrandtsen, G. Sahana, M.S. Lund, and G. Su. 2014. Strategies for imputation to whole genome sequence using a single or multi-breed reference population in cattle. BMC Genomics 15:728.

Brøndum, R.F., G. Su, L. Janss, G. Sahana, B. Guldbrandtsen, D. Boichard, and M.S. Lund. 2015. Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. J. Dairy Sci. 98:4107–4116.

Calus, M.P.L., J. Vandenplas, and J. Ten Napel. 2015. Ever-growing data sets pose (new) challenges to genomic prediction models. J. Anim. Breed. Genet. 132:407–408.

Calus, M.P.L., and R.F. Veerkamp. 2011. Accuracy of multi-trait genomic selection using different methods. Genet. Sel. Evol. 43:26.

Cameron, N.D. 1997. Selection indices and prediction of genetic merit in animal breeding. CAB International, Wallingford, Oxon, UK.

Charlier, C., W. Li, C. Harland, M. Littlejohn, W. Coppieters, F. Creagh, S. Davis, T. Druet, P. Faux, F. Guillaume, L. Karim, M. Keehan, N.K. Kadri, N. Tamma, R. Spelman, and M. Georges. 2016. NGS-based reverse genetic screen for common embryonic lethal mutations compromising fertility in livestock. Genome Res. 26:1333–1341.

Cole, J.B. 2015. A simple strategy for managing many recessive disorders in a dairy cattle breeding program. Genet. Sel. Evol. 47:94.

Cole, J.B., and D.J. Null. 2010. Age at first calving in Holstein cattle in the United States. J. Dairy Sci. 93(E-Suppl. 1):594. (Abstr.)

Cole, J.B., D.J. Null, C. Sun, and P.M. VanRaden. 2015. Assignment of polled status using single nucleotide polymorphism genotypes and predicted gene content. J. Dairy Sci. 98(Suppl. 2):800. (Abstr.)

Cole, J.B., and VanRaden, P.M. 2011. Use of haplotypes to estimate Mendelian sampling effects and selection limits. J. Anim. Breed. Genet. 128:448–445.

Connor, E.E., J.L. Hutchison, H.D. Norman, K.M. Olson, C.P. Van Tassell, J.M. Leith, and R.L. Baldwin VI. 2013. Use of residual feed intake in Holsteins during early lactation shows potential to improve feed efficiency through genetic selection. J. Anim. Sci. 91:3978–3988.

Connor, E.E., J.L. Hutchison, K.M. Olson, and H.D. Norman. 2012. Triennial Lactation Symposium: Opportunities for improving milk production efficiency in dairy cattle. J. Anim. Sci. 90:1687–1694.

Cooper, T.A., G.R. Wiggans, and P.M. VanRaden. 2015. *Short communication:* Analysis of genomic predictor population for Holstein dairy cattle in the US—Effects of sex and age. J. Dairy Sci. 98:2785–2788.

Daetwyler, H.D., A. Capitan, H. Pausch, P. Stothard, R. van Binsbergen, R.F. Brøndum, X. Liao, A. Djari, S.C. Rodriguez, C. Grohs, D. Esquerré, O. Bouchez, M.-N. Rossignol, C. Klopp, D. Rocha, S. Fritz, A. Eggen, P.J. Bowman, D. Coote, A.J. Chamberlain, C. Anderson, C.P. Van Tassell, I. Hulsegge, M.E. Goddard, B. Guldbrandtsen, M.S. Lund, R.F. Veerkamp, D.A. Boichard, R. Fries, and B.J. Hayes. 2014. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nat. Genet. 46:858–865.

Davis, S.R., K.A. Macdonald, G.C. Waghorn, and R.J. Spelman. 2014. Residual feed intake of lactating Holstein-Friesian cows predicted from high-density genotypes and phenotyping of growing heifers. J. Dairy Sci. 97:1436–1445.

de Haas, Y., M.P.L. Calus, R.F. Veerkamp, E. Wall, M.P. Coffey, H.D. Daetwyler, B.J. Hayes, and J.E. Pryce. 2014. Improved accuracy of genomic prediction for dry matter intake of dairy cattle from combined European and Australian data sets. J. Dairy Sci. 95:6103–6112.

de Haas, Y., J.E. Pryce, M.P.L. Calus, E. Wall, D.P. Berry, P. Løvendahl, N. Krattenmacher, F. Miglior, K. Weigel, D. Spurlock, K.A. Macdonald, B. Hulsegge, and R.F. Veerkamp. 2015. Genomic prediction of dry matter intake in dairy cattle from an international data set consisting of research herds in Europe, North America, and Australasia. J. Dairy Sci. 98:6522–6534.

De Donato, M., S.O. Peters, S.E. Mitchell, T. Hussain, and I.G. Imumorin. 2013. Genotyping-by-sequencing (GBS): A novel, efficient and cost-effective genotyping method for cattle using next-generation sequencing. PLoS ONE. 8:e62137.

Dhakal K., C. Maltecca, J.P. Cassady, G. Baloche, C.M. Williams, and S.P. Washburn. 2013. Calf birth weight, gestation length, calving ease, and neonatal calf mortality in Holstein, Jersey, and crossbred cows in a pasture system. J Dairy Sci. 96:690–698.

Dhakal, K., F. Tiezzi, J.S. Clay, and C. Maltecca. 2015. Short communication: Genomic selection for hoof lesions in first-parity US Holsteins. J. Dairy Sci. 98:3502–3507.

Dhurandhar, E.J., A.I. Vazquez, G.A. Argyropoulos, and D.B. Allison. 2015. Even modest prediction accuracy of genomic models can have large clinical utility. Front. Genetics. 5:417.

Eaglen, S.A.E., M.P. Coffey, J.A. Woolliams, and E. Wall. 2013. Direct and maternal genetic relationships between calving ease, gestation length, milk production, fertility, type, and lifespan of Holstein-Friesian primiparous cows. J. Dairy Sci. 96:4015–4025.

Egger-Danner, C., J.B. Cole, J. Pryce, N. Gengler, B. Heringstad, A. Bradley, L. Andrews, and K.F. Stock. 2014. *Invited review:* Overview of new traits and phenotyping strategies in dairy cattle with a focus on functional traits. Animal 9:191–207.

Ellen, E.D., T.B. Rodenburg, G.A.A. Albers, J.E. Bolhuis, I. Camerlink, N. Duijvesteijn, E.F. Knol, W.M. Muir, K. Peeters, I. Reimert, E. Sell-Kubiak, J.A.M. Van Arendonk, J. Visscher, and P. Bijma. 2014. The prospects of selection for social genetic effects to improve welfare and productivity in livestock. Frontiers Genet. 5:377.

Fernando, R.L., J.C.M. Dekkers, and D.J. Garrick. 2014. A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. Genet. Sel. Evol. 46:50.

Fritz, S., A. Capitan, A. Djari, S.C. Rodriguez, A. Barbat, A. Baur, C. Grohs, B. Weiss, M. Boussaha, D. Esquerré, C. Klopp, D. Rocha, and D. Boichard. 2013. Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2. PLoS ONE 8:e65550.

Fuerst-Waltl, B., and M.K. Sorensen. 2010. Genetic analysis of calf and heifer losses in Danish Holstein. Interbull Bull. 40:117–122.

Gao, Z., D. Waggoner, M. Stephens, C. Ober, and M. Przeworski. 2015. An estimate of the average number of recessive lethal mutations carried by humans. Genetics 199:1243–1254.

García-Ruiz, A., J.B. Cole, P.M. VanRaden, G.R. Wiggans, F.J. Ruiz-López, and C.P. Van Tassell. 2016. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. Proc. Natl. Acad. Sci. USA 113:E3995–E4004.

Gengler, N., S. Abras, C. Verkenne, S. Vanderick, M. Szydlowski, and R. Renaville. 2008. Accuracy of prediction of gene content in large animal populations and its use for candidate gene detection and genetic evaluation. J. Dairy Sci. 91:1652–1659.

Gianola, D. 2013. Priors in whole-genome regression: The Bayesian alphabet returns. Genetics 194:573–596.

Gonzalez-Recio, O., M.P. Coffey, and J.E. Pryce. 2014a. On the value of the phenotypes in the genomic era. J. Dairy Sci. 97:7905–7915.

Gonzalez-Recio, O., J.E. Pryce, M. Haile-Mariam, and B.J. Hayes. 2014b. Incorporating heifer feed efficiency in the Australian selection index using genomic selection. J. Dairy Sci. 97:3883–3893.

Gorjanc, G., M.A. Cleveland, R.D. Houston, and J.M. Hickey. 2015. Potential of genotyping-by-sequencing for genomic selection in livestock populations. Genet. Sel. Evol. 47:12.

Hansen, M., P. Madsen, J. Jensen, J. Pedersen, and L.G. Christensen. 2003. Genetic parameters of postnatal mortality in Danish Holstein calves. J. Dairy Sci. 86:1807–1817.

Harris, B.L., and D.L. Johnson. 2010. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. J. Dairy Sci. 93:1243–1252.

Hayes, B.J., I.M. MacLeod, H.D. Daetwyler, P.J. Bowman, A.J. Chamberlian, C.J. Vander Jagt, A. Capitan, H. Pausch, P. Stothard, X. Liao, C. Schrooten , E. Mullaart, R. Fries, B. Guldbrandtsen, M.S. Lund, D.A. Boichard, R.F. Veerkamp, C.P. VanTassell, B. Gredler, T. Druet, A. Bagnato, J. Vilkki, D.J. deKoning, E. Santus, and M.E. Goddard. 2014. Genomic prediction from whole genome sequence in livestock: The 1000 Bull Genomes Project. Proc. 10th World Congr. Genet. Appl. Livest. Prod., Commun. 183.

Henderson, L., F. Miglior, A. Sewalem, D. Kelton, A. Robinson, and K. Leslie. 2011. Estimation of genetic parameters for measures of calf survival in a population of dairy calves in New York State. J. Dairy Sci. 91:461–470.

Hickey, J.M. 2013. Sequencing millions of animals for genomic selection 2.0. J. Anim. Breed. Genet. 130:331–332.

Hou, Y., D.M. Bickhart, M.L. Hvinden, C. Li, J. Song, D.A. Boichard, S. Fritz, A. Eggen, S. DeNise, G.R. Wiggans, T.S. Sonstegard, C.P. Van Tassell, and G.E. Liu. 2012. Fine mapping of copy number variations on two cattle genomes assembles using high density SNP array. BMC Genomics 13:376.

Hou, Y., G.E. Liu, D.M. Bickhart, M. Cardone, K. Wang, E.-S. Kim, L.K. Matukumalli, M. Ventura, J. Song, P.M. VanRaden, T.S. Sonstegard, and C.P. Van Tassell. 2011. Genomic characteristics of cattle copy number variations. BMC Genomics 12:127.

Hozé, C., S. Fritz, F. Phocas, D. Boichard, V. Ducrocq, and P. Croiseau. 2014. Efficiency of multi-breed genomic selection for dairy cattle breeds with different sizes of reference population. J. Dairy Sci. 97:3918–3929.

Jamrozik, J., A. Koeck, G.J. Kistemaker, and F. Miglior. 2016. Multiple-trait estimates of genetic parameters for metabolic disease traits, fertility disorders, and their predictors in Canadian Holsteins. J. Dairy Sci. 99:1990–1998.

Jenkins, G.M., P. Amer, K. Stachowicz, and S. Meier. 2015. Phenotypic associations between gestation length and production, fertility, survival, and calf traits. J. Dairy Sci. 99(1):418–426.

Jenko, J., G. Gorjanc, M.A. Cleveland, R.K. Varshney, C. Whitelaw, J.A. Woolliams, and J.M. Hickey. 2015. Potential of promotion of alleles by genome editing to improve quantitative traits in livestock breeding programs. Genet. Sel. Evol. 47:55.

Jia, Y., and J.-L. Jannink. 2012. Multiple-trait genomic selection methods increase genetic value prediction accuracy. Genetics 192:1513–1522.

Karoui, S., M.J. Carabaño, C. Díaz, and A. Legarra. 2012. Joint genomic evaluation of French dairy cattle breeds using multiple-trait models. Genet. Sel. Evol. 44:39.

Kemper, K.E., C.M. Reich, P.J. Bowman, C.J. vander Jagt, A.J. Chamberlain, B.A. Mason, B.J. Hayes, and M.E. Goddard. 2015. Improved precision of QTL mapping using a nonlinear Bayesian method in a multi-breed population leads to greater accuracy of across-breed genomic predictions. Genet. Sel. Evol. 47:29.

Kichaev, G., W.-Y. Yang, S. Lindstrom, F. Hormozdiari, E. Eskin, A.L. Price, P. Kraft, and B. Pasaniuc. 2014. Integrating functional data to prioritize causal variants in statistical fine-mapping studies. PLoS Genet. 10:e1004722.

Kipp, S., D. Segelke, S. Schierenbeck, F. Reinhardt, R. Reents, C. Wurmser, H. Pausch, R. Fries, G. Thaller, J. Tetens, J. Pott, D. Haas, B.B. Raddatz, M. Hewicker-Trautwein, I. Proios, M. Schmicke, and W. Grünberg. 2016. Identification of a haplotype associated with cholesterol deficiency and increased juvenile mortality in Holstein cattle. J. Dairy Sci. 99:8915–8931.

Koeck, A., S. Loker, F. Miglior, D.F. Kelton, J. Jamrozik, and F.S. Schenkel. 2014. Genetic relationships of clinical mastitis, cystic ovaries, and lameness with milk yield and somatic cell score in first-lactation Canadian Holsteins. J. Dairy Sci. 97:5806–5813.

Koivula, M., I. Strandén, J. Pösö, G.P. Aamand, and E.A. Mäntysaari. 2015. Single-step genomic evaluation using multitrait random regression model and test-day data. J. Dairy Sci. 98:2775–2784.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.

Lourenco, D.A.L., I. Misztal, S. Tsuruta, I. Aguilar, T.J. Lawlor, S. Forni, and J.I. Weller. 2014. Are evaluations on young genotyped animals benefiting from the past generations? J. Dairy Sci. 97:3930–3942.

Lu, Y., M.J. Vandehaar, D.M. Spurlock, K.A. Weigel, L.E. Armentano, C.R. Staples, E.E. Connor, Z. Wang, N.M. Bello, and R.J. Tempelman. 2015. An alternative approach to modeling genetic merit of feed efficiency in dairy cattle. J. Dairy Sci. 98:6535–6551.

Lund, M. S., G. Su, L. Janss, B. Guldbrandtsen, and R. F. Brøndum, 2014. *Invited review:* Genomic evaluation of cattle in a multi-breed context. Livest. Sci. 166:101–110.

Ma, L., J.R. O'Connell, P.M. VanRaden, B. Shen, A. Padhi, C. Sun, D.M. Bickhart, J.B. Cole, D.J. Null, G.E. Liu, Y. Da, and G.R. Wiggans. 2015a. Cattle sex-specific recombination and genetic control from a large pedigree analysis. PLoS Genet. 11:e1005387.

Ma, P., M.S. Lund, U.S. Nielsen, G.P. Aamand, and G. Su. 2015b. Single-step genomic model improved reliability and reduced the bias of genomic predictions in Danish Jersey. J. Dairy Sci. 98:9026–9034.

MacLeod, I.M., P.J. Bowman, C.J. Vander Jagt, M. Haile-Mariam, K.E. Kemper, A.J. Chamberlain, C. Schrooten, B.J. Hayes, and M.E. Goddard. 2016. Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. BMC Genomics 17:144.

Makgahlela, M.L., E.A. Mäntysaari, I. Strandén, M. Koivula, U.S. Nielsen, M.J. Sillanpaa, and J. Juga. 2013. Across breed multi-trait random regression genomic predictions in the Nordic Red dairy cattle. J. Anim. Breed. Genet. 130:10–19.

Maltecca, C., K.A. Gray, K.A. Weigel, J.P. Cassady, and M. Ashwell. 2011. A genome-wide association study of direct gestation length in US Holstein and Italian Brown populations. Anim. Genet. 42:585–591.

Masuda, Y., I. Misztal, S. Tsuruta, A. Legarra, I. Aguilar, D.A.L. Lourenco, B.O. Fragomeni, and T.J. Lawlor. 2016a. Implementation of genomic recursions in single-step genomic best linear unbiased predictor for US Holsteins with a large number of genotyped animals. J. Dairy Sci. 99:1968–1974.

Masuda, Y., I. Misztal, and P.M. VanRaden. 2016b. Single-step GBLUP using APY inverse for protein yield in U.S. Holstein with a large number of genotyped animals. J. Dairy Sci. 99(E-Suppl. 1):141. (Abstr.)

Matilainen, K., M. Koivula, I. Strandén, G.P. Aamand, and E.A. Mäntysaari. 2016. Managing genetic groups in single-step genomic evaluations applied on female fertility traits in Nordic red dairy cattle. Interbull Bull. 50:71–75.

Matukumalli, L.K., C.T. Lawley, R.D. Schnabel, J.F. Taylor, M.F. Allan, M.P. Heaton, J. O'Connell, S.S. Moore, T.P. Smith, T.S. Sonstegard, and C.P. Van Tassell. 2009. Development and characterization of a high density SNP genotyping assay for cattle. PLoS ONE 4:e5350-10.

McClure, M.C., D. Bickhart, D. Null, P. VanRaden, L. Xu, G. Wiggans, G. Liu, S. Schroeder, J. Glasscock, J. Armstrong, J.B. Cole, C.P. Van Tassell, and T.S. Sonstegard. 2014. Bovine exome sequence analysis and targeted SNP genotyping of recessive fertility defects BH1, HH2, and HH3 reveal a putative causative mutation in SMC2 for HH3. PLoS ONE 9:e92769.

McClure, M., E. Kim, D. Bickhart, D. Null, T. Cooper, J. Cole, G. Wiggans, P. Ajmone-Marsan, L. Colli, E. Santus, G.E. Liu, S. Schroeder, L. Matukumalli, C. Van Tassell, and T. Sonstegard. 2013a. Fine mapping for Weaver Syndrome in Brown Swiss cattle and the identification of 41 concordant mutations across NRCAM, PNPLA8 and CTTNBP2. PLoS ONE 8:e59251.

McClure, M.C., T.S. Sonstegard, G.R. Wiggans, A.L. Van Eenennaam, K.L. Weber, C.T. Penedo, D.P. Berry, J. Flynn, J.F. Garcia, A.S. Carmo, L.C.A. Regitano, M. Albuquerque, M.V.G.B. Silva, M.A. Machado, M. Coffey, K., Moore, M.-Y. Boscher, L. Genestout, R. Mazza, J.F. Taylor, R.D. Schnabel, B. Simpson, E. Marques, J.C. McEwan, A. Cromie, L.L. Coutinho, L.A. Kuehn, J.W. Keele, E.K. Piper, J. Cook, R. Williams, Bovine HapMap Consortium, and C.P. Van Tassell. 2013b. Imputation of microsatellite alleles from dense SNP genotypes for parentage verification across multiple *Bos taurus* and *Bos indicus* breeds. Front. Genet. 4:176.

McDaneld, T.G., L.A. Kuehn, M.G. Thomas, W.M. Snelling, T.P.L. Smith, E.J. Pollak, J.B. Cole, and J.W. Keele. 2014. Genomewide association study of reproductive efficiency in female cattle. J. Anim. Sci. 92:1945–1957.

Misztal, I. 2016. Inexpensive computation of the inverse of the genomic relationship matrix in populations with small effective population size. Genetics 202:401–409.

Misztal, I., Z.G. Vitezica, A. Legarra, I. Aguilar, and A.A. Swan. 2013. Unknown-parent groups in single-step genomic evaluation. J. Anim. Breed. Genet. 130:252–258.

Norman, H.D., J.R. Wright, J.L. Hutchison, and J.M. Mattison. 2014. Selection changes in the United States due to genomics. Interbull Bull. 48:37–39.

Norman, H.D., J.R. Wright, M.T. Kuhn, S.M. Hubbard, J.B. Cole, and P.M. VanRaden. 2009. Genetic and environmental factors that affect gestation length in dairy cattle. J. Dairy Sci. 92:2259–2269.

Norman, H.D., J.R. Wright, and R.H. Miller. 2011. Potential consequences of selection to change gestation length on performance of Holstein cows. J. Dairy Sci. 94:1005–1010.

O'Connell, J.R., M.E. Tooker, D.M. Bickhart, and P.M. VanRaden. 2016. Selection of sequence variants to improve genomic predictions. Interbull Bull. 50:58–66.

Olson, K.M., P.M. VanRaden, and M.E. Tooker. 2012. Multibreed genomic evaluations using purebred Holsteins, Jerseys, and Brown Swiss. J. Dairy Sci. 95:5378–5383.

Olson, K.M., P.M. VanRaden, M.E. Tooker, and T.A. Cooper. 2011. Differences among methods to validate genomic evaluations for dairy cattle. J. Dairy Sci. 94:2613–2620.

Parker Gaddis, K.L., J.B. Cole, J.S. Clay, and C. Maltecca. 2012. Incidence validation and relationship analysis of producer-recorded health event data from on-farm computer systems in the United States. J. Dairy Sci. 95:5422–5435.

Parker Gaddis, K.L., Cole, J.B., Clay, J.S., and Maltecca, C. Genomic selection for producer-recorded health event data in US dairy cattle. J. Dairy Sci. 97(5):3190–3199. 2014.
Parker Gaddis, K.L., J.B. Cole, J.S. Clay, and C. Maltecca. 2016. Benchmarking dairy herd health status using routinely recorded herd summary data. J. Dairy Sci. 99:1298–1314.
Patry, C., and V. Ducrocq. 2011. Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle. Genet. Sel. Evol. 43:30.
Pausch, H., C. Wurmser, F. Reinhardt, R. Emmerling, and R. Fries. 2015. *Short communication:* Validation of 4 candidate causative trait variants in 2 cattle breeds using targeted sequence imputation. J. Dairy Sci. 98:4162–4167.
Pérez-Enciso, M., J.C. Rincón, and A. Legarra. 2015. Sequence- vs. chip-assisted genomic selection: Accurate biological information is advised. Genet. Sel. Evol. 47:43.
Přibyl, J., J. Bauer, V. Čermák, P. Pešek, J. Přibylová, J. Šplíchal, H. Vostrá-Vydrová, L. Vostrý, and L. Zavadilová. 2015. Domestic estimated breeding values and genomic enhanced breeding values of bulls in comparison with their foreign genomic enhanced breeding values. Animal 9:1635–1642.
Pryce, J.E., O. Gonzalez-Recio, G. Nieuwhof, W.J. Wales, M.P. Coffey, B.J. Hayes, and M.E. Goddard. 2015. *Hot topic:* Definition and implementation of a breeding value for feed efficiency in dairy cows. J. Dairy Sci. 98:7340–7350.
Sargolzaei, M., J. Chesnais, and F. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15:478.
Shook, G.E. 2006. Major advances in determining appropriate selection goals. J. Dairy Sci. 89:1349–1361.
Stouthard, P., X. Liao, A.S. Arantes, M. De Pauw, C. Coros, G.S. Plastow, M. Sargolzaei, J.J. Crowley, J.A. Basarab, F. Schenkel, S. Moore, and S.P. Miller. 2015. A large and diverse collection of bovine genome sequences from the Canadian Cattle Genome Project. GigaScience 4:49.
Sun, C., and P.M. VanRaden. 2014. Increasing long-term response by selecting for favorable minor alleles. PLoS ONE 9:e88510.
Tsuruta, S., I. Misztal, D.A.L. Lourenco, and T.J. Lawlor. 2014. Assigning unknown parent groups to reduce bias in genomic evaluations of final score in US Holsteins. J. Dairy Sci. 97:5814–5821.
United States Supreme Court. 2013. Ass'n for Molecular Pathology v. Myriad Genetics. http://www.supremecourt.gov/opinions/12pdf/12-398_1b7d.pdf.
Vallimont, J.E., C.D. Dechow, J.M. Daubert, M.W. Dekleva, J.W. Blum, W. Liu, G.A. Varga, A.J. Heinrichs, and C.R. Baumrucker. 2012. *Short communication:* Feed utilization and its associations with fertility and productive life in 11 commercial Pennsylvania tie-stall herds. J. Dairy Sci. 96:1251–1254.
van Binsbergen, R., M.C. Bink, M.P. Calus, F.A. van Eeuwijk, B.J. Hayes, I. Hulsegge, and R.F. Veerkamp. 2014a. Accuracy of imputation to whole-genome sequence data in Holstein Friesian cattle. Genet. Sel. Evol. 46:41.
van Binsbergen, R., M.P.L. Calus, M.C.A.M. Bink, C. Schrooten, F.A. van Eeuwijk, and R.F. Veerkamp. 2014b. Genomic prediction with 12.5 million SNPs for 5503 Holstein Friesian bulls. Proc. 10th World Congr. Genet. Appl. Livest. Prod., Commun. 664.
van den Berg, I., D. Boichard, and M.S. Lund. 2016. Comparing power and precision of within-breed and multibreed genome-wide association studies of production traits using whole-genome sequence data for 5 French and Danish dairy cattle breeds. J. Dairy Sci. 99:8932–8945.
VandeHaar, M.J., L.E. Armentano, K. Weigel, D.M. Spurlock, R.J. Tempelman, and R. Veerkamp. 2016. Harnessing the genetics of the modern dairy cow to continue improvements in feed efficiency. J. Dairy Sci. 99:4941–4954.

Vandenplas, J., F.G. Colinet, G. Glorieux, C. Bertozzi, and N. Gengler. 2015. Integration of external estimated breeding values and associated reliabilities using correlations among traits and effects. J. Dairy Sci. 98:9044–9050.
VanRaden, P.M. 2004. *Invited review:* Selection on net merit to improve lifetime profit. J. Dairy Sci. 87:3125–3131.
VanRaden, P.M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423.
VanRaden, P.M. 2016. Practical implications for genetic modeling in the genomics era. J. Dairy Sci. 99:2405–2412.
VanRaden, P.M., and D.M. Bickhart. 2016. Fast single-pass alignment and variant calling using sequencing data. Plant Anim. Genome XXIV Conf., Abstr. W161.
VanRaden, P.M., and J.B. Cole. 2014. Net merit as a measure of lifetime profit: 2014 revision. AIP Res. Rep. NM$5(10-14). https://aipl.arsusda.gov/reference/nmcalc-2014.htm.
VanRaden, P.M., and T.A. Cooper. 2015. Genomic evaluations and breed composition for crossbred U.S. dairy cattle. Interbull Bull. 49:14–18.
VanRaden, P.M., J.R. Wright, M.E. Tooker, and H.D. Norman. 2016. Value of selecting for cow and calf livability. Interbull Bull. 50:30–33.
VanRaden, P.M., and Miller, R.H. 2008. The USDA Animal Improvement Programs Laboratory: A century old and just getting started. AIPL Res. Rep. HIST1(10-08). https://aipl.arsusda.gov/ Centennial/AIPL%20history_ARR%20HIST1.pdf.
VanRaden, P.M., D.J. Null, M. Sargolzaei, G.R. Wiggans, M.E. Tooker, J.B. Cole, T.S. Sonstegard, E.E. Connor, M. Winters, J.B.C.H.M. van Kaam, A. Valentini, B.J. Van Doormaal, M.A. Faust, and G.A. Doak. 2013. Genomic imputation and evaluation using high density Holstein genotypes. J. Dairy Sci. 96:668–678.
VanRaden, P.M., J.R. O'Connell, G.R. Wiggans, and K.A. Weigel. 2011a. Genomic evaluations with many more genotypes. Genet. Sel. Evol. 43:10.
VanRaden, P.M., K.M. Olson, D.J. Null, and J.L. Hutchison. 2011b. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. J. Dairy Sci. 94:6153–6161.
VanRaden, P.M., and A.H. Sanders. 2003. Economic merit of crossbred and purebred US dairy cattle. J. Dairy Sci. 86:1036–1044.
VanRaden, P.M., M.E. Tooker, J.R. Wright, C. Sun, and J.L. Hutchison. 2014. Comparison of single-trait to multi-trait national evaluations for yield, health, and fertility. J. Dairy Sci. 97:7952–7962.
VanRaden, P.M., and J.R. Wright. 2013. Measuring genomic pre-selection in theory and in practice. Interbull Bull. 47:147–150.
Veerkamp, R.F., M.P.L. Calus, G. de Jong, R. van der Linde, and Y. De Haas. 2014. Breeding value for dry matter intake for Dutch bulls based on DGV for DMI and BV for predictors. Proc. 10th World Congr. Genet. Appl. Livest. Prod., Commun. 554.
Ventura, R.V., D. Lu, F.S. Schenkel, Z. Wang, C. Li, and S.P. Miller. 2014. Impact of reference population on accuracy of imputation from 6K to 50K single nucleotide polymorphism chips in purebred and crossbreed beef cattle. J. Anim. Sci. 92:1433–1444.
Vitezica, Z.G., I. Aguilar, I. Misztal, and A. Legarra. 2011. Bias in genomic predictions for populations under selection. Genet. Res. 93:357–366.
Vukasinovic, N., N. Bacciu, C.A. Przybyla, P. Boddhireddy, and S.K. DeNise. 2017. Development of genetic and genomic evaluation for wellness traits in US Holstein cows. J. Dairy Sci. 100:428–438.
Weber, A., E. Stamer, W. Junge, and G. Thaller. 2013. Genetic parameters for lameness and claw and leg diseases in dairy cows. J. Dairy Sci. 96:3310–3318.
Wientjes, Y.C.J., P. Bijma, R.F. Veerkamp, and M.P.L. Calus. 2016. An equation to predict the accuracy of genomic values by combining data from multiple traits, populations, or environments. Genetics 202:799–823.

Wiggans, G.R., T.A. Cooper, C.P. Van Tassell, T.S. Sonstegard, and E.B. Simpson. 2013. *Technical note:* Characteristics and use of the Illumina BovineLD and GeneSeek Genomic Profiler low-density bead chips for genomic evaluation. J. Dairy Sci. 96:1258–1263.

Wiggans, G.R., T.A. Cooper, P.M. VanRaden, C.P. Van Tassell, D.M. Bickhart, and T.S. Sonstegard. 2016a. Increasing the number of single nucleotide polymorphisms used in genomic evaluation of dairy cattle. J. Dairy Sci. 99:4504–4511.

Wiggans, G.R., P.M. VanRaden, D.M. Bickhart, and M.E. Tooker. 2016b. Strategy for incorporating newly discovered causative genetic variants into genomic evaluations. J. Dairy Sci. 99(E–Suppl. 1):137. (Abstr.)

Wiggans, G.R., P.M. VanRaden, and T.A. Cooper. 2011. The genomic evaluation system in the United States: Past, present, future. J. Dairy Sci. 94:3202–3211.

Winkelman, A.M., D.L. Johnson, and B.L. Harris. 2015. Application of genomic evaluation to dairy cattle in New Zealand. J. Dairy Sci. 998:659–675.

Womack, J.E. 2005. Advances in livestock genomics: Opening the barn door. Genome Res. 15:1699–1705.

Wright, J.R., and P.M. VanRaden. 2015. Genetic interactions for heat stress and herd yield level: Predicting foreign genetic merit from domestic data. J. Dairy Sci. 98(Suppl. 2):350. (Abstr.)

Xu, L., J.B. Cole, D.M. Bickhart, Y. Hou, J. Song, P.M. VanRaden, T.S. Sonstegard, C.P. Van Tassell, and G.E. Liu. 2014. Genome wide CNV analysis reveals additional variants associated with milk production traits in Holsteins. BMC Genomics 15:683.

Xue Y, Y. Chen, Q. Ayub, N. Huang, E.V. Ball, M. Mort, A.D. Phillips, K. Shaw, P.D. Stenson, D.N. Cooper, C. Tyler-Smith, and 1000 Genomes Project Consortium. 2012. Deleterious- and disease-allele prevalence in healthy individuals: Insights from current predictions, mutation databases, and population-scale resequencing. Am. J. Hum. Genet. 91:1022–1032.

Yang, W., and R.J. Tempelman. 2012. A Bayesian antedependence model for whole genome prediction. Genetics 190:1491–1501.

Zhou, L., B. Heringstad, G. Su, B. Guldbrandtsen, T.H.E. Meuwissen, M. Svendsen, H. Grove, U.S. Nielsen, and M.S. Lund. 2014. Genomic predictions based on a joint reference population for the Nordic Red cattle breeds. J. Dairy Sci. 97:4485–4496.

Zhu, Z., F. Zhang, H. Hu, A. Bakshi, M.R. Robinson, J.E. Powell, G.W. Montgomery, M.E. Goddard, N.R Wray, P.M. Visscher, and J. Yang. 2016. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nature Genetics 48:481–487.

Zwald, N.R., K.A. Weigel, Y.M. Chang, R.D. Welper, and J.S. Clay. 2004. Genetic selection for health traits using producer-recorded data. I. Incidence rates, heritability estimates, and sire breeding values. J. Dairy Sci. 87:4287–4294.

Past Accomplishments of Investigators

Past Accomplishments of Paul M. VanRaden

Education

1981 University of Illinois, B.S., dairy science
1984 Iowa State University, M.S., animal breeding
1986 Iowa State University, Ph.D., animal breeding

Experience

1982–86 Graduate Research Assistant, Iowa State University, Ames, IA
1986 Postdoctoral Research Associate, Iowa State University, Ames, IA
1987–88 Postdoctoral Research Associate, University of Wisconsin, Madison, WI
1988–present Research Geneticist (Animal), USDA, ARS, Beltsville, MD

Accomplishments

Dr. VanRaden has improved several important areas of animal breeding such as selection for more traits, avoidance of inbreeding, computation of genomic evaluations, and discovery of several lethal recessive genes. His methods were made directly available to U.S. dairy producers, and computer programs for routine use have made those advances available to other scientists and producers worldwide. He developed the new multitrait evaluation software implemented for U.S. national evaluations in 2014. He cooperated with Dr. John Cole to introduce heifer conception and cow conception rates into national merit indexes in 2014 and developed cow livability as a new trait in 2016. Official genomic evaluations were implemented using Dr. VanRaden's methods and programs for U.S. dairy cattle in January 2009, the first such system in the world. Several other countries quickly adopted those same methods or programs, and the Interbull Centre (Uppsala, Sweden) now uses them for genomic multitrait across-country evaluation (GMACE) of the Holstein breed and the global genomic evaluation of the Brown Swiss breed. Genotypes from many low-density marker panels have been imputed to higher marker densities since 2010 using methods developed by Dr. VanRaden. His imputation methods also proved to be very efficient for predicting higher density genotypes and sequence data in large populations. Dr. VanRaden's methods to split genotypes into haplotypes led to the discovery of five new lethal recessive defects in 2011 and 10 additional haplotypes now reported to breeders for all genotyped animals. Since 2012, Dr. VanRaden has (co)authored 29 scientific journal research papers (senior or sole author of 5) and 26 research abstracts. He has made 12 scientific presentations (including 9 invited papers) at national and international scientific meetings and formal research presentations in 8 countries. He has (co)authored 16 articles in international proceedings and 7 in USDA and popular trade publications. He has made over 21 presentations at international, national, industry, and university meetings. Dr. VanRaden's accomplishments and publications have resulted in national and international awards, including ADSA's Most Cited Award in 2012 and 2016, and Thomson-Reuters Highly Cited Researcher awards in 2014-2016.

Publications: (20 selected)

VanRaden, P.M., and G.R. Wiggans. 1991. Derivation, calculation, and use of national animal model information. J. Dairy Sci. 74:2737–2746.

VanRaden, P.M. 1992. Accounting for inbreeding and crossbreeding in genetic evaluation of large populations. J. Dairy Sci. 75:3136–3144.

VanRaden, P.M., and A.H. Sanders. 2003. Economic merit of crossbred and purebred US dairy cattle. J. Dairy Sci. 86:1036–1044.

VanRaden, P.M., A.H. Sanders, M.E. Tooker, R.H. Miller, H.D. Norman, M.T. Kuhn, and G.R. Wiggans. 2004. Development of a national genetic evaluation for cow fertility. J. Dairy Sci. 87:2285–2292.

VanRaden, P.M. 2004. *Invited review:* Selection on net merit to improve lifetime profit. J. Dairy Sci. 87:3125–3131.

VanRaden, P.M., M.E. Tooker, J.B. Cole, G.R. Wiggans, and J. H. Megonigal Jr. 2007. Genetic evaluations for mixed-breed populations. J. Dairy Sci. 90:2434–2441.

VanRaden, P.M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423.

VanRaden, P.M., C.P. Van Tassell, G.R. Wiggans, T.S. Sonstegard, R.D. Schnabel, J.F. Taylor, and F.S. Schenkel. 2009. *Invited review:* Reliability of genomic predictions for North American Holstein bulls. J. Dairy Sci. 92:16–24.

VanRaden, P.M., and P. Sullivan. 2010. International genomic evaluation methods for dairy cattle. Genet. Sel. Evol. 42:7.

VanRaden, P.M., J.R. O'Connell, G.R. Wiggans, and K.A. Weigel. 2011. Genomic evaluations with many more genotypes. Genet. Sel. Evol. 43:10.

Wiggans, G.R., **P.M. VanRaden**, and T.A. Cooper. 2011. The genomic evaluation system in the United States: Past, present, future. J. Dairy Sci. 94:3202–3211.

VanRaden, P.M., K.M. Olson, G.R. Wiggans, J.B. Cole, and M.E. Tooker. 2011. Genomic inbreeding and relationships among Holsteins, Jerseys, and Brown Swiss. J. Dairy Sci. 94:5673–5680.

VanRaden, P.M., K.M. Olson, D.J. Null, and J.L. Hutchison. 2011. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. J. Dairy Sci. 94:6153–6161.

Olson, K.M., **P.M. VanRaden**, and Tooker, M.E. 2012. Multibreed genomic evaluations using purebred Holsteins, Jerseys, and Brown Swiss. J. Dairy Sci. 95:5378–5383.

VanRaden, P.M., D.J. Null, M. Sargolzaei, G.R. Wiggans, M.E. Tooker, J.B. Cole, T.S. Sonstegard, E.E. Connor, M. Winters, J.B.C.H.M. van Kaam, A. Valentini, B.J. Van Doormaal, M.A. Faust, and G.A. Doak. 2013. Genomic imputation and evaluation using high density Holstein genotypes. J. Dairy Sci. 96:668–678.

VanRaden, P.M., T.A. Cooper, G.R. Wiggans, J.R. O'Connell, and L.R. Bacheller. 2013. Confirmation and discovery of maternal grandsires and great-grandsires in dairy cattle. J. Dairy Sci. 96:1874–1879.

Sun, C., **P.M. VanRaden**, J.R. O'Connell, K.A. Weigel, and D. Gianola. 2013.Mating programs including genomic relationships and dominance effects. J. Dairy Sci. 96:8014–8023.

VanRaden, P.M., M.E. Tooker, J.R. Wright, C. Sun, and J.L. Hutchison. 2014. Comparison of single-trait to multi-trait national evaluations for yield, health, and fertility. J. Dairy Sci. 97:7952–7962.

VanRaden, P.M., C. Sun, and J.R. O'Connell. 2015. Fast imputation using medium or low-coverage sequence data. BMC Genetics 16:82.

VanRaden, P.M. 2016. Practical implications for genetic modeling in the genomics era. J. Dairy Sci. 99:2405–2412.

Past Accomplishments of John B. Cole

Education

1994 Louisiana State University, B.S., animal production systems (dairy)
1996 Louisiana State University, M.S., animal, dairy, and poultry sciences
2003 Louisiana State University, Ph.D., animal and dairy sciences

Experience

1994–96	Graduate Research Assistant, Louisiana State University, Baton Rouge, LA
1996–2000	Graduate Assistant, University of Minnesota, St. Paul, MN
2000–02	Computer Analyst II (Webmaster), Louisiana State University, Baton Rouge, LA
2001	Instructor, Louisiana State University, Baton Rouge, LA
2002–03	Data Manager, Southern Regional Climate Center, Baton Rouge, LA
2003–2016	Research Geneticist (Animal), USDA, ARS, Beltsville, MD
2016–present	Acting Research Leader, USDA, ARS, Beltsville, MD
2010–11	Legislative Fellow, Senator Mark L. Pryor, U.S. Senate, Washington, DC

Accomplishments

Dr. Cole's research program has focused on calving traits, lactation persistency, health traits, and use of climatological data in dairy cattle evaluation. He introduced calving-ease evaluations for Brown Swiss bulls in 2005 and stillbirth evaluations for Holstein bulls in 2006 and worked with Dr. Paul VanRaden on revision of the lifetime net merit index to include those calving traits. He developed the first U.S. evaluation to use data routinely from crossbred animals and was part of the team that developed the U.S. all-breed evaluation in 2007. Dr. Cole estimated variance components for lactation persistency, genetic correlations with yield traits, and breeding values for six U.S. dairy cattle breeds and showed that selection for improved lactation persistency would not adversely affect yield. He also collaborated with university scientists on relationships between persistency and early-lactation metabolic diseases. Dr. Cole developed a database of climate data from around the United States in conjunction with the Hydrology and Remote Sensing Laboratory (Beltsville, MD) and the Southern Regional Climate Center (Baton Rouge, LA) for research on regional climate effects and genotype-by-environment interaction. He also developed the PyPedal software package for pedigree validation and analysis in 2007. In collaboration with regional research project S-1040, Dr. Cole revised the net merit, fluid merit, and cheese merit selection indexes in 2009 to reflect current and predicted future economic conditions. Dr. Cole worked extensively with high-density DNA marker data and published the first estimates of selection limits and Mendelian sampling effects based on haplotypes in 2009. In addition, he developed methods for visualizing high-dimensionality genomics data. Using those data, he identified a QTL associated with dystocia, conformation, longevity, and lifetime economic merit in Holsteins and proposed a physiological explanation for the QTL effect based on comparative bioinformatics with the human and the mouse. In 2011, Dr. Cole developed multiplicative adjustment factors for correcting milk, fat, and protein test-day data to account for effects of region- and season-of-calving, which were added to best prediction programs used to compute lactation yields from test-day data. Those tools are being used by researchers at the University of Florida to study genotype-by-environment interactions affecting yield and fertility. Since 2002, Dr. Cole has (co)authored 139 publications and is senior or sole author of 50 of those, including 23 scientific journal articles, 11 proceeding papers, 3 software packages, 4 popular publications, 5 Laboratory research reports, and 4 book chapters; he has also authored 95 abstracts.

Publications: (20 selected)

Cole, J.B., D.E. Franke, and E.A. Leighton. 2004. Population structure of a colony of dog guides. J. Anim. Sci. 82:2906–2912.

Cole, J.B., R.C. Goodling Jr., G.R. Wiggans, and P.M. VanRaden. 2005. Genetic evaluation of calving ease for Brown Swiss and Jersey bulls from purebred and crossbred calvings. J. Dairy Sci. 88:1529–1539.

Cole, J.B., G.R. Wiggans, and P.M. VanRaden. 2007. Genetic evaluation of stillbirth in United States Holsteins using a sire-maternal grandsire threshold model. J. Dairy Sci. 90:2480–2488.

Cole, J.B., D.J. Null, and P.M. VanRaden. 2009. Best prediction of yields for long lactations. J. Dairy Sci. 92:1796–1810.

Cole, J.B., P.M. VanRaden, J.R. O'Connell, C.P. Van Tassell, T.S. Sonstegard, T.S., R.D. Schnabel, J.F. Taylor, and G.R. Wiggans. 2009. Distribution and location of genetic effects for dairy traits. J. Dairy Sci. 92:2931–2946.

Cole, J.B., and P.M. VanRaden. 2010. Visualization of results from genomic evaluations. J. Dairy Sci. 93:2727–2740.

Cole, J.B., and P.M. VanRaden. 2011. Use of haplotypes to estimate Mendelian sampling effects and selection limits. J. Anim. Breed. Genet. 128:407–497.

Cole, J.B., G.R. Wiggans, L. Ma, T.S. Sonstegard, T.J. Lawlor, Jr., B.A. Crooker, C.P. Van Tassell, J. Yang, S. Wang, L.K. Matukumalli, and Y. Da. 2011. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. BMC Genomics 12:408.

Cochran, S.D., **J.B. Cole**, D.J. Null, and P.J. Hansen. 2013. Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. BMC Genetics 14:49.

Sonstegard, T.S., **J.B. Cole**, P.M. VanRaden, C.P. Van Tassell, D.J. Null, S.G. Schroeder, D. Bickhart, and M.C. McClure. 2013. Identification of a nonsense mutation in CWC15 associated with decreased reproductive efficiency in Jersey cattle. PLoS ONE 8:e54872.

Egger-Danner, C., **J.B. Cole**, J. Pryce, N. Gengler, B. Heringstad, A. Bradley, L. Andrews, and K.F. Stock. 2014. Invited review: overview of new traits and phenotyping strategies in dairy cattle with a focus on functional traits. Animal 9:191–207.

Cole, J.B., B. Waurich, M. Wensch-Dorendorf, D.M. Bickhart, and H.H. Swalve. 2014. A genome-wide association study of calf birth weight in Holstein cattle using single nucleotide polymorphisms and phenotypes predicted from auxiliary traits. J. Dairy Sci. 97:3156–3172.

Parker Gaddis, K.L., **J.B. Cole**, J.S. Clay, and C. Maltecca. 2014. Genomic selection for producer-recorded health event data in U.S. dairy cattle. J. Dairy Sci. 97:3190–3199.

Gay, K.D., N.J. Widmar, T.D. Nennich, A.P. Schinckel, **J.B. Cole**, and M.M. Schutz. 2014. Development of a lifetime merit-based selection index for US dairy grazing systems. J. Dairy Sci. 97:4568–4578.

Parker Gaddis, K.L., F. Tiezzi, **J.B. Cole**, J.S. Clay, and C. Maltecca. 2015. Genomic prediction of disease occurrence using producer-recorded health data: A comparison of methods. Genet. Sel. Evol. 47:41.

Cole, J.B. 2015. A simple strategy for managing many recessive disorders in a dairy cattle breeding program. Genet. Sel. Evol. 47:94.

Ma, L., J.R. O'Connell, P.M. VanRaden, B. Shen, A. Padhi, C. Sun, D.M. Bickhart, **J.B. Cole**, D.J. Null, G. Liu, Y. Da, and G.R. Wiggans. 2015. Cattle sex-specific recombination and genetic control from a large pedigree analysis. PLoS Genet. 11:31005387.

García-Ruiz, A., **J.B. Cole**, P.M. VanRaden, G.R. Wiggans, F.J. Ruiz-López, and C.P. Van Tassell. 2016. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. Proc. Natl. Acad. Sci. USA 113:E3995–E4004.

Pryce, J., K. Parker Gaddis, A. Koeck, C. Bastin, M. Abdelsayed, N. Gengler, F. Miglior, B. Heringstad, C. Egger-Danner, K. Stock, A. Bradley, and **J. Cole**. 2016. *Invited review:* Opportunities for genetic improvement of metabolic diseases. J. Dairy Sci. 99:6855–6873.

Cole, J.B., D.J. Null, and P.M. VanRaden. 2016. Phenotypic and genetic effects of recessive haplotypes on yield, longevity, and fertility. J. Dairy Sci. 99:7274–7288.

Past Accomplishments of Curtis P. Van Tassell

Education

B.S. (Honors and Distinction), Cornell University, Animal Science, 1986
M.S., Iowa State University, Animal Breeding, 1989
Ph.D., Cornell University, Animal Breeding, 1994

Experience

Research Geneticist, USDA-ARS US Meat Animal Research Center, and University of Nebraska, Lincoln, NE, 1994–1996
Research Geneticist, USDA-ARS Animal Improvement Programs Laboratory, Beltsville, MD, 1996–1997
Research Geneticist, USDA-ARS Animal Genomics and Improvement Laboratory, Bovine Functional Genomics Laboratory, and Animal Improvement Programs Laboratory, Beltsville, MD, 1998–present

Accomplishments

Dr. Van Tassell has consistently demonstrated high levels of originality and leadership in his career at ARS. He has contributed crucial theoretical advances in animal breeding, statistics, computational methodology, and bioinformatics. Computer algorithms developed by Dr. Van Tassell enable estimation of variance components and genetic prediction using extremely flexible and complex models. He has completed two of the largest parameter estimation projects of their kind ever conducted. Dr. Van Tassell led development of a methodology for one-step single nucleotide polymorphism (SNP) discovery and validation by next-generation sequencing featured as the cover article of the prominent journal *Nature Methods.* This approach has propelled the abilities of researchers across species to develop SNP resources. His foresight to develop the BovineSNP50 with uniform marker density and robust assay performance proved to be critical not only for scientific success in genome selection and predictions of genetic merit in dairy cattle, but also launched the field of agricultural genomics with over 2 million animals genotyped with this assay, making it the *de facto* standard for bovine genomics. Key to his success has been the ability to lead teams with divergent skills and capacities and shepherd them to a common goal. An example of this leadership can be seen in the coordination of the goat genome assembly, where the team included PacBio, a third-generation sequencing platform, academic groups from University of Washington and Virginia State University, as well as government laboratories. Although there were natural competitors among the team, Dr. Van Tassell was able to mold this group of cutting-edge researchers into a highly effective group and to deliver, arguably, the best mammalian assembly (with the exception of the human and mouse), for under $250,000 – a fraction of previous assembly costs.

Publications: (20 selected)

Van Tassell, C.P., and L.D. Van Vleck. 1991. Estimates of genetic selection differentials and generation intervals for four paths of selection. J. Dairy Sci. 74:1078–1086.
Van Tassell, C.P., and L.D. Van Vleck. 1996. Multiple trait Gibbs sampler for animal models: flexible programs for Bayesian and likelihood-based (co)variance component inference. J. Anim. Sci. 74:2586–2597.
Van Tassell, C.P., L.D. Van Vleck, and K.E. Gregory. 1998. Bayesian analysis of twinning and ovulation rates using a multiple-trait threshold model and Gibbs sampling. J. Anim. Sci. 76:2048–2061.
Van Tassell, C.P., G.R. Wiggans, and H.D. Norman. 1999. Method R estimates of heritability for milk, fat, and protein yields of United States dairy cattle. J. Dairy Sci. 82:2231–2237.

Van Tassell, C.P., I. Misztal, and L. Varona. 2000. Method R estimates of additive genetic, dominance genetic, and permanent environmental fraction of variance for yield and health traits of Holsteins. J. Dairy Sci. 83:1873–1877.

Van Tassell, C.P., G.R. Wiggans, and I. Misztal. 2003.Implementation of a sire-maternal grandsire model for evaluation of calving ease in the United States. J. Dairy Sci. 86:3366–3373.

Van Tassell, C.P., T.S. Sonstegard, and M.S. Ashwell. 2004. Mapping quantitative trait loci affecting dairy conformation to chromosome 27 in two Holstein grandsire families. J. Dairy Sci. 87:450–457.

Matukumalli, L.K., J.J. Grefenstette, D.L. Hyten, I.Y. Choi, P.B. Cregan, and **C.P. Van Tassell**. 2006. Application of machine learning in SNP discovery. BMC. Bioinformatics 7:4.

Matukumalli, L.K., J.J. Grefenstette, D.L. Hyten, I.Y. Choi, P.B. Cregan, and **C.P. Van Tassell**. 2006. SNP-PHAGE--High throughput SNP discovery pipeline. BMC. Bioinformatics 7:468.

Van Tassell, C.P., T.P. Smith, L.K. Matukumalli, J.F. Taylor, R.D. Schnabel, C.T. Lawley, C.D. Haudenschild, S.S. Moore, W.C. Warren, and T.S. Sonstegard. 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. Nat. Methods. 5:247–252.

Bovine Genome Sequencing and Analysis Consortium. 2009. The genome sequence of taurine cattle: A window to ruminant biology and evolution. Science 324:522–528.

Bovine HapMap Consortium (R.A. Gibbs, J.F. Taylor, and **C.P. Van Tassell,** corresponding authors). 2009. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. Science 324:528–532.

Matukumalli, L.K., C.T. Lawley, R.D. Schnabel, J.F. Taylor, M.F. Allan, M.P. Heaton, J. O'Connell, S.S. Moore, T.P. Smith, T.S. Sonstegard, and **C.P. Van Tassell**. 2009. Development and characterization of a high density SNP genotyping assay for cattle. PLoS ONE 4:e5350.

VanRaden, P.M., **C.P. Van Tassell**, G.R. Wiggans, T.S. Sonstegard, R.D. Schnabel, J.F. Taylor, and F.S. Schenkel. 2009. *Invited review:* Reliability of genomic predictions for North American Holstein bulls. J. Dairy Sci. 92:16–24.

McClure, M.C., E. Kim, D. Bickhart, D. Null, T. Cooper, J. Cole, G. Wiggans, P. Ajmone-Marsan, L. Colli, E. Santus, G.E. Liu, S. Schroeder, L. Matukumalli, **C. Van Tassell**, and T. Sonstegard. 2013. Fine mapping for Weaver Syndrome in Brown Swiss cattle and the identification of 41 concordant mutations across NRCAM, PNPLA8 and CTTNBP2. PLoS ONE 8:e59251.

McClure, M.C., Sonstegard, T.S., Wiggans, G.R., Van Eenennaam, A.L., Weber, K.L., Penedo, M.C.T., Berry, D., Flynn, J., Garcia, J.F., Carmo, A.S., Regitano, L.C.A., Albuquerque, M., Silva, M.V.G.B., Machado, M.A., Coffey, M., Moore, K., Boscher, M.Y., Genestout, L., Mazza, R., Taylor, J.F., Schnabel, R.D., Simpson, B., Marques, E., McEwan, J., Cromie, A., Coutinho, L.L., Kuehn, L., Keele, J., Piper, E., Cook, J., Williams, R., and **Van Tassell, C.P.** Imputation of Microsatellite Alleles from Dense SNP Genotypes for Parentage Verification Across Multiple Bos taurus and Bos indicus breeds. Front. Genet. 4:176. 2013.

Daetwyler, H.D., A. Capitan, H. Pausch, P. Stothard, R. van Binsbergen, R.F. Brøndum, X. Liao, A. Djari, S.C. Rodriguez, C. Grohs, D. Esquerre, O. Bouchez, M.N. Rossignol, C. Klopp, D. Rocha, S. Fritz, A. Eggen, P.J. Bowman, D. Coote, A.J. Chamberlain, C. Anderson, **C.P. Van Tassell**, I. Hulsegge, M.E. Goddard, B. Guldbrandtsen, M.S. Lund, R.F. Veerkamp, D.A. Boichard, R. Fries, and B.J. Hayes. 2014. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nat. Genet. 46:858–865.

McClure, M.C., D. Bickhart, D. Null, P. VanRaden, L. Xu, G. Wiggans, G. Liu, S. Schroeder, J. Glasscock, J. Armstrong, J.B. Cole, **C.P. Van Tassell**, and T.S. Sonstegard. 2014. Bovine exome sequence analysis and targeted SNP genotyping of recessive fertility

defects BH1, HH2, and HH3 reveal a putative causative mutation in SMC2 for HH3. PLoS ONE 9:e92769.

Benavides, M.V., T.S. Sonstegard, S. Kemp, J.M. Mugambi, J.P. Gibson, R.L. Baker, O. Hanotte, K. Marshall, and **C.P. Van Tassell**. 2015. Identification of novel loci associated with gastrointestinal parasite resistance in a Red Maasai x Dorper backcross population. PLoS ONE 10:e0122797.

García-Ruiz, A., J.B. Cole, P.M. VanRaden, G.R. Wiggans, F.J. Ruiz-López, and **C.P. Van Tassell**. 2016. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. Proc. Natl. Acad. Sci. USA 113:E3995–E4004.

Issues of Concern Statement

Animal Care. Not applicable.

Endangered Species. Not applicable.

Environmental Impact Statement. On the basis that this Federal project is undertaken for the sole purpose of conducting research, this project is categorically excluded, in accordance with the National Environmental Policy Act (NEPA). The project facilitates obtaining the national milk supply with a smaller dairy cattle population, thereby reducing any adverse environmental impact that animals may have on the environment.

Human Study Procedure. Not applicable.

Laboratory Hazards. Not applicable.

Occupational Safety and Health. Safety and health concerns are those related to an office setting. A safety inspection of the Laboratory's assigned building area for this project is conducted annually by the Beltsville Area's Occupational Health and Safety Unit. Employees participate in building fire drills (at least one annually).

Recombinant DNA procedures. Not applicable.

Intellectual property issues: The genomic evaluation program is based on genotypes that are not publicly available. Various agreements among industry groups as well as a nonfunded cooperative agreement between ARS and the Council on Dairy Cattle Breeding specify that industry authorizes access to the materials. The ARS nonfunded cooperative agreement also addresses control of research results originating from the material. Nondisclosure agreements will be concluded with industry partners in the development of genotyping chips to permit them to control release of information about their products.

Appendices

Appendix 1

Nonfunded Cooperative Agreement, Council on Dairy Cattle Breeding

UNITED STATES DEPARTMENT OF AGRICULTURE - RESEARCH, EDUCATION, AND ECONOMICS
Award Face Sheet

Agency Control No. 31800

Agency
AGRICULTURAL RESEARCH SERVICE

Type of Instrument
Non-Funded Cooperative Agreement

Authority
7 U.S.C.3318(b)

Title of Project
National Dairy Genetic Evaluation Program

Agreement Number/ FAIN
58-1245-3-228-N

Type of Action
New

Correction
Y

CFDA No.
10.001

Period of Performance
Start 12/31/2012 End 12/30/2017

(Reserved)

Total Federal Amount
$0.00

Federal Amount Obligated by This Action
$0.00

Indirect Cost Rate
0.00%

Agency Administrative Point of Contact/ ADO
MARY L WILKINS
USDA, ARS, AFM, EBSC
BUDGET, TRAVEL, AND AGREEMENTS BRANCH, EBSC
10300 BALTIMORE AVE, BG 003/ RM 331, BARC-WEST
BELTSVILLE MD 20705
Phone: Fax:
E-mail:

Non-Federal Entity/ Federal Agency (Legal Name and Address)
COUNCIL ON DAIRY CATTLE BREEDING
6468 East Main Street
REYNOLDSBURG OH 43068-2362 US
POC:
Phone: 614-861-3636 Fax:
E-mail: omeland@accelgen.com DUNS ID:

Agency Principal Investigator
PAUL M VAN RADEN
10300 BALTIMORE AVENUE
BLDG. 005, RM. 306, BARC-WEST
BELTSVILLE MD 20705
Phone: Fax: 301-504-8092
E-mail: paul.vanraden@ars.usda.gov

Non-Federal Entity/ Federal Agency Principal Investigator
OLE M MELAND
6468 East Main Street
REYNOLDSBURG OH 43068-2362
Phone: 614-861-3636 Fax: 614-861-8040
E-mail: omeland@accelgen.com

Agency Finance Office

Method of Payment
- ☐ HHS/ Payment Management System
- ☐ EFT/ Treasury Check
- ☐ Agency Receives Funds
- ☐ ASAP
- ☐ Advance Payment Authorized
- ☐ Pre-Award Costs Authorized
- ☐ UES (for FAS awards only)
- ☐ IPAC

PROVISIONS

This Agreement incorporates the following:
- ☐ Statement/ Scope of Work
- ☐ Proposal
- ■ Non-Federal Entity Proposal/ Award/ Agreement
- ☐ Research & Related Budget (Total Fed + Non-Fed) or REE-454
- ☐ Research & Related Budget or REE-455
- ☐ Prime Award attached (for subrecipients)
- ☐ Comments (REE-451, page 2)

These are available at http://www.afm.ars.usda.gov/agreements/partnership.htm:
- ■ Conflict of Interest Policy
- ■ USDA Civil Rights Policy Statement
- ■ USDA Civil Rights Poster (AD-475-C)
- ☐ ARS-157 - Research Support Agreement Management Report Template

Reporting Requirements:
Submit to: ■ Agency PI ☐ ADO ☐ Non-Federal Entity/ Federal Agency

Performance Reports	Financial Reports	Management Reports
☐ Quarterly	☐ Quarterly	☐ Monthly
☐ Semi-Annual	☐ Semi-Annual	☐ Quarterly
■ Annual	☐ Annual	☐ Semi-Annual
☐ Final	☐ Final	☐ Final

- ☐ Intellectual Property Reports (www.iEdison.gov)
- ☐ Form SF-428-B Tangible Personal Property Report - Final Report

Applicable Regulations, Terms and Conditions, and Required Certifications (available at http://www.afm.ars.usda.gov/agreements/partnership.htm)
- ☐ 2 CFR Part 200 and 2 CFR Part 400
- ■ 7 CFR Part 550 - General Administrative Policy for Non-Assistance Cooperative Agreements
- ☐ General Provisions, Research Support Agreement (REE-452R)
- ☐ General Provisions, Trust Fund and Reimbursable Cooperative Agreements (REE-22)
- ☐ AD-1047 - Certification Regarding Debarment, Suspension and other Responsibility Matters - Primary Covered Transactions
- ☐ AD-1048 - Certification Regarding Debarment, Suspension, Ineligibility and Voluntary Exclusion - Lower Tier Covered Transactions
- ☐ AD-1049 - Certification Regarding Drug-Free Workplace Requirements (Grants) - Alt I - For Grantees Other Than Individuals
- ☐ AD-1050 - Certification Regarding Drug-Free Workplace Requirements (Grants) - Alt II - For Grantees Who Are Individuals
- ☐ AD-1052 - Certification Regarding Drug-Free Workplace State and State Agencies
- ☐ AD-3031 - Assurance Regarding Felony Conviction or Tax Delinquent Status for Corporate Applicants
- ☐ Certification Regarding Lobbying
- ■ REE-26 - Organization Information, Representations, Assurances & Certifications

This agreement, subject to the provisions above, is executed by the United States Department of Agriculture:

Signature	ADO Name	Federal Award Date
	MARY L WILKINS	

By signing this agreement, the signor certifies that they are vested with the authority to enter into this agreement.

Non-Federal Entity/ Federal Agency Signature	Name and Title	Date
Non-Federal Entity/ Federal Agency Signature	Name and Title	Date

REE-451 10/ 2015 USDA-REE Page: 1

01/12/2017

U.S. DEPARTMENT OF AGRICULTURE
RESEARCH, EDUCATION, AND ECONOMICS

Statement of Work

Agreement No.: 58-1245-3-228-N **Type of Action:** New Correction
Cooperator: COUNCIL ON DAIRY CATTLE BREEDING **Control No.:** 31800

Objective:

The objective of this agreement is to improve the productivity, efficiency, conformation, and health of the national dairy herd as well as the composition and quality of resulting products for the benefit of dairy producers and the consuming public by transferring the research knowledge of USDA-ARS scientists to cooperator on the appropriate genetic models and computational procedures for the analysis of the Phenotypic and Genomic Data contained in the cooperator Database controlled by cooperator.

Approach:

A. Collect phenotypic, genomic, and proteomic data.
B. Edit collected data for accuracy prior to inclusion in the cooperator database and maintain data in a database to support calculation of genetic evaluations (CGE).
C. Analyze collected data through the CGE to produce estimates of genetic merit.
D. Distribute estimates of generic merit.

STATEMENT OF MUTUAL INTEREST:

Both parties want highly accurate and comprehensive genetic evaluations of dairy cattle to improve the productivity, efficiency, conformation, and health of the national dairy herd. The parties agree that meeting the objectives of this project will strengthen and enhance ongoing research within the scope of this agreement. ARS has research expertise in developing and implementing methodology for national genetic evaluations of dairy animals. The cooperator can acquire the resources to provide the services that the industry deems worthwhile and provide incentives to acquire data on additional traits.

THE COOPERATOR AGREES TO:

A. Continue to collect Phenotypic and Genomic Data and collect Proteomic Data from any party with a release providing data to the Cooperator Database.

B. Edit and maintain collected data:

1. Assure data are edited for accuracy prior to inclusion in the Cooperator Database.

2. Conduct an effective quality certification program to assure that only high quality data are added to the Cooperator Database.

3. Maintain and control the Cooperator Database accessed as part of CGE with access by USDA-ARS only for non-commercial research purposes.

C. Carry out the CGE applying appropriate analysis procedures to produce Estimates of Genetic Merit from data in the Cooperator Database.

D. Distribute Estimates of Genetic Merit, for a fee if appropriate.

E. Host cooperator website, for a fee if appropriate, for access by those agreeing to the terms and conditions of the cooperator.

F. Supply its own administrative support as the USDA-ARS funded program support assistant is responsible for USDA-ARS support only.

G. Supply its own e-mail, web addresses, and internet connection.

H. Within two years of the Effective Date of this agreement, supply a computer environment capable of supporting the increasing workload of its responsibilities under this Agreement.

AGENCY AGREES TO:

A. AIPL will research data quality issues and develop methods to insure data added to the Cooperator Database are of high quality.

B. AIPL will develop effective analysis procedures for CGE to compute Estimates of Genetic Merit of Dairy Animals from

01/12/2017

U.S. DEPARTMENT OF AGRICULTURE
RESEARCH, EDUCATION, AND ECONOMICS
Statement of Work

Agreement No.: 58-1245-3-228-N **Type of Action:** New Correction
Cooperator: COUNCIL ON DAIRY CATTLE BREEDING **Control No.:** 31800

datasets in the Cooperator Database provided by cooperator.

C. AIPL will periodically publish documentation of enhancements for procedures used to compute Estimates of Genetic Merit of Dairy Animals and provide summaries of AIPL Algorithms to others for educational purposes as appropriate.

D. AIPL will determine when the information and potential benefit is sufficient to develop AIPL Algorithms for new traits.

MUTUAL AGREEMENTS:

1. See applicable Provisions on the website:
   https://www.afm.ars.usda.gov/agreements/partnership.htm
2. Correspondence and documentation submitted by the Cooperator to the Agency PI in reference to this agreement should cite Agreement No. 58-1245-3-228-N .

Both parties mutually agree to continue collection of genotypes and phenotypes and begin collection of new phenotypes and proteomic data to improve the accuracy and comprehensiveness of the national dairy database accessed for the purposes of the calculation of genetic evaluations; both parties will work to improve accuracy of prediction of economically important traits and investigate methods to incorporate high-density genomic data.

Appendix 2

Trust Fund Cooperative Agreement, Council on Dairy Cattle Breeding

UNITED STATES DEPARTMENT OF AGRICULTURE - RESEARCH, EDUCATION, AND ECONOMICS
Award Face Sheet

Agency Control No. 34164

Agency
AGRICULTURAL RESEARCH SERVICE

Type of Instrument
Trust Fund Cooperative Agreement

Authority
7 U.S.C. 450a

Title of Project
Improvement of Genetic Evaluation of Dairy Cattle through Expanded Genomic Data and Improved Computing Procedures

Agreement Number/ FAIN
58-1245-4-002

Type of Action Correction
Amendment 03 N

CFDA No.
10.001

Period of Performance
Start 10/01/2013 End 07/23/2018

Total Amount
$130,000.00

Amount Obligated by This Action
$0.00

Indirect Cost Rate
0.00%

Agency Administrative Point of Contact/ ADO
MARY L WILKINS
USDA, ARS, AFM, EBSC
BUDGET, TRAVEL, AND AGREEMENTS BRANCH, EBSC
10300 BALTIMORE AVE, BG 003/ RM 331, BARC-WEST
BELTSVILLE MD 20705
Phone: Fax:
E-mail:

Non-Federal Entity/ Federal Agency (Legal Name and Address)
COUNCIL ON DAIRY CATTLE BREEDING
6468 E. Main St
REYNOLDSBURG OH 43068-2362 US
POC: JOAO DURR
CDCB CEO
Phone: 240-334-7744 Fax:
E-mail: joao.durr@cdcb.us DUNS ID:

Agency Principal Investigator
PAUL M VAN RADEN
AGIL
BLDG 005, RM 306, BARC-WEST, 10300 BALTIMORE AVE
BELTSVILLE MD 20705
Phone: Fax: 301-504-8092
E-mail: paul.vanraden@ars.usda.gov

Non-Federal Entity/ Federal Agency Principal Investigator
OLE MELAND
6468 E. Main St
REYNOLDSBURG OH 43068-2362
Phone: 608-365-3332 Fax: 608-365-6644
E-mail: omeland@accelgen.com

Agency Finance Office
JENNIFER C COURCHAINE
USDA, ARS, AFM, EASTERN BUSINESS SERVICE CENTER
BUDGET, TRAVEL, AND AGREEMENTS BRANCH
10300 BALTIMORE AVE, BLDG 3
Beltsville MD 20705
Phone: 301-504-5978 Fax: 301-504-5535
E-mail: jennifer.courchaine@ars.usda.g

Method of Payment
- ☐ HHS/ Payment Management System
- ☐ EFT/ Treasury Check
- ■ Agency Receives Funds
- ☐ ASAP
- ☐ Advance Payment Authorized
- ☐ Pre-Award Costs Authorized
- ☐ UES (for FAS awards only)
- ☐ IPAC

PROVISIONS

This Agreement incorporates the following:
- ☐ Statement/ Scope of Work
- ☐ Proposal
- ☐ Non-Federal Entity Proposal/ Award/ Agreement
- ■ Research & Related Budget (Total Fed + Non-Fed) or REE-454
- ☐ Research & Related Budget or REE-455
- ☐ Prime Award attached (for subrecipients)
- ☐ Comments (REE-451, page 2)

These are available at http://www.afm.ars.usda.gov/agreements/partnership.htm:
- ☐ Conflict of Interest Policy
- ☐ USDA Civil Rights Policy Statement ☐ USDA Civil Rights Poster (AD-475-C)
- ☐ ARS-157 - Research Support Agreement Management Report Template

Reporting Requirements:
Submit to: ☐ Agency PI ☐ ADO ■ Non-Federal Entity/ Federal Agency

Performance Reports	Financial Reports	Management Reports
☐ Quarterly	☐ Quarterly	☐ Monthly
☐ Semi-Annual	☐ Semi-Annual	☐ Quarterly
☐ Annual	☐ Annual	☐ Semi-Annual
☐ Final	☐ Final	☐ Final

- ☐ Intellectual Property Reports (www.iEdison.gov)
- ☐ Form SF-428-B Tangible Personal Property Report - Final Report

Applicable Regulations, Terms and Conditions, and Required Certifications (available at http://www.afm.ars.usda.gov/agreements/partnership.htm)
- ☐ 2 CFR Part 200 and 2 CFR Part 400
- ☐ 7 CFR Part 550 - General Administrative Policy for Non-Assistance Cooperative Agreements
- ☐ General Provisions, Research Support Agreement (REE-452R)
- ■ General Provisions, Trust Fund and Reimbursable Cooperative Agreements (REE-22)
- ☐ AD-1047 - Certification Regarding Debarment, Suspension and other Responsibility Matters - Primary Covered Transactions
- ☐ AD-1048 - Certification Regarding Debarment, Suspension, Ineligibility and Voluntary Exclusion - Lower Tier Covered Transactions
- ☐ AD-1049 - Certification Regarding Drug-Free Workplace Requirements (Grants) - Alt I - For Grantees Other Than Individuals
- ☐ AD-1050 - Certification Regarding Drug-Free Workplace Requirements (Grants) - Alt II - For Grantees Who Are Individuals
- ☐ AD-1052 - Certification Regarding Drug-Free Workplace State and State Agencies
- ☐ AD-3031 - Assurance Regarding Felony Conviction or Tax Delinquent Status for Corporate Applicants
- ☐ Certification Regarding Lobbying
- ☐ REE-26 - Organization Information, Representations, Assurances & Certifications

This agreement, subject to the provisions above, is executed by the United States Department of Agriculture:

Signature	ADO Name	Date
	MARY L WILKINS	

By signing this agreement, the signor certifies that they are vested with the authority to enter into this agreement.

Non-Federal Entity/ Federal Agency Signature	Name and Title	Date
Non-Federal Entity/ Federal Agency Signature	Name and Title	Date

REE-451 10/ 2015 USDA-REE Page: 1

01/12/17

U.S. DEPARTMENT OF AGRICULTURE - RESEARCH, EDUCATION, AND ECONOMICS
Award Face Sheet
REE-451, Page 2

Agreement Number/FAIN: 58-1245-4-002
Type of Action: Amendment 03
Project Number: 8042-31000-101-09T
Accession No.: 425556

Agency Control No.: 34164
BOC: 0199
FMMI Customer Code :

Agency Funds Chargeable - Agency Use Only

Account Code	FY	Amount

Comments:

This amendment is being issued to change ADODR from George R. Wiggans to Paul M. VanRaden due to the retirement of Dr. Wiggans.

01/12/2017

U.S. DEPARTMENT OF AGRICULTURE
RESEARCH, EDUCATION, AND ECONOMICS

Statement of Work

Agreement No.: 58-1245-4-002 **Type of Action:** Amendment 03

Cooperator: COUNCIL ON DAIRY CATTLE BREEDING **Control No.:** 34164

Objective:

Improve genetic evaluations provided to the U.S. dairy industry by the Council on Dairy Cattle Breeding through international collaboration by BA researchers on procedures and sharing of full-sequence genotypic data and through BA development of improved computing procedures to process and deliver genome-based information.

Approach:

Cooperator funding will be used to support travel of BA collaborators to international meetings where advanced procedures for genomic evaluation are presented, purchase and support of computer equipment required for processing of full-sequence data by BA collaborators, and acquisition of sequence data and data on novel traits of economic interest to the U.S. dairy industry.

THE COOPERATOR AGREES TO:

Sponsor's support letter page 5 of this package.

AGENCY AGREES TO:

Conduct these portions of the research effort:

a. Will consult with colleagues on advanced procedures for genomic evaluation. Create the hardware and software environment required for processing full sequence data, obtain full sequence data and obtain data on novel traits of economic interest to the U.S. dairy industry.

b. Investigate improved genetic evaluaton procedures.

c. Provide expertise to COOPERATOR in genetics and breeding.

MUTUAL AGREEMENTS:

1. Agency receives funds.
2. See applicable Provisions on the website:
   https://www.afm.ars.usda.gov/agreements/partnership.htm
3. Correspondence and documentation submitted by the Cooperator to the Agency PI in reference to this agreement should cite Agreement No. 58-1245-4-002 .

U.S. DEPARTMENT OF AGRICULTURE
RESEARCH, EDUCATION, AND ECONOMICS

AGREEMENT BUDGET

COOPERATOR: COUNCIL ON DAIRY CATTLE BREEDING **AGREEMENT NO.:** 58-1245-4-002 **TYPE OF ACTION:** Amendment 03	AGENCY TO RECEIVE	COOPERATOR CONTRIBUTION(S)
A. Salaries and Wages		
1. Senior/ Key Person(s)	$0.00	$0.00
2. Other Personnel (Post-Doctoral Associates, Graduate Students, Undergraduate Students)	$0.00	$0.00
3. Support Personnel/ Secretarial/ Clerical	$0.00	$0.00
Total Salaries and Wages →	$0.00	$0.00
B. Fringe Benefits (If charged as Direct Costs)	$0.00	$0.00
C. Total Salaries, Wages, and Fringe Benefits (A plus B) →	$0.00	$0.00
D. Equipment (Provide supporting data; list items and dollar amounts for each item exceeding $5,000)	$0.00	$0.00
E. Materials and Supplies	$0.00	$0.00
F. Travel (List destination and amount for each trip)		
1. Domestic (Include Canada, Mexico, and U.S. Possessions)	$0.00	$0.00
2. Foreign	$0.00	$0.00
G. Publication Costs	$0.00	$0.00
H. ADP/ Computer Services	$0.00	$0.00
I. Subawards	$0.00	$0.00
J. All Other Direct Costs (Provide supporting data. List items and dollar amounts for each item.)	$0.00	$0.00
K. Total Direct Costs (C through J) →	$0.00	$0.00
L. Indirect Costs (Specify rate and base) Rate : 0.00 % Base : $0.00	$0.00	$0.00
M. Total Costs (K plus L) →	$0.00	$0.00

NOTES:

1. A separate budget is required for each year.
2. Federal Statute (7 U.S.C. 3318 (b)(1)(B)) requires a contribution of resources by all parties toward meeting the objectives of the Cooperative Agreement.
3. The Cooperator's contribution must be no less than 20 percent of the total of the resource contributions under the cooperative agreement. Resource contributions of the Cooperator must consist of a sufficient amount of itemized direct costs to substantiate a true stake in the project as determined by the ADO. The Cooperator's contribution must be maintained at 20 percent of Federal funding throughout the period of performance.
4. Under 7 USC 3319, USDA is prohibited from reimbursing State Cooperative Institutions for indirect costs or tuition remission in connection with non-assistance cooperative agreements awarded under the authority of 7 USC 3318(b).
5. Indirect costs will be reimbursed only upon receipt of a current approved Negotiated Indirect Cost Rate Agreement for all non-State Cooperative Institutions.
6. Unrecovered indirect costs may be used to meet a portion of the resource contribution requirement toward the cooperative effort. Indirect costs only for the resource contribution requirement is not allowed.
7. Unallowable costs as defined in 2 CFR Part 200, Subpart E - Cost Principles, cannot be considered a resource contribution.

REE-454 (10/15)

Appendix 3

Nonfunded Cooperative Agreement, Department of Environment and Primary Industries

UNITED STATES DEPARTMENT OF AGRICULTURE - RESEARCH, EDUCATION, AND ECONOMICS
Award Face Sheet

Agency Control No. 36141

Agency: AGRICULTURAL RESEARCH SERVICE	Type of Instrument: Non-Funded Cooperative Agreement	Authority: 7 U.S.C.3318(b)
Title of Project: 1000 Bull Genomes Project	Agreement Number/ FAIN: 58-1245-4-064-FN; Type of Action: Amendment 01; Correction: N	CFDA No.: 10.001
	Period of Performance: Start 03/01/2014 End 03/01/2019	
	Total Federal Amount: $0.00; Federal Amount Obligated by This Action: $0.00	Indirect Cost Rate: 0.00%

Agency Administrative Point of Contact/ ADO
DEBORAH L WATSON
USDA, REE, ARS, AFM, FMAD
5601 SUNNYSIDE AVE./ MS 5110
Room 3-2170A
BELTSVILLE MD 20705
Phone: Fax:
E-mail:

Agency Principal Investigator
JOHN BRUCE COLE
AGIL
BLDG 005, RM 306, BARC-WEST, 10300 BALTIMORE AVE
BELTSVILLE MD 20705
Phone: Fax: 301-504-8092
E-mail: john.cole@ars.usda.gov

Agency Finance Office

Non-Federal Entity/ Federal Agency (Legal Name and Address)
DEPT OF ENVIRONMENT & PRIMARY INDUSTRIES
1 Park Drive, Victoria, Australia, 3086 AS
POC:
Phone: +136 186 Fax:
E-mail: ben.hays@depi.vic.gov.au DUNS ID:

Non-Federal Entity/ Federal Agency Principal Investigator
BEN J HAYES
Department of Environment and Primary Industries
1 Park Drive
VICTORIA 3086
Phone: +136 186 Fax:
E-mail: ben.hayes@dpi.vic.gov.au

Method of Payment
- ☐ HHS/ Payment Management System
- ☐ EFT/ Treasury Check
- ☐ Agency Receives Funds
- ☐ ASAP
- ☐ Advance Payment Authorized
- ☐ Pre-Award Costs Authorized
- ☐ UES (for FAS awards only)
- ☐ IPAC

PROVISIONS

This Agreement incorporates the following:
- ■ Statement/ Scope of Work
- ☐ Proposal
- ☐ Non-Federal Entity Proposal/ Award/ Agreement
- ☐ Research & Related Budget (Total Fed + Non-Fed) or REE-454
- ☐ Research & Related Budget or REE-455
- ☐ Prime Award attached (for subrecipients)
- ■ Comments (REE-451, page 2)

These are available at http://www.afm.ars.usda.gov/agreements/partnership.htm:
- ■ Conflict of Interest Policy
- ■ USDA Civil Rights Policy Statement ■ USDA Civil Rights Poster (AD-475-C)
- ☐ ARS-157 - Research Support Agreement Management Report Template

Reporting Requirements:
Submit to: ■ Agency PI ☐ ADO ☐ Non-Federal Entity/ Federal Agency

Performance Reports	Financial Reports	Management Reports
☐ Quarterly	☐ Quarterly	☐ Monthly
☐ Semi-Annual	☐ Semi-Annual	☐ Quarterly
■ Annual	☐ Annual	☐ Semi-Annual
■ Final	☐ Final	☐ Final

- ☐ Intellectual Property Reports (www.iEdison.gov)
- ☐ Form SF-428-B Tangible Personal Property Report - Final Report

Applicable Regulations, Terms and Conditions, and Required Certifications (available at http://www.afm.ars.usda.gov/agreements/partnership.htm)
- ☐ 2 CFR Part 200 and 2 CFR Part 400
- ■ 7 CFR Part 550 - General Administrative Policy for Non-Assistance Cooperative Agreements
- ☐ General Provisions, Research Support Agreement (REE-452R)
- ☐ General Provisions, Trust Fund and Reimbursable Cooperative Agreements (REE-22)
- ☐ AD-1047 - Certification Regarding Debarment, Suspension and other Responsibility Matters - Primary Covered Transactions
- ☐ AD-1048 - Certification Regarding Debarment, Suspension, Ineligibility and Voluntary Exclusion - Lower Tier Covered Transactions
- ☐ AD-1049 - Certification Regarding Drug-Free Workplace Requirements (Grants) - Alt I - For Grantees Other Than Individuals
- ☐ AD-1050 - Certification Regarding Drug-Free Workplace Requirements (Grants) - Alt II - For Grantees Who Are Individuals
- ☐ AD-1052 - Certification Regarding Drug-Free Workplace State and State Agencies
- ☐ AD-3031 - Assurance Regarding Felony Conviction or Tax Delinquent Status for Corporate Applicants
- ☐ Certification Regarding Lobbying
- ☐ REE-26 - Organization Information, Representations, Assurances & Certifications

This agreement, subject to the provisions above, is executed by the United States Department of Agriculture:

Signature	ADO Name	Federal Award Date
	DEBORAH L WATSON	

By signing this agreement, the signor certifies that they are vested with the authority to enter into this agreement.

Non-Federal Entity/ Federal Agency Signature	Name and Title	Date
Non-Federal Entity/ Federal Agency Signature	Name and Title	Date

REE-451 10/2015 USDA-REE Page: 1

01/12/17

U.S. DEPARTMENT OF AGRICULTURE - RESEARCH, EDUCATION, AND ECONOMICS
Award Face Sheet
REE-451, Page 2

Agreement Number/ FAIN: 58-1245-4-064-FN
Type of Action: Amendment 01
Project Number: 8042-31000-101-11N
Accession No.: 426790

Agency Control No.: 36141
BOC: 4500
PO No.:

Comments:

1. This amendment 1 is being issued to change ARS' PI from George R. Wiggans to John B. Cole due to the retirement of Dr. Wiggans.

01/12/2017

U.S. DEPARTMENT OF AGRICULTURE
RESEARCH, EDUCATION, AND ECONOMICS

Statement of Work

Agreement No.: 58-1245-4-064-FN **Type of Action:** Amendment 01

Cooperator: DEPT OF ENVIRONMENT & PRIMARY INDUSTRIES **Control No.:** 36141

Objective:

The objective of this collaboration is to exchange bull genome sequence data in support of research in cattle genetics.

Approach:

Each participating organization will contribute whole genome sequence data and exome sequence data for individuals from the genus Bos or Bos Taurus, not including phenotypic data or breeding values. Each organization must contribute a specified minimum number of whole genome sequences per phase during the term of the agreement. The project aims to develop a resource to allow the project partners to impute full genome sequence in bulls and cows that have been genotyped with 800K and 50K SNP arrays. This could be for the purposes of genomic prediction, genome wide association, and discovery of causal mutations.

STATEMENT OF MUTUAL INTEREST:

Both parties are actively engaged in independent research on increasing our knowledge of dairy catle genetics based on full sequence data. The parties agree that meeting the objectives of this project will strengthen and enhance ongoing research within the scope of this agreement.

THE COOPERATOR AGREES TO:

1. Facilitate the sharing of sequence data among the particiopants.

2. Process raw sequence data to prepare haplotype labraries.

AGENCY AGREES TO:

1. Contribute sequence data to the repository for at least the number of bulls required.

2. Conduct research on the improving processing of sequence data and share the results.

MUTUAL AGREEMENTS:

1. See applicable Provisions on the website:
   https://www.afm.ars.usda.gov/agreements/partnership.htm

2. Correspondence and documentation submitted by the Cooperator to the Agency PI in reference to this agreement should cite Agreement No. 58-1245-4-064-FN .

4. Both parties will respect the confidentiality of the data, only releasing it to those approved to receive it.

5. Both parties will collaborate in inproving the processing of sequence data.

6. Both parties including the individual members of the consortium as appropriate will publish results jointly in peer-reviewed journals.

Appendix 4

Letter of Collaboration, Dr. Christian Maltecca, North Carolina State University

Department of Animal Science
College of Agriculture and Life Sciences

NC STATE UNIVERSITY

Campus Box 7621
Raleigh, NC 27695-7621
Phone: 919-515-2755
Fax: 919.515.6884

January 17, 2017

Dr. Paul M. VanRaden
USDA-ARS
Animal Genomics and Improvement Laboratory
10300 Baltimore Avenue
Bldg. 005, Rm. 306, BARC-West
Beltsville, MD 20705-2350

Dear Paul:

I am writing to indicate my willingness to collaborate with you on your new research project plan entitled "Improving Dairy Animals by Increasing Accuracy of Genomic Prediction, Evaluating New Traits, and Redefining Selection Goals". Specifically, I will assist with research to identify genomic regions that may contain undesirable haplotypes using my HaploFinder software. My group has successfully used our software to identify regions associated with inbreeding depression in Jersey cattle and study runs of homozygosity in purebred and crossbred pigs, and we think it will be helpful to identify causal variants associated with early embryonic loss and decreased fertility.

Sincerely,

Christian Maltecca, Ph.D.
Associate Professor, Faculty Scholar
Animal Science Dept., Genetics Program
North Carolina State University
Campus Box 7621
Raleigh, NC, 27695

Appendix 5

Recessive Loci to Be Evaluated for Inclusion in Selection Indexes

Breed	Haplotype	Functional/*gene* name	Frequency (%)	Chromosome	Region (bp)
Ayrshire	AH1	PIRM/*UBE3B*	13.0	17	65,921,497
Brown Swiss	BH1	—	6.67	7	42,811,272–47,002,161
	BH2	*TUBD1*	7.78	19	11,063,520
	BHD	SDM/*SPAST*	2.19	11	14,742,058
	BHM	SMA/*KDSR*(*FVT1*)	3.61	24	62,118,139–62,156,760
	BHP	Polledness/*POLLED*	0.4	1	1,705,834–1,989,480
	BHW	Weaver/*PNPLA8*	1.56	4	49,878,773
Holstein	HBR	Black/red coat color/*MC1R*(*MSHR*)	0.8	18	14,757,332–14,759,082
	HCD	Cholesterol deficiency/*APOB*	2.5	11	77,958,995
	HDR	Dominant red coat color	0.04	3	9,479,761
	HH0	Brachyspina/*FANCI*	2.76	21	21,184,869–21,188,198
	HH1	*APAF1*	1.92	5	63,150,400
	HH2	—	1.66	1	94,860,836–96,553,339
	HH3	*SMC2*	2.95	8	95,410,507
	HH4	*GART*	0.37	1	1,277,227
	HH5	*TFB1M*	2.22	9	93,223,651–93,370,998
	HHB	BLAD/*ITGB2*	0.25	1	145,119,004
	HHC	CVM/*SLC35A3*	1.37	3	43,412,427
	HHD	DUMPS/*UMPS*	0.01	1	69,757,801
	HHM	Mulefoot/*LRP4*	0.07	15	77,663,790–77,701,209
	HHP	Polledness/*POLLED*	0.71	1	1,705,834–1,989,480
	HHR	Red coat color/*MC1R*(*MSHR*)	5.42	18	14,758,207
Jersey	JH1	*CWC15*	12.10	15	15,707,169
	JH2	—	1.3	26	8,812,759–9,414,082
	JHP	Polledness/*POLLED*	2.2	1	1,705,834–1,989,480

[1]**Source:** AIP Research Report GENOMIC3 (https://aipl.arsusda.gov/reference/recessive_haplotypes_ARR-G3.html)

Appendix 6

Non-Assistance Cooperative Agreement, University of Maryland

UNITED STATES DEPARTMENT OF AGRICULTURE - RESEARCH, EDUCATION, AND ECONOMICS
Award Face Sheet

Agency Control No. 36337

Agency: AGRICULTURAL RESEARCH SERVICE	Type of Instrument: Non-Assistance Cooperative Agreement	Authority: 7 U.S.C.3318(b)
Title of Project: Improvement of Algorithms and Software to Process Very Large Genomic Datasets	Agreement Number/ FAIN: 58-1245-4-070	Type of Action: Amendment 02 Correction: N
	CFDA No.: 10.001	
	Period of Performance: Start 08/01/2014 End 07/31/2018	
	Total Federal Amount: $135,722.00	Federal Amount Obligated by This Action: $35,750.00
	Indirect Cost Rate: 10.00%	

Agency Administrative Point of Contact/ ADO
MARY L WILKINS
USDA, ARS, AFM, EBSC
BUDGET, TRAVEL, AND AGREEMENTS BRANCH, EBSC
10300 BALTIMORE AVE, BG 003/ RM 331, BARC-WEST
BELTSVILLE MD 20705
Phone: Fax:
E-mail:

Agency Principal Investigator
PAUL M VAN RADEN
USDA, ARS, BARC, AGIL
10300 Baltimore Ave., BLDG. 005, RM. 306, BARC-W
BELTSVILLE MD 20705
Phone: Fax: 301-504-8334
E-mail: paul.vanraden@ars.usda.gov

Agency Finance Office
JENNIFER C COURCHAINE
USDA, ARS, AFM, EBSC
BUDGET, TRAVEL, AND AGREEMENTS BRANCH
10300 BALTIMORE AVE, BLDG 3
Beltsville MD 20705
Phone: 301-504-5978 Fax: 301-504-5535
E-mail: jennifer.courchaine@ars.usda.gov

Non-Federal Entity/ Federal Agency (Legal Name and Address)
UNIVERSITY OF MARYLAND
Office of Sponsored Programs
620 W. Lexington Street, 4th. Floor
BALTIMORE MD 21201-1531 US
POC: DANIELLE BROWN
Manager, Sponosored Programs Administration
Phone: 410-706-6723 Fax:
E-mail: team-yellow2@ordmail.umaryland.edu DUNS ID: 188435911

Non-Federal Entity/ Federal Agency Principal Investigator
JEFFREY R O'CONNELL
University of Maryland, School of Medicine
660 Redwood Street, Room 021
BALTIMORE MD 20201-3183
Phone: 410-706-8959 Fax: 410-706-1622
E-mail: joconnel@medicine.umaryland.edu

Method of Payment
- ☐ HHS/ Payment Management System
- ■ EFT/ Treasury Check
- ☐ Agency Receives Funds
- ☐ ASAP
- ☐ Advance Payment Authorized
- ☐ Pre-Award Costs Authorized
- ☐ UES (for FAS awards only)
- ☐ IPAC

PROVISIONS

This Agreement incorporates the following:
- ☐ Statement/ Scope of Work
- ☐ Proposal
- ☐ Non-Federal Entity Proposal/ Award/ Agreement
- ■ Research & Related Budget (Total Fed + Non-Fed) or REE-454
- ☐ Research & Related Budget or REE-455
- ☐ Prime Award attached (for subrecipients)
- ■ Comments (REE-451, page 2)

These are available at http://www.afm.ars.usda.gov/agreements/partnership.htm:
- ■ Conflict of Interest Policy
- ■ USDA Civil Rights Policy Statement ■ USDA Civil Rights Poster (AD-475-C)
- ☐ ARS-157 - Research Support Agreement Management Report Template

Reporting Requirements:
Submit to: ■ Agency PI ☐ ADO ☐ Non-Federal Entity/ Federal Agency

Performance Reports	Financial Reports	Management Reports
☐ Quarterly	☐ Quarterly	☐ Monthly
☐ Semi-Annual	☐ Semi-Annual	☐ Quarterly
■ Annual	■ Annual	☐ Semi-Annual
■ Final	☐ Final	☐ Final

- ■ Intellectual Property Reports (www.iEdison.gov)
- ☐ Form SF-428-B Tangible Personal Property Report - Final Report

Applicable Regulations, Terms and Conditions, and Required Certifications (available at http://www.afm.ars.usda.gov/agreements/partnership.htm)
- ☐ 2 CFR Part 200 and 2 CFR Part 400
- ■ 7 CFR Part 550 - General Administrative Policy for Non-Assistance Cooperative Agreements
- ☐ General Provisions, Research Support Agreement (REE-452R)
- ☐ General Provisions, Trust Fund and Reimbursable Cooperative Agreements (REE-22)
- ☐ AD-1047 - Certification Regarding Debarment, Suspension and other Responsibility Matters - Primary Covered Transactions
- ☐ AD-1048 - Certification Regarding Debarment, Suspension, Ineligibility and Voluntary Exclusion - Lower Tier Covered Transactions
- ☐ AD-1049 - Certification Regarding Drug-Free Workplace Requirements (Grants) - Alt I - For Grantees Other Than Individuals
- ☐ AD-1050 - Certification Regarding Drug-Free Workplace Requirements (Grants) - Alt II - For Grantees Who Are Individuals
- ☐ AD-1052 - Certification Regarding Drug-Free Workplace State and State Agencies
- ☐ AD-3031 - Assurance Regarding Felony Conviction or Tax Delinquent Status for Corporate Applicants
- ■ Certification Regarding Lobbying
- ☐ REE-26 - Organization Information, Representations, Assurances & Certifications

This agreement, subject to the provisions above, is executed by the United States Department of Agriculture:

Signature	ADO Name	Federal Award Date
	MARY L WILKINS	

By signing this agreement, the signor certifies that they are vested with the authority to enter into this agreement.

Non-Federal Entity/ Federal Agency Signature	Name and Title	Date
Non-Federal Entity/ Federal Agency Signature	Name and Title	Date

REE-451 10/2015 USDA-REE Page: 1

01/12/17

U.S. DEPARTMENT OF AGRICULTURE - RESEARCH, EDUCATION, AND ECONOMICS
Award Face Sheet
REE-451, Page 2

Agreement Number/ FAIN: 58-1245-4-070
Type of Action: Amendment 02
Project Number: 8042-31000-101-12S
Accession No.: 426953

Agency Control No.: 36337
BOC: 4521
PO No.: 4500062421
FMMI Vendor Code : 1102511589

Agency Funds Chargeable - Agency Use Only

Account Code	FY	Amount	FMMI Fund Code	Cost Center	WBS Element
601-8042-530	2016	$35,750.00	AR0001BASE	AR80420530	AR.MU.8042.01.0530

Comments:

This Amendment to extend period of agreement from 7/30/2017 to 07/30/2018 and increase agreement amount by an additional $35,750.

01/12/2017

U.S. DEPARTMENT OF AGRICULTURE
RESEARCH, EDUCATION, AND ECONOMICS

Statement of Work

Agreement No.: 58-1245-4-070 **Type of Action:** Amendment 02

Cooperator: UNIVERSITY OF MARYLAND **Control No.:** 36337

Objective:

To 1) develop DNA sequence analysis methods to estimate effects of sequence variants that impact economically important traits in the US dairy industry; 2) refine imputation methods to efficiently and accurately combine genotypes from a variety of input data with differing densities or coverage; and 3) test for parent of origin effects to determine if genes are expressed differently when inherited from the sire or from the dam.

Approach:

Hundreds or thousands of animals may have full DNA sequence data in the near future. Efficient computer methods to process such data will be developed and tested on simulated data before the actual sequences become available. Many more animals already have medium or high density genotypes and can be used to estimate the effects of new sequence variants if these can be accurately imputed. Fast algorithms have been derived and will be updated and refined to impute genotype probabilities and to estimate genetic effects from larger datasets. These imputation methods can determine which allele was inherited from the sire or the dam. Effects of parent of origin (imprinting) will be estimated for several traits using software (MMAP) already developed by the Cooperator.

STATEMENT OF MUTUAL INTEREST:

Both parties are actively engaged in research on statistical and computational methods to evaluate the rapidly growing genomic datasets. The parties agree that the objectives of this agreement will strengthen and enhance the ongoing exchange of research ideas between bovine and human genetic analyses. Both parties are interested in developing and applying new software to datasets containing many individuals, many genetic variants, and a wide variety of missing data patterns. Both parties have cooperated formally and informally on a wide variety of similar research projects over the past 7 years.

THE COOPERATOR AGREES TO:

1. Work closely with ARS in planning and conducting the research outlined herein:

2. Conduct at Cooperator facilities and elsewhere as appropriate, research directed toward the completion of this project to include:

(a) Update computer programs as needed to process new types of data and larger datasets;

(b) Apply the updated methods to research datasets to test their accuracy and computational feasibility;

(c) Assist in deriving new algorithms and methods of analysis for routine genomic predictions.

Text Messaging:

AGENCY AGREES TO:

1. Work closely with the Cooperator in planning and conducting the research outlined herein:

2. Conduct research on the following aspects of the project:

(a) Maintain, edit, and provide actual and simulated data needed to test the new statistical methods and evaluation procedures;

(b) Share access to ARS computer resources to conduct joint analyses of data held in ARS database;

(c) Derive and program new algorithms with more efficient use of parallel processing to allow timely evaluation of larger datasets.

U.S. DEPARTMENT OF AGRICULTURE
RESEARCH, EDUCATION, AND ECONOMICS

AGREEMENT BUDGET

COOPERATOR: UNIVERSITY OF MARYLAND **AGREEMENT NO.:** 58-1245-4-070 **TYPE OF ACTION:** Amendment 02	AGENCY TO REIMBURSE	COOPERATOR CONTRIBUTION(S)
A. Salaries and Wages		
1. Senior/ Key Person(s)	$22,000.00	$5,000.00
2. Other Personnel (Post-Doctoral Associates, Graduate Students, Undergraduate Students)	$0.00	$0.00
3. Support Personnel/ Secretarial/ Clerical	$0.00	$0.00
Total Salaries and Wages →	$22,000.00	$5,000.00
B. Fringe Benefits (If charged as Direct Costs)	$7,500.00	$1,000.00
C. Total Salaries, Wages, and Fringe Benefits (A plus B) →	$29,500.00	$6,000.00
D. Equipment (Provide supporting data; list items and dollar amounts for each item exceeding $5,000)	$0.00	$0.00
E. Materials and Supplies	$0.00	$0.00
F. Travel (List destination and amount for each trip)		
1. Domestic (Include Canada, Mexico, and U.S. Possessions)	$0.00	$0.00
2. Foreign	$3,000.00	$0.00
G. Publication Costs	$0.00	$0.00
H. ADP/ Computer Services	$0.00	$0.00
I. Subawards	$0.00	$0.00
J. All Other Direct Costs (Provide supporting data. List items and dollar amounts for each item.)	$0.00	$0.00
K. Total Direct Costs (C through J) →	$32,500.00	$6,000.00
L. Indirect Costs (Specify rate and base) Rate : 10.00 % Base : $0.00	$3,250.00	$0.00
M. Total Costs (K plus L) →	$35,750.00	$6,000.00

NOTES:

1. A separate budget is required for each year.
2. Federal Statute (7 U.S.C. 3318 (b)(1)(B)) requires a contribution of resources by all parties toward meeting the objectives of the Cooperative Agreement.
3. The Cooperator's contribution must be no less than 20 percent of the total of the resource contributions under the cooperative agreement. Resource contributions of the Cooperator must consist of a sufficient amount of itemized direct costs to substantiate a true stake in the project as determined by the ADO. The Cooperator's contribution must be maintained at 20 percent of Federal funding throughout the period of performance.
4. Under 7 USC 3319, USDA is prohibited from reimbursing State Cooperative Institutions for indirect costs or tuition remission in connection with non-assistance cooperative agreements awarded under the authority of 7 USC 3318(b).
5. Indirect costs will be reimbursed only upon receipt of a current approved Negotiated Indirect Cost Rate Agreement for all non-State Cooperative Institutions.
6. Unrecovered indirect costs may be used to meet a portion of the resource contribution requirement toward the cooperative effort. Indirect costs only for the resource contribution requirement is not allowed.
7. Unallowable costs as defined in 2 CFR Part 200, Subpart E - Cost Principles, cannot be considered a resource contribution.

REE-454 (10/15)

U.S. DEPARTMENT OF AGRICULTURE
RESEARCH, EDUCATION, AND ECONOMICS

AGREEMENT BUDGET JUSTIFICATION
FOR 58-1245-4-070
Amendment No. 2

Salary Summary

salary	$29,500.00

Travel - Foreign

Chile (1 @ $3000)	$3,000.00

Indirect Costs

IPSC	$3,250.00

Appendix 7

Reimbursable Cooperative Agreement, University of Maryland

UNITED STATES DEPARTMENT OF AGRICULTURE - RESEARCH, EDUCATION, AND ECONOMICS
Award Face Sheet

Agency Control No. 41717

Agency
AGRICULTURAL RESEARCH SERVICE

Type of Instrument
Reimbursable Cooperative Agreement

Authority
7 U.S.C. 450a

Title of Project
Sequence-Based Big Data Genomic Discovery and Application to Improve Dairy Fertility

Agreement Number/ FAIN
58-8042-6-034

Type of Action
New

Correction
Y

CFDA No.
10.001

Period of Performance
Start 02/15/2016 End 02/14/2019

Incoming Award Number
Z5773002

Total Amount
$22,000.00

Amount Obligated by This Action
$0.00

Indirect Cost Rate
0.00%

Agency Administrative Point of Contact/ ADO
MARY L WILKINS
USDA, ARS, AFM, EBSC
BUDGET, TRAVEL, AND AGREEMENTS BRANCH
10300 BALTIMORE AVE, BG 003/ RM 331, BARC-WEST
BELTSVILLE MD 20705-2350
Phone: 301-504-7647 Fax:
E-mail: mary.wilkins@ars.usda.gov

Non-Federal Entity/ Federal Agency (Legal Name and Address)
UNIVERSITY OF MARYLAND
Office of Research Administration
3112 Lee Building
COLLEGE PARK MD 20742-5141 US
POC: SUSAN GROSSMAN
Senior Contract Administrator
Phone: 301-405-5736 Fax: 301-314-9569
E-mail: oraa@umd.edu; sgossman@umd.edu DUNS ID: 790934285

Agency Principal Investigator
JOHN BRUCE COLE
10300 BALTIMORE AVENUE
BLDG. 005, RM. 306, BARC-WEST
BELTSVILLE MD 20705
Phone: Fax: 301-504-8092
E-mail: john.cole@ars.usda.gov

Non-Federal Entity/ Federal Agency Principal Investigator
LI MA
Assistant Professor
UMD, Department of Animal and Avian Sciences
2123 Animal Science Building
COLLEGE PARK MD 20742-0001
Phone: 301-405-1389 Fax:
E-mail: lima@umd.edu

Agency Finance Office
JENNIFER C COURCHAINE
USDA, ARS, AFM, EASTERN BUSINESS SERVICE CENTER
BUDGET, TRAVEL, AND AGREEMENTS BRANCH
10300 BALTIMORE AVE, BLDG 3
Beltsville MD
Phone: 301-504-5978 Fax: 301-504-5535
E-mail: jennifer.courchaine@ars.usda.g

Method of Payment
- ☐ HHS/ Payment Management System
- ☐ EFT/ Treasury Check
- ■ Agency Receives Funds
- ☐ ASAP
- ☐ Advance Payment Authorized
- ☐ Pre-Award Costs Authorized
- ☐ UES (for FAS awards only)
- ☐ IPAC

PROVISIONS

This Agreement incorporates the following:
- ■ Statement/ Scope of Work
- ☐ Proposal
- ☐ Non-Federal Entity Proposal/ Award/ Agreement
- ☐ Research & Related Budget (Total Fed + Non-Fed) or REE-454
- ☐ Research & Related Budget or REE-455
- ☐ Prime Award attached (for subrecipients)
- ☐ Comments (REE-451, page 2)

These are available at http://www.afm.ars.usda.gov/agreements/partnership.htm:
- ☐ Conflict of Interest Policy
- ☐ USDA Civil Rights Policy Statement ☐ USDA Civil Rights Poster (AD-475-C)
- ☐ ARS-157 - Research Support Agreement Management Report Template

Reporting Requirements:
Submit to: ☐ Agency PI ☐ ADO ■ Non-Federal Entity/ Federal Agency

Performance Reports	Financial Reports	Management Reports
☐ Quarterly	☐ Quarterly	☐ Monthly
☐ Semi-Annual	☐ Semi-Annual	☐ Quarterly
☐ Annual	☐ Annual	☐ Semi-Annual
■ Final	■ Final	☐ Final

- ☐ Intellectual Property Reports (www.iEdison.gov)
- ☐ Form SF-428-B Tangible Personal Property Report - Final Report

Applicable Regulations, Terms and Conditions, and Required Certifications (available at http://www.afm.ars.usda.gov/agreements/partnership.htm)
- ☐ 2 CFR Part 200 and 2 CFR Part 400
- ☐ 7 CFR Part 550 - General Administrative Policy for Non-Assistance Cooperative Agreements
- ☐ General Provisions, Research Support Agreement (REE-452R)
- ■ General Provisions, Trust Fund and Reimbursable Cooperative Agreements (REE-22)
- ☐ AD-1047 - Certification Regarding Debarment, Suspension and other Responsibility Matters - Primary Covered Transactions
- ☐ AD-1048 - Certification Regarding Debarment, Suspension, Ineligibility and Voluntary Exclusion - Lower Tier Covered Transactions
- ☐ AD-1049 - Certification Regarding Drug-Free Workplace Requirements (Grants) - Alt I - For Grantees Other Than Individuals
- ☐ AD-1050 - Certification Regarding Drug-Free Workplace Requirements (Grants) - Alt II - For Grantees Who Are Individuals
- ☐ AD-1052 - Certification Regarding Drug-Free Workplace State and State Agencies
- ☐ AD-3031 - Assurance Regarding Felony Conviction or Tax Delinquent Status for Corporate Applicants
- ☐ Certification Regarding Lobbying
- ☐ REE-26 - Organization Information, Representations, Assurances & Certifications

This agreement, subject to the provisions above, is executed by the United States Department of Agriculture:

Signature	ADO Name	Date
	MARY L WILKINS	

By signing this agreement, the signor certifies that they are vested with the authority to enter into this agreement.

Non-Federal Entity/ Federal Agency Signature	Name and Title	Date
Non-Federal Entity/ Federal Agency Signature	Name and Title	Date

REE-451 10/2015 USDA-REE Page: 1

01/12/2017

U.S. DEPARTMENT OF AGRICULTURE
RESEARCH, EDUCATION, AND ECONOMICS

Statement of Work

Agreement No.: 58-8042-6-034 **Type of Action:** New Correction
Cooperator: UNIVERSITY OF MARYLAND **Control No.:** 41717

Objective:

The overall goal of this research is to discover causative genomic variations and apply those genomic discoveries to improve dairy fertility. Specific aims supporting the overall objective are: 1) Identify genomic regions associated with dairy fertility using sequencing-based GWAS and selection signature analysis of contemporary Holsteins and those unselected for 50 years; and 2) Pinpoint and apply causative genetic variants to develop optimal strategies of genomic selection to improve dairy fertility.

Approach:

ARS will process 110 semen samples from sires of importance, and genomic DNA will be extracted from semen using an adapted protocol from Qiagen. These 110 bulls will be sequenced to a coverage depth of 10X using the established NGS facility at USDA AGIL. The 125 libraries will be pooled by group and within-SNP effect and trait to target a depth of about 10X sequence coverage per animal (30 billion base pairs) using paired end 150-bp reads on the NextSeq. The goal of the SNP and sequence-based GWAS analysis is to identify locations of causal loci affecting fertility traits. The SNP-based GWAS analysis includes 60K SNP genotype and 15 PTA phenotypes of more than 760,000 Holstein cows and bulls, while the sequence-based GWAS includes millions of SNPs and 15 yield deviation (YD) phenotypes of over 10,000. The SNP-based analysis will use PTA to test for additive effects and additive × additive interactions, while the sequence-based analysis will use YD to test for additional effects including dominance, additive × dominance, dominance × dominance, and imprinting effects. Selection signature analysis will be conducted within each of the three groups to identify loci under selection.

THE COOPERATOR AGREES TO:

Perform SNP-based and sequence-based GWAS, identify signatures of selection associated with fertility, and discovery causal polymorphisms of fertility traits.

AGENCY AGREES TO:

Provide fertility phenotypes for the GWAS analysis, as well as SNP genotype and whole-genome sequence data for the signatures of selection and variant discovery studies.

MUTUAL AGREEMENTS:

1. Agency receives funds.
2. Correspondence and documentation submitted by the Cooperator to the Agency PI in reference to this agreement should cite Agreement No. 58-8042-6-034 .

3. Share equally in the exchange of research data obtained, evaluate the analysis and interpretation of results, and jointly submit manuscripts to peer-reviewed journals.

U.S. DEPARTMENT OF AGRICULTURE
RESEARCH, EDUCATION, AND ECONOMICS

AGREEMENT BUDGET

COOPERATOR: UNIVERSITY OF MARYLAND **AGREEMENT NO.:** 58-8042-6-034 **TYPE OF ACTION:** New Correction	AGENCY TO RECEIVE	COOPERATOR CONTRIBUTION(S)
A. Salaries and Wages		
1. Senior/ Key Person(s)	$0.00	$0.00
2. Other Personnel (Post-Doctoral Associates, Graduate Students, Undergraduate Students)	$0.00	$0.00
3. Support Personnel/ Secretarial/ Clerical	$0.00	$0.00
Total Salaries and Wages →	$0.00	$0.00
B. Fringe Benefits (If charged as Direct Costs)	$0.00	$0.00
C. Total Salaries, Wages, and Fringe Benefits (A plus B) →	$0.00	$0.00
D. Equipment (Provide supporting data; list items and dollar amounts for each item exceeding $5,000)	$0.00	$0.00
E. Materials and Supplies	$0.00	$0.00
F. Travel (List destination and amount for each trip)		
1. Domestic (Include Canada, Mexico, and U.S. Possessions)	$15,000.00	$0.00
2. Foreign	$5,000.00	$0.00
G. Publication Costs	$0.00	$0.00
H. ADP/ Computer Services	$0.00	$0.00
I. Subawards	$0.00	$0.00
J. All Other Direct Costs (Provide supporting data. List items and dollar amounts for each item.)	$0.00	$0.00
K. Total Direct Costs (C through J) →	$20,000.00	$0.00
L. Indirect Costs (Specify rate and base) Rate : 0.00 % Base : $0.00	$2,000.00	$0.00
M. Total Costs (K plus L) →	$22,000.00	$0.00

NOTES:

1. A separate budget is required for each year.
2. Federal Statute (7 U.S.C. 3318 (b)(1)(B)) requires a contribution of resources by all parties toward meeting the objectives of the Cooperative Agreement.
3. The Cooperator's contribution must be no less than 20 percent of the total of the resource contributions under the cooperative agreement. Resource contributions of the Cooperator must consist of a sufficient amount of itemized direct costs to substantiate a true stake in the project as determined by the ADO. The Cooperator's contribution must be maintained at 20 percent of Federal funding throughout the period of performance.
4. Under 7 USC 3319, USDA is prohibited from reimbursing State Cooperative Institutions for indirect costs or tuition remission in connection with non-assistance cooperative agreements awarded under the authority of 7 USC 3318(b).
5. Indirect costs will be reimbursed only upon receipt of a current approved Negotiated Indirect Cost Rate Agreement for all non-State Cooperative Institutions.
6. Unrecovered indirect costs may be used to meet a portion of the resource contribution requirement toward the cooperative effort. Indirect costs only for the resource contribution requirement is not allowed.
7. Unallowable costs as defined in 2 CFR Part 200, Subpart E - Cost Principles, cannot be considered a resource contribution.

REE-454 (10/15)

U.S. DEPARTMENT OF AGRICULTURE
RESEARCH, EDUCATION, AND ECONOMICS

AGREEMENT BUDGET JUSTIFICATION
FOR 58-8042-6-034
Correction

Travel - Domestic

Travel	$15,000.00

Travel - Foreign

Travel	$5,000.00

Indirect Costs

IPSC	$2,000.00