Non-Confidential Final Outcomes Report

Submitted to Emissions Reduction Alberta

By

Agriculture and Agri-Food Canada
on behalf of the project team

October 31, 2024

1. Project information:

1.1. Project ID: B0160957

1.2. Title of project: Demonstrating the impact of genomics-enhanced whole herd genetic management platform on reducing beef greenhouse gas emissions

1.3. Name and contact information of Recipient contact: Agriculture and Agr-Food Canada (AAFC) (Dr. Changxi Li, email: changxi.li@ualberta.ca or changxi.li@agr.gc.ca)

1.4. Name of ERA Project Advisor: Ibironke Popoola (September 1, 2021 to August 4 2023), Erin Crerar (August 4, 2023 to February 8, 2024), Prashant Pandey (February 8, 2024 to August 31, 2024 and post project)

1.5. Completion date of the project: August 31, 2024

1.6. Total project costs: $1,178,999.58; **ERA contributions:** $487,370

1.7. Final outcomes report submission date: October 31, 2024

2. Table of Contents

Contents

2.1. List of Tables

2.2. List of Figures

3. Executive Summary

Agricultural emissions account for approximately 8.5% of Canada's total greenhouse gas (GHG) emissions, with cattle being one of the major contributors within the agricultural sector. Cattle performance, including feed efficiency and GHG emissions, is partially determined by genetics, which includes individual genes passed down from the parents (genetic merit) and gene combinations of the parents that boost in performance (retained hybrid vigour or heterosis). More efficient cattle have been shown to not only consume less feed for the same amount of beef produced but also produce less GHG emissions than inefficient ones. Furthermore, cattle with higher retained hybrid vigour or heterosis have improved production efficiency, leading to less GHG emissions. These genetic variations among animals present an opportunity to reduce GHG emissions through selective breeding. However, beef producers lack easy access to science-based tools to effectively select for better genetic merit and breed cattle with optimized hybrid vigour to improve production efficiency. To provide beef producers these tools, researchers from Agriculture and Agri-Food Canada (AAFC), Alberta Agriculture and Forestry, and the University of Alberta, have developed genomic prediction tools including a genomics enhanced Whole Herd Genetic Management Platform. The project aimed to demonstrate the genomics enhanced Whole Herd Genetic Management Platform to the beef industry through genotyping 10,000 (+/- 10%) beef cattle from participating beef producers. Based on the genotypes, the Platform predicted genetic merit, such as genomic expected progeny difference (gEPD) for traits like growth, feed efficiency, carcass, and fertility, as well as retained hybrid vigour, measured by genomic retained heterozygosity (gRHET). The predicted genomic information was provided to the beef producers through a web portal, which enabled them to maximize the genetic potential of their herds by selecting and managing genetics to improve production efficiency and thus reduce GHG emissions.

The project has successfully completed proposed activities with all objectives and deliverables achieved. Major accomplishments of the project include:

(1). The Emissions Reduction Alberta (ERA) project has successfully recruited 10,427 cattle from 88 beef producers and organizations, of which 10,415 cattle were successfully genotyped on the Bovine GGP-100K SNP panel for downstream genomic analysis (ERA cattle). The number of cattle and producers recruited has exceeded the original plan, with more than 10,000 cattle and 40-60 producers participating. This increase reflects both the strong interest from producers in the genomic tools being offered and the team's dedicated efforts in communicating the tools and their benefits to producers.

(2). The project has successfully developed a very reliable genomic prediction on individual animal's breed composition (gBC) using the refined single nucleotide polymorphisms (SNP) reference data set of 14 pure breed beef cattle along with improved breed composition prediction methods/algorithms including the "small batch" method and "projection analyses" methods. The predicted breed compositions, based on the refined SNP reference panel and the improved methods, show a high correlation (r>0.90) with known animal breed composition or with herd breed composition information submitted by the producers, demonstrating the Platform's reliability from SNP genotyping, SNP data consolidation, genomic analyses, to output of the final results. The high accuracy of the predicted gBC enables reliable estimation of individual animals' retained heterozygosity (gRHET), which is an indicator of retained hybrid vigor.

(3). By leveraging a large, refined data set of over 11,000 beef cattle for feed efficiency and over 7,200 for carcass merit traits and along with a high-density panel of 87,941 SNPs, the team has achieved significant advances in genomic prediction accuracy. Through optimizing genomic prediction methods jointly developed under the co-funding BCRC Genomic Prediction Platform Development project (BCRC-FDE.05.20, PI: Changxi Li, co-PI: John Basarab, 2021-2024) the team has attained moderate to moderately high genomic prediction accuracy for gEPDs on average in the ERA cattle, ranging from 0.41 for RFI to 0.49 DMI. Remarkably, maximum accuracies for gEPDs ranged from 0.41 for AFAT to 0.73 for DMI, with three traits (RFI, ADG, DMI) achieving accuracies of 0.65 or greater, demonstrating a major step forward in genomic selection for feed efficiency and performance.

(4). Building on historical beef cow data and by leveraging new data and resources from the co-funding BCRC Fertility DMI project (BCRC- FDE.06.17, led by Basarab and Li, 2018-2023), the team has developed and refined a genomic prediction reference data set of 81,689 SNPs for female fertility traits, ranging from 1,802 cows with AFC to 2,792 for DUR traits. The refined reference population size for female fertility traits has doubled the size of what was originally planned in the project for most traits. This data set is able to predict 9 female feed intake and fertility traits gEPDs with an average accuracy ranging from 0.26 for AFC to 0.38 for PBWT on the ERA cattle. The maximum accuracy of gEPDs for all female fertility trait has exceeded 0.65, highlighting a strong genetic link between some ERA cattle and the animals in the reference data set. This high level of accuracy underscores the effectiveness of the genomic predictions and the close alignment of industry cattle with the reference population.

(5). In collaboration with co-funding projects of the BCRC Genomic Prediction Platform Development project and BCRC Fertility DMI project, the team has successfully refined a feeder profitability index (FPI) with accuracy from 0.22 to 0.57 (average of 0.38), and a replacement heifer profitability index (RHPI) with an accuracy from 0.22 to 0.78 (average of 0.38) for the ERA cattle. Furthermore, the team developed an RHPI score that combines RHPI and hybrid vigor. The team also introduced a flexible, customizable multiple trait selection index calculation method based on feed efficiency, carcass, and female fertility traits, combined with their economic weights.

(6). Through a complementary project “Gentec Feeder Profit Index validation and deployment” (RDAR-2022T001R, led by Basarab and Li, 2022-2024), the team has validated the genetic gains and economic benefits of the selection index using independent industry data. The validation on the feeder profitability index (FPI) showed that one unit of the index value equaled $1.54 in feeder cattle economic net return, or selection of cattle with an index value of 100 units above the average herd index value could lead to an increase of $154 in economic net return per animal of feeder cattle. The index also has a high and negative correlation (-0.74) with residual feed intake (RFI). Therefore, selection of cattle with FPI above the herd average will result in improvement on feed efficiency, which will lead to production of more efficient cattle with less GHG emissions.

(7). Additionally, through the co-funding BCRC Fertility DMI project, the team has also successfully validated the genetic gain and economic benefits of the replacement heifer profitability index (RHPI) score. The validation showed that heifers with a greater than 48.4 RHPI score produced 9 more calves per 100 replacement heifers exposed to natural breeding, which resulted in an increase of $22,725 net revenue from 100 heifers that were selected based on the RHPI score. Moreover, as the RHPI score puts more weight on the hybrid vigor score selection of heifers with a higher RHPI score will reduce GHG emission as demonstrated by Basarab et al. Animals 2012; 2(2):195-220.

(8). Further to the above genomic profiles, including the FPI and RHPI score, the team has developed a virtual sire mating system that evaluates genetic potential of a bull to produce offspring with greater hybrid vigor score in the herd, reported as the average Herd Progeny Heterosis (AHPH). This tool enables producers to match bulls with heifers/cows based on genotypes to maximize retained hybrid vigor in their herd for improved performance.

(9). The team has successfully refined the Genomics-enhanced Whole Herd Genetic Management Platform under the co-funding BCRC Genomics Platform Development project. This refinement optimized key genomic analysis pipelines including SNP quality control, SNP merging, gBC prediction, gEPD prediction, multiple trait index calculation, the virtual sire mating system and the website portal where producers can submit their animal information and samples, and access genomic profiles results. Using this improved Platform, the current project analyzed 10, 415 genotyped cattle for the ERA project, predicting all genomic profiles as originally planned. The team has released the gBC, gRHET, gEPDs of the feed efficiency, carcass traits and FPI for all cattle, as well as gEPDs for female fertility traits, RHPI, and RHPI score for all heifers/cows, and AHPH for bulls. These results have been shared with the respective producers to support their genetic selection, helping improve feed efficiency, female reproductive performance, profitability, and ultimately reducing GHG emissions.

4. Project Description

4.1 Introduction and background

Agricultural emissions account for approximately 8.5% of the total annual greenhouse gases (GHG) in Canada (Environment and Climate Change Canada, 2018), with beef cattle being one of the major contributors, accounting for 31% of agricultural emissions. Cattles genetics partially controls their performance, including determining the basic or metabolic level of GHG emissions. Variation in GHG emissions among cattle is commonly observed, and this variation is associated with the animals' genetic makeup which affects feed intake, feed efficiency and rumen microbial populations, i.e. an animal's ability to convert consumed feed into saleable meat. Animals that are more feed efficient consume less feed for the same amount of body weight and growth, thus producing less GHG emissions. Residual feed intake (RFI), a common measure of feed efficiency, is the difference between actual and expected feed or dry matter intake required for maintenance and growth (Koch et al. 1963. J. Anim. Sci. 22:486–494). Cattle with low RFI produce 6-28% less methane than inefficient ones (high RFI). This reduction in methane emissions depends on the degree of improvement for RFI (Nkrumah et al. 2006, J Anim Sci. 84, 145–153, Hegarty et al. 2007, J Anim Sci. 85, 1479–1486, Fitzsimons et al. 2013, J Anim Sci 91, 5789–5800, Manafiazar et al. 2020, Can. J. Anim. Sci. 100: 522–535). Therefore, RFI is an indicator trait for genetic selection to reduce GHG emissions. The variation among animals in their genetic makeup comes from genes that they inherit from their parents. In cattle, sires and dams have different sets of genes or gene alleles to pass down to the next generation. Therefore, as breeding stock, they vary in their genetic merit for a particular trait such as feed efficiency or GHG emissions. This genetic merit is commonly referred to as expected progeny difference (EPD), which is the difference between the average performance of offspring from the sire or dam and the average performance of all animals in the population. To improve animal performance, such as feed efficiency and GHG emissions, sires and dams with the best EPDs should be selected as breeding stock to produce the next generation of animals.

Another measure of genetic variation among animals is the level of retained heterozygosity (RHET), which is linearly related to heterosis or retained hybrid vigour. The genetic basis of heterosis is the combination of gene alleles from the two parents. In beef cattle, cross breeding is a strategy used to produce animals with the best combination of gene alleles from the parents to enhance animal performance in addition to using EPDs. Studies have shown that enhanced retained heterozygosity in a cattle herd will not only improve reproductive performance but also contribute to reduction of GHG emissions. A preliminary life cycle greenhouse gas assessment on a cattle herd has shown that total GHG emissions from increasing genomic retained heterozygosity (gRHET) predicted based on DNA markers (i.e. SNPs) were reduced by 639 kg CO2e per youthful animal slaughtered (Basarab et al. Animals 2012, 2, 195-220; Basarab et al. 2019, gEPD Sustainability project report, Genome Alberta).

The effects of EPD and RHET on GHG reduction indicate that selection and breeding for both more efficient and high RHET cattle herds using genomic tools would be one of the most beneficial management practices for reducing the carbon footprint of beef production. However, according to a survey conducted by Weaber et al. (2016, https://digitalcommons.unl.edu/animalscifacpub/908), 41.2% of beef producers were not knowledgeable of methods to select for improved feed efficiency, an economically important and GHG emission related trait that is difficult and expensive-to-measure (i.e. estimated at $250/head to measure feed efficiency for individual animals). In Canada, EPDs are not on the list of the top three criteria for bull selection in the western provinces and it is ranked only the 4th on bull selection in the Atlantic regions (BCRC, 2019, Adoption Rates of Recommended Practices by Cow-Calf Operators in Canada). Furthermore, science-based crossbreeding to optimize RHET is not commonly practiced by commercial beef producers as it requires pedigree information and breeding records of multiple generations. Therefore, development of reliable, accessible and affordable genomic tools is needed to enable beef producers to make more informed genetic selection and breeding decisions.

4.2. Literature review

In beef production, selection of breeding stock with superior genetics represents the first and important step to improve beef production efficiency and quality with reduced environmental footprints. Traditionally, prediction of genetic merit or expected progeny difference (EPD) relies on either animal performance measures or pedigree records or both. However, this approach is not feasible for traits that are expensive or difficult to measures such as feed efficiency, fertility and GHG emissions. Furthermore, it is not always practical for commercial beef producers to track down pedigree information. In 2001, a method named genomic selection was proposed to allow prediction of traits using genome-wide high-density DNA markers (Meuwissen et al., 2001, Genetics 157:1819–1829). Genomic selection or genomic prediction uses a training population of individuals with both phenotypic and DNA marker information to estimate marker effects and subsequently to predict total genetic value of selection candidates based on their marker genotypes (Meuwissen et al., 2001, Genetics 157:1819–1829), therefore it does not require measurement of the trait or pedigree of the candidate animals. Genomic selection has been widely used for the past 10 years in dairy cattle and the rates of genetic gain per year have increased from 50–100% for milk yield traits and from threefold to fourfold for lowly heritable traits such as fertility (García-Ruiza et al. 2016, Proc Natl Acad Sci. 113(28): E3995-4004.). The accelerated rate of genetic improvement is attributed to early selection of breeding stock with increased accuracy of EPDs through incorporating DNA marker information into the process of EPD estimation. In beef cattle, genomic selection on feed efficiency and carcass merit traits have been investigated in Canadian beef cattle (Mujibi et al. 2011, J. Anim. Sci. 89:3353–3361. Chen et al. 2013, J. Anim. Sci. 91:1-10, Chen et al. 2015, Animal Genetics, 46: 55-59, Lu et al 216, J. Anim. Sci. 94(4), 1342-1353), and the results have shown great potential to improve traits that are difficult and expensive to measure such as feed efficiency.

Hybrid vigor or heterosis represents another important component of genetics in beef cattle herds. Hybrid vigor is caused by combining two different gene alleles from the parents and interactions between genes in the genome, often achieved by crossing two distinct breeds. Retained hybrid vigor (RHV) is hybrid vigor that persists in later generation. RHV is calculated as the proportion of retained heterozygosity multiplied by the typical first-generation hybrid vigor for a trait. Traditionally, retained heterozygosity in beef cattle can be calculated based on pedigree-based breed fraction of the sire (PSi) and dam (PDi) of the animal as 1-(sum(PSi x PDi)) over 1 to n breeds (Dickerson,1973, Proc. Animal Breeding and Genetics Symposium, Champaign, IL, USA. pp. 54–77). Research in genomic profiling based on DNA markers has made it possible to probe the ancestry of an individual (Alexander et al. 2009, Genome Res. 19(9):1655-64). This allows for prediction of breed composition (gBC) and thus genomic retained heterozygosity (gRHET) of individual animals when pedigree data and breeding records are unavailable or incomplete (Akanno et al. 2017, Can. J. Anim. Sci. 97: 431–438). This is calculated by using gRHET=1- (sum(Pi*Pi)) over n breeds, where Pi is the fraction of breed i from gBC predicted based on animal DNA markers, and n is the number of breeds contributing to the individual. Unlike gEPDs, gRHET cannot be passed to the next generation. However, it contributes to the performance of the individual over their whole life cycle. A study by Basarab et al. (2018, Can. J. Anim. Sci. 98: 642–655) showed that each 10% increase in gRHET resulted in 51 ± 20 days longer survival in the herd and 35.7 ± 15.2 kg more calf weaning weight per cow exposed to breeding when summed over five parities. At present, a complete set of EPDs (growth, carcass traits, feed efficiency, fertility) for beef cattle are rarely available from breed associations/beef breeders. Although one commercial product, EnVigour HX™, provides estimation of gBC and gRHET, it does not provide EPDs of the breeding stock.

Accurate prediction of the above gEPDs and gRHET in beef cattle depends on large genomic prediction reference data sets of high quality and reliable statistical methods. During the past 25 years feed intake, feed efficiency, and other beef performance data were collected through continuous efforts of researchers from multiple Canadian institutions including Alberta Agriculture and Forestry, Agriculture and Agri-Food Canada (AAFC), and the University of Alberta. These animals were also genotyped on DNA marker (SNP) chips making them useful for genomic prediction studies. Meanwhile, various genomic prediction statistical methods have been investigated based on the reference data sets of Canadian beef cattle (Mujibi et al. 2011, J. Anim. Sci. 89:3353–3361. Chen et al. 2013, J. Anim. Sci. 91:1-10, Chen et al. 2015, Animal Genetics, 46: 55-59, Lu et al 216, J. Anim. Sci. 94(4), 1342-1353). Furthermore, genetic architecture of quantitative traits in beef cattle was revealed by genome wide association studies of imputed whole genome sequence variants for feed efficiency and carcass merit traits in large Canadian beef cattle populations (Zhang et al. 2020, BMC Genomics 21:36; Wang et al. 2020, BMC Genomics 21:38). These Canadian data sets and genomic prediction methods offer a preliminary version of genomic tools that enable prediction of EPDs for a variety of traits including feed efficiency, carcass merit and female fertility traits based on their DNA markers without measuring their phenotype or recording pedigree.

Prediction of genomic breed composition of individual cattle and subsequently gRHET requires a well-defined DNA marker panel of all possible pure breeds used in crossbred beef cattle productions. Through mining cattle sequence data and genotyping animals of purebred beef cattle, a preliminary purebred single nucleotide polymorphisms (SNP) panel was used to predict genomic breed composition for Canadian beef cattle based on cattle genotypes (Akanno et al. 2017, Can. J. Anim. Sci. 97: 431–438; Basarab et al. 2018, Can. J. Anim. Sci. 98: 642–655). However, the reference genotypes for genomic breed composition prediction included a limited number of pure breeds (4 to 6 pure breeds). As a result, many other pure breeds in Canadian beef cattle were still missing in early studies (Akanno et al. 2017, Can. J. Anim. Sci. 97: 431–438; Basarab et al. 2018, Can. J. Anim. Sci. 98: 642–655).

Research outcomes of previous studies have laid a foundation to develop genomic tools. However, the accuracy of genomic prediction on gEPDs and on gRHET still requires improvement before it can be deployed for wider industry application. In addition, genomic prediction is a complex process and it requires in-depth understanding of quantitative genetics/genomics. Accurate genomic prediction involves multiple steps of genomic analyses, which has likely slowed down the implementation of genomic prediction on industry beef cattle. Furthermore, end-users, or beef producers, need a channel to access these genomic tools. Therefore, the overall goal of this project was to develop and refine a genomics-enhanced whole herd genetic management platform (Platform) that integrates the genomic prediction reference data sets, genomic selection tools, genomic analysis pipeline, and end-user website interface, and subsequently to promote the Platform through a demonstration of genomic analysis on more than 10,000 beef cattle from commercial beef producers.

4.3 Detailed technology description

The above proposed genomics-enhanced whole herd genetic management platform consists of 4 key components (1). The genomic prediction reference data sets, (2): Genomic selection tools, (3) Genomic analysis pipeline, and (4): End-user website interface.

4.3.1 Genomic prediction reference data sets

Three beef cattle data sets were consolidated from previous projects and were initially refined under previous research projects for genomic prediction. The first data set included genotypes on 14,698 cattle and phenotypic data from 13,028 to 13, 842 animals on feed intake and growth traits, 8,839 to 9,630 animals on carcass traits. The original data of feed intake, growth, feed efficiency, and carcass merit traits were collected on Canadian beef cattle over the last 25 years by researchers from the Alberta Agriculture and Forestry, Agriculture and Agri-Food Canada (AAFC), and the University of Alberta under multiple previous research projects (e.g. phenomics gap project, sustainable Beef-MBVs RFI carcass project, candidate gene project, Bovine Genome project, and Kinsella Breeding RFI project etc.). The original phenotype data and 50K single nucleotide polymorphisms (SNP) genotypes were consolidated and initially refined for genomic prediction under the project “Development of functional genomic prediction platform for industry application”, Genome Alberta/AAF, A3GP_37 (2018-2021, PI: Changxi Li, co-PI: John Basarab). The refined data set provides a foundation to develop and implement genomic prediction on genetic merit for growth, feed intake, feed efficiency and carcass merit traits. The refined data set (Table 1) was referred to as “Refined beef cattle genomic prediction reference data set for feed efficiency and carcass merit traits (Version 1, 2021)”, and is the subject matter of an AAFC invention disclosure document filed on September 30th, 2021).

Table 1. Summary of current Canadian beef cattle genotype and phenotype data from different populations

Trait/Data set	Angus	Charolais	Kinsella composite	Lacombe crossbreed	Commercial herds	Total
50K SNP	1,776	1,060	2,907	1,682	7,282	14,698
Imputed HD SNP	1,776	1,060	2,906	1,682	7,225	14,649
Imputed WGS SNP	1,776	1,060	2,906	1,682	7,225	14,649
BWT	1,761	1,057	2,917	1,683	5,610	13,028
WWT	1,779	1,056	2,899	1,664	8,444	13,842
ADG	1,761	1,056	2,918	1,675	5,813	13,223
DMI	1,760	1,038	2,901	1,675	5,785	13,159
MWT	1,763	1,038	2,915	1,674	5,789	13,179
RFI	1,760	1,037	2,900	1,674	5,725	13,096
HCW	1,115	707	1,490	550	5,768	9,630
AFAT	1,116	704	1,487	530	5,154	8,991
CREA	1,113	706	1,444	550	5,733	9,546
LMY	1,111	701	1,442	530	5,055	8,839
CMAR	1,117	698	1,488	550	5,595	9,448

SNP: single nucleotide polymorphisms; HD SNPs: high density SNPs (444,558 SNPs); WGS: whole genome sequence SNPs (47,833,012 SNPs including Indel of run 7 from the international 1000 Bull Genomes project); BWT: birth weight, kg; WWT: weaning weight, kg; ADG: average daily gain, kg/day; DMI: dry matter intake, kg/day; MWT: metabolic body weight, kg; RFI: residual feed intake, kg DM/day ; HCW: hot carcass weight, kg; AFAT: average carcass backfat, cm; CREA: carcass rib-eye area, cm2; LMY: lean meat yield. %; CMAR: carcass marbling score, 100 (trace marbling) to 900 (more marbling); Angus, Charolais, Kinsella composite animals were from Roy Berg Kinsella Research Station, University of Alberta. The Lacombe crossbreed animals were from AAFC Lacombe Research Station while the commercial herds were represented by multiple beef cattle populations from Canada.

The second data set included female feed intake and fertility traits as in Table 2. The original data was collected on ~1600 research heifers and cows, ~2000 commercial crossbred heifers, and fertility phenotypes on 12,769 female mating opportunities from the Roy Berg Kinsella Research Station herd (6062 mating opportunities, 2012-2022 breeding seasons), Lacombe Research and Development Centre herd (5613 mating opportunities, 2005-2022 breeding seasons), and 10 commercial cow-calf herds (1104 mating opportunities). This phenotypic data was collected under multiple projects led by Dr. John Basarab (Alberta Agriculture and Forestry/University of Alberta)), and were compiled under the co-funding BCRC Fertility DMI project. These same animals were genotyped on various SNP panels in various projects. The data set (Table 2) is currently referred to as “Cow Fertility and Feed Efficiency Reference Data Set (Version 1 2023)”.

Table 2. Summary of Canadian beef cattle reference data set for female feed intake and fertility related traits

Trait/Data set	Lacombe Crossbreed and Kinsella Composite	Commercial beef cattle	Total
SNP Panel/Replacement heifer trait	Illumina Bovine 50k, GeneSeek 50K, GeneSeek GGP-F250, GeneSeek UltraHigh Density 130k, Affymetrix BOS 1, and GeneSeek Genomic Profiler (GGP) Bovine 100K	GeneSeek Genomic Profiler (GGP) Bovine 100K	2384
BRWT	1291	767	2058
JulianDT	1305	1451	2756
WT200D	1290	1390	2680
DMI	1272	1504	2776
RFIfat	1250	1457	2707
DUR	1288	1504	2792
PBWT	1157	1493	2650
PBBF	1123	1500	2623
AFC	1003	799	1802

BRWT: Birth Weight, kg; JulianDT: Julian date of birth, days; WT200d: 200 days wean weight, kg; DMI: On-test DMI, kg DM/day; RFIfat: On-test RFIfat, i.e. RFI adjusted for off-test backfat thickness, kg DM/day; DUR: On-test Feeding duration, min/day; PBWT: Pre-breeding weight at first parity, kg; PBBF: Pre-breeding backfat at first parity, mm; AFC: Age at first calving, day.

The third data sets included a SNP reference panel of 14 pure breeds including Black Angus, Red Angus, Charolais, Simmental, Hereford, Limousin, Gelbvieh, Salers, Maine Anjou, Shorthorn, Holstein, Brown

Swiss, Jersey, and Galloway (Table 3). The SNP genotypes of the 14 breeds were obtained through mining genotypes and cattle sequence data of multiple previous projects and was further expended and refined under the project "Development of functional genomic prediction platform for industry application", Genome Alberta/AAF, A3GP_37 (2018-2021, PI: Changxi Li, co-PI: John Basarab). The refined SNP reference panel of 14 pure breeds provide a very valuable data set to predict genomic breed composition for Canadian beef cattle. The refined data set (Table 3) was referred to as "Refined beef cattle pure breed SNP reference panel for breed composition prediction (Version 1, 2021)", and is the subject matter of an AAFC invention disclosure document filed on September 30th, 2021.

Table 3. List of 14 Canadian cattle breeds and number of animals with 50K SNP genotypes in the reference population for breed composition prediction

Breed	Number of Animals		Breed	Number of Animals
Black-Angus	1346		Gelbvieh	113
Red-Angus	371		Holstein	321
Charolais	452		Jersey	159
Hereford	657		Limousin	118
Shorthorn	490		Maine-Anjou	105
Brown-Swiss	130		Salers	113
Galloway	100		Simmental	244

The above three genomic prediction reference data sets are governed by a Data Sharing Reciprocal Agreement between AAFC and the University of Alberta signed in August 2024.

4.3.2. Genomic selection tools

Main genomic selection tools include genetic merit or gEPDs of single beef performance traits as listed Table 1 and Table 2. Other genomic tools include the genomic breed composition (gBC) and gRHET of individual cattle, a multiple trait feeder profitability index (FPI), and a replacement heifer profitability index (RHPI).

Genetic merit or a gEPD for a single trait can be predicted using the genomic prediction reference data sets and optimal statistical methods (Mujibi et al. 2011, J. Anim. Sci. 89:3353–3361. Chen et al. 2013, J. Anim. Sci. 91:1-10, Chen et al. 2015, Animal Genetics, 46: 55-59, Lu et al 216, J. Anim. Sci. 94(4), 1342-1353). Prediction of genomic breed composition (gBC) can be conducted using the refined pure breed SNP panel that includes 14 Canadian cattle breeds listed in Table 3. As each pure breed has a genome signature of SNP genotype profiles, the exact percentage of the genome that each beef animal inherited from those 14 breeds can be determined by comparing the genome signature of those purebreds to the animal genome, allowing more accurate prediction of gBC (Alexander et al. 2009, Genome Res. 19(9):1655-64). The gBC can then be used to predict gRHET, a key determinant of hybrid vigour or heterosis (Dickerson et al. 1973, Proc. Animal Breeding and Genetics Symposium, Champaign, IL, USA. pp. 54–77; Akanno et al. 2017, Can. J. Anim. Sci. 97: 431–438).

A preliminary version of the multiple trait feeder profitability selection index (FPI) was developed by multiple previous projects including the "Kinsella breeding project" (2013-2018) led by Dr. Graham Plastow from the University of Alberta with the development of FPI led by Dr. Changxi Li, and then further refined under the “MBVs/gEPDs for feed efficiency and carcass trait project” (2015-2019) led by Basarab and Li et al. The FPI was developed to select bulls to improve feed efficiency and profits of feeder cattle, and it included 6 traits:

FPI = b1*RFI_mbv + b2*ADG_mbv + b3*HCW_mbv + b4*AFAT_mbv + b5*LMY_mbv + b6*CMAR_mbv

A preliminary version of the replacement heifer profitability index (RHPI) was developed by the co-funding BCRC Fertility DMI project. The initial RHPI included 9 traits:

RHPI = b1* JulianDT _mbv + b2* BRWT _mbv + b3* WT200d _mbv + b4* DMI _mbv + b5* RFIfat _mbv + b6* DUR _mbv + b7* PBWT _mbv + b8* PBBF _mbv + b9* AFC _mbv

To further maximize both gEPDs and gRHET in a herd, a mate selection analysis is conducted to select bulls from the herd and/or from commercial breeders that would produce offspring with optimal gRHET in the herd. This mate selection tool allows producers to maximize hybrid vigor in the cattle herd, leading to improved cattle performance and reduced GHG emissions.

4.3.3. Genomic analysis pipeline

The genomic analysis pipeline integrates genomic prediction of multiple modules. This includes SNP data quality check (module 1), genomic breed composition and genomic retained heterozygosity prediction (module 2), genomic prediction for gEPD (module 3), multiple trait heifer profit index calculation (module 4), and multiple trait feedlot profit index calculation (module 5). This enables a data analyst to run all these modules under a window interface as shown in Figure 1. This genomic analysis pipeline makes the complex steps involved in genomic prediction a feasible process for any data analysts or industry genomic service provider to conduct genomic analyses so long as they provide a reference data set and animal SNP genotypes (Figure 1).

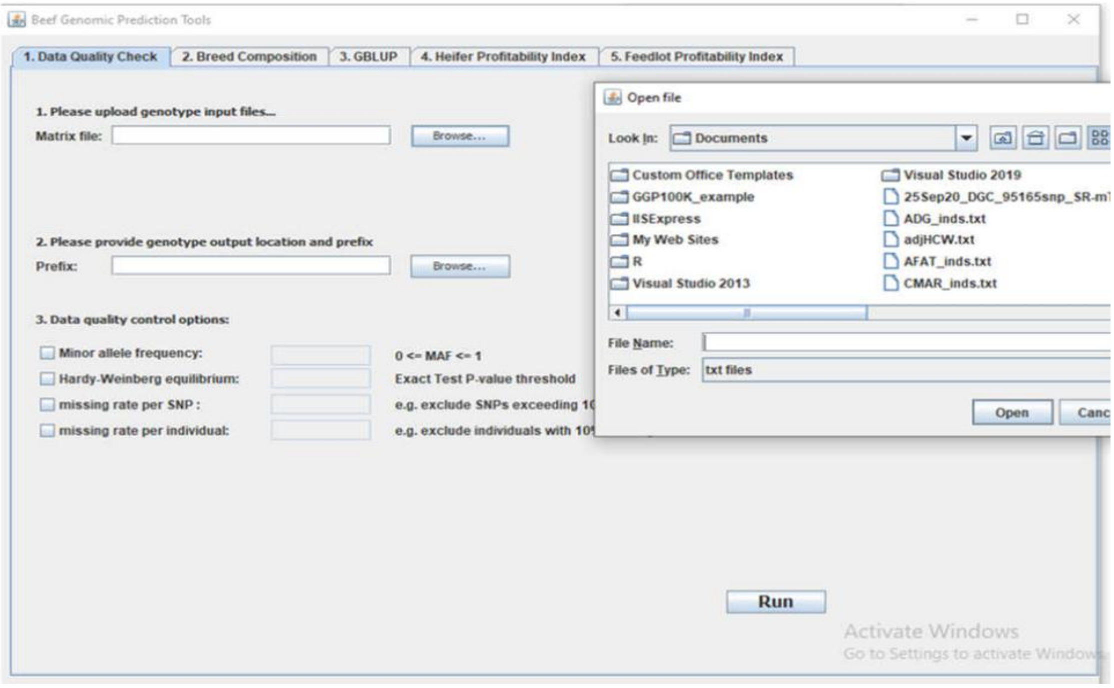
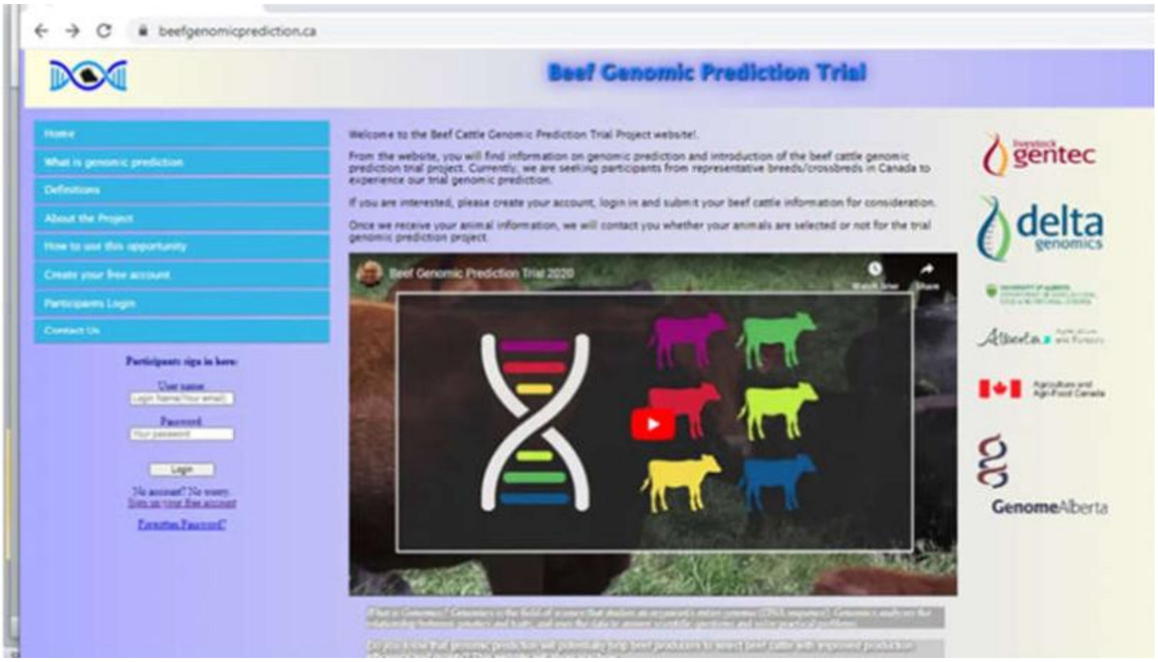


Figure 1. A screen shot to show a genomic analysis pipeline that allows genomic prediction of multiple modules under a window interface.

4.3.4. End-user website interface.

To make the genomic tool more accessible to the beef industry, an end-user website interface was also developed. The website portal (https://www.beefgenomicprediction.ca/) is for industry end users (e.g. beef producers) to provide inputs on their herd and to receive reports. It has six functions/tabs including (1): End user account, (2(: End user animal submission, (3): End user defines selection goals (i.e. desired traits to improve), (4): End user receives herd genomic profiles gEPDs and gRHET, (5): End user receives animal ranking on their gEPDs and gRHET, and (6): End user receives predicted genetic gain or single $ index of each candidate animal, and recommendation to improve gEPDs and gRHET for their herds. To use the genomic tools, end users only need to provide materials for SNP genotyping or existing SNP genotypes of 50K or greater, and traits they want to improve via the industry website interface. A screen shot of the user interface for industry users to create an account and a screen shot of the user interface for industry users to submit basic animal information and to view and receive genomic prediction results are presented in Figure 2 and Figure 3, respectively.



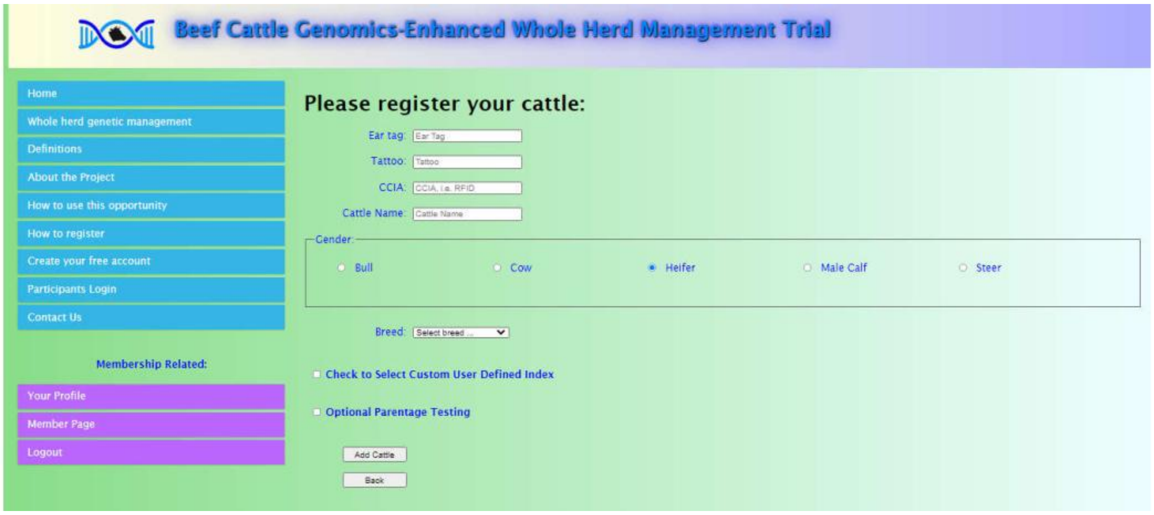
Welcome to the Beef Cattle Genomic Prediction Trial Project website!.

From the website, you will find information on genomic prediction and introduction of the beef cattle genomic prediction trial project. Currently, we are seeking participants from representative breeds/crossbreds in Canada to experience our trial genomic prediction.

If you are interested, please create your account, login in and submit your beef cattle information for consideration.

Once we receive your animal information, we will contact you whether your animals are selected or not for the trial genomic prediction project.

Figure 2. A screen shot of the user interface for industry users to create an account.



Please register your cattle:

Ear tag: Ear Tag
Tattoo: Tattoo
CCIA: CCIA, i.e. RFID
Cattle Name: Cattle Name

Gender:
○ Bull ○ Cow ● Heifer ○ Male Calf ○ Steer

Breed: Select breed ...

☐ Check to Select Custom User Defined Index

☐ Optional Parentage Testing

Figure 3. A screen shot of the user interface for industry users to submit basic animal information and to view and receive genomic prediction results.

4.3.5. A genomics-enhanced Whole Herd Genetic Management Platform

The proposed genomics-enhanced Whole Herd Genetic Management Platform integrates the industry end-user website interface, the genomic prediction reference data sets, the genomic analysis pipeline, and the genomic selection tools into one Platform. A framework or workflow of a genomics-enhanced Whole Herd Genetic Management Platform is presented in Figure 4.

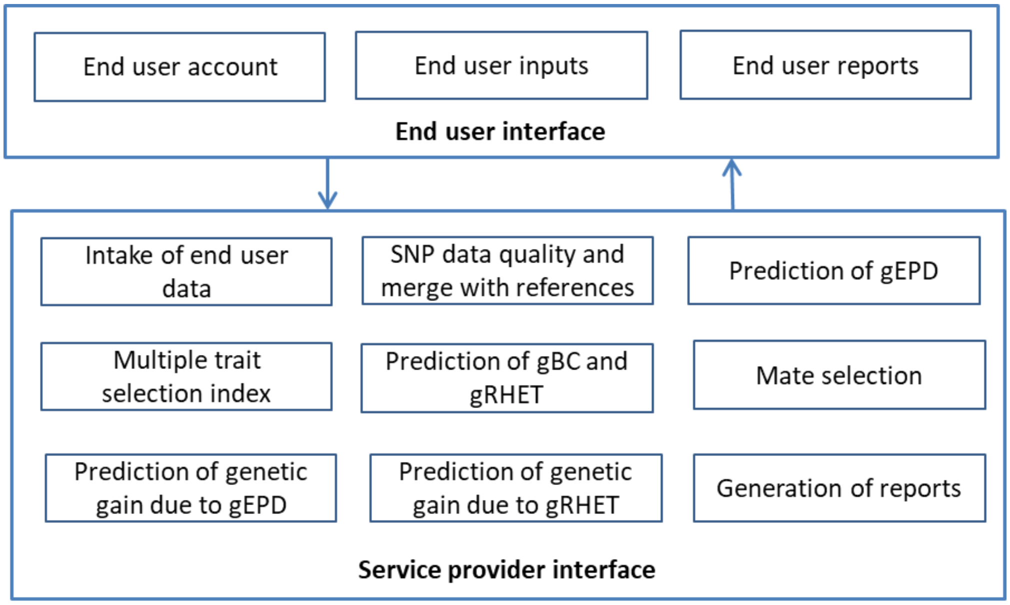


Figure 4. A framework or workflow of a genomics-enhanced Whole Herd Genetic Management Platform. SNP: single nucleotide polymorphisms; gBC: genomic breed composition of the animal; gEPD: genomic expected progeny difference; gRHET: genomic retained heterozygosity.

The website portal of the Platform (i.e. the End user interface) serves as a channel to communicate with the industry end-users and for them to access the genomic tools. The service provider interface is for data analysts to take animal information from the industry end users, combine with the DNA genotypes, merge with the genomic prediction reference data sets, perform various genomic prediction analyses, and release genomic prediction results to the industry end users. It has nine functions/tabs including (1): Intake of end user data, (2): SNP data quality control and merging with reference data, (3): Prediction of gEPD (genomic expected progeny difference), (4): Calculation of multiple trait selection index, (5): Prediction of gBC (genomic breed composition) and gRHET (genomic retained heterozygosity); (6): Mate selection to optimize gEPDs and gRHET, (7): Prediction of genetic gain due to gEPD selection, (8): Prediction of gain due to gRHET increase, and (9): Generation of reports including a single $ index for each candidate animal and recommendation on genetic selection to improve gEPDs and gRHET. The computational process under each function is controlled from a window-based interface run on a local workstation. Prediction of gEPDs can be conducted for a variety of traits including growth, feed efficiency and female fertility traits as listed in Table 1 and Table 2. The predicted gEPDs for the traits can also be provided through single trait or multiple trait selection indexes, enabling producers to make genetic selection based on their breeding objective.

The development of the genomics-enhanced Whole Herd Genetic Management Platform started with the support by the AAFC A-base peer-reviewed project “Development and application of functional genomic prediction for feed efficiency and carcass traits in beef cattle” (funded by AAFC A-base fund, 2016-2019, PI: Dr. Changxi Li, AAFC Project ID: J-001274) and further by project “Development of functional genomic prediction platform for industry application” (funded by Genome Alberta and Alberta Agriculture and Forestry, 2019-2021, PI: Changxi Li, co-PI: John Basarab, AAFC Project ID: J-002320). The genomics-enhanced Whole Herd Genetic Management Platform has been referred to as “Beef cattle genomics enhanced whole herd genetic management platform" and it is the subject matter of an AAFC invention disclosure document filed on September 30th, 2021,

4.4 Project objectives (including the objectives from original contribution agreement and any evolution/revisions made over the course of the project)

The overall objectives of the project were to genotype 10,000 (+/-10%) cattle from beef producers and to demonstrate the application of the genomics-enhanced Whole Herd Genetic Management Platform in generating genomic values on genetic merit, hybrid vigor, multiple trait selection indexes, i.e. genomic selection tools, based on cattle DNA marker genotypes to assist beef producers with their genetic selection and breeding to improve profitability and reduce GHG emissions. To achieve maximum accuracy of predicted genomic profiles and deliver a valuable Platform, the project proposed to further refine the genomic prediction tools and the Platform described above through collaboration with the co-funding BCRC Fertility DMI project, the co-funding BCRC Genomics Platform Development project and other complementary projects. Specifically, the objectives included: (1): recruiting and training beef producers on the application of the genomics-enhanced Whole Herd Genetic Management Platform; (2): characterization of whole herd genomic profiles for producers' herd with advanced genomic prediction tools; (3): construction of flexible multiple trait selection index for producers to make a genetic selection; (4): assignment of mates in the herd to maximize both gEPDs and retained hybrid vigour, i.e. genomic retained heterozygosity (gRHET); (5): conducting economic analysis for producers resulting from the application of the Whole Herd Genetic Management Platform (6): Further refine the genomics-enhanced Whole Herd Genetic Management Platform; (7): Deliver decision making tools to participating producers on 10,000 (+/- 10%) beef cattle, analyze feedback and prepare for commercialization of the Platform. All the project activities were conducted in Alberta with core genomic analyses performed at Livestock Gentec, University of Alberta where the AAFC genomic researchers are co-located.

4.5. Work scope overview

To achieve the objectives and deliverables, this ERA project has utilized the above three genomic prediction reference data sets as in Table 1, Table 2 and Table 3 that were consolidated and preliminarily refined under previous projects. The project has also leveraged additional beef cattle data and resources from two co-funding projects "Genetic analyses of feed intake, feed efficiency, female fertility, and cow lifetime productivity in beef cattle raised under two environments" (BCRC-FDE.06.17, led by Basarab and Li, 2018-2023, referred to as the BCRC Fertility DMI project), and "Development and demonstration of a genomics-enhanced whole herd genetic management platform to improve beef production efficiency and quality (BCRC-FDE.05.20, PI: Changxi Li, co-PI: John Basarab, 2021-2024, referred to as the BCRC Genomic Prediction Platform project", and from complementary projects including "Gentec Feeder Profit Index validation and deployment" (RDAR-2022T001R, led by Basarab and Li, 2022-2024, referred to as RDAR Feeder Profit Index validation project). These co-funding and complementary projects and the current ERA project shared similar goals of developing genomic tools for Canadian beef cattle. The co-funding BCRC Fertility DMI project focused on data collection of female feed intake and fertility traits, generation of economic weights of fertility related traits for replacement heifer profitability index construction, and validation of the genomic prediction results including replacement heifer selection tools using independent industry data sets. The co-funding BCRC Genomic Prediction Platform project was centred on refining genomic prediction tools. This included exploring methods to enhance genomic prediction accuracy, developing and refining multiple trait selection indexes and genomic prediction pipelines, and finally developing a genomics-enhanced whole herd genetic management platform (referred to as the Platform) that integrates the web-site portal for industry access and the refined genomic prediction pipelines for routine genomic prediction services. The RDAR Feeder Profit Index validation project focused on data collection from industry partners and validation of the genomic prediction results from the Platform project using independent industry feeder cattle performance data. The current ERA project aimed at recruiting 10,000 beef cattle from producers, genotyping the cattle, analyzing their genomic profiles using the Platform and releasing the genomic profiles to the respective producers. In the following sections of work scope, activities, and results achieved by joint efforts of multiple projects will be specifically indicated.

Activity 1: Recruiting and training beef producers on the application of the genomics-enhanced Whole Herd Genetic Management Platform

This activity was conducted by the ERA project in close collaboration with the co-funding BCRC Genomic Prediction Platform project and with assistance from the Livestock Gentec (University of Alberta) technology transfer team. The joint efforts aimed to recruit 10,000 beef cattle (bulls, replacement heifers/cows, 5000 animals per year for the first two years of the project) from ~50 beef producers, genotype the animals on a 100K SNP panel for subsequent characterization of whole herd genomic profiles using the Platform. To support recruitment, beef producers were trained through in-person or virtual meetings, as well as through recorded videos and pamphlets made available on the project website (https://www.beefgenomicprediction.ca/). The team also engaged producers by attending industry meetings to further promote the project and its benefits. The training focused on knowledge and understanding of gEPDs, gBC, gHRET, and how producers can apply these genomic tools to improve genetics of their beef herds. The training also included how to submit requests via the website and how to collect and send animal tissues for genotyping, how to receive result reports through the website, and economic benefits or returns of using the platform. Upon receiving animal tissues (hair, ear tag, semen, etc), the team coordinated genotyping of these animals on a 100K SNP chip with Neogen Canada, a SNP genotyping service provider. Producers paid $15 per animal for genotyping while the remaining proportion of genotyping cost ($30) was shared by this ERA Genomics Platform Demonstration project (~9600 cattle) and by the co-funding BCRC Genomic Prediction Platform project (N=400 cattle).

Activity 2: Characterization of whole herd genomic profiles for producers' herds with advanced genomic prediction tools

Activity 2.1. Animal populations and phenotype data consolidation:

Genotype and phenotypic genomic prediction reference data sets consisted of >13,000 animals on feed efficiency, ~9,000 animals on carcass, >1,800 animals on female feed intake and fertility related traits, and 50K genotypes of 4,719 cattle of 14 Canadian cattle pure breeds as in Table 1, Table 2, and Table 3. The team also obtained a data set of 520 animals with genotypes and GHG emission traits. The GHG data set was collected and consolidated under multiple previous projects including "Methane emissions from beef cattle bred for low residual feed intake, 2014R073R, 2014-2017, led by John Basarab) and "gGreenBeefCow" project (2016R033R, 2016 to 2019 led by Carolyn Fitzsimmons).

Activity 2.2. Enhance genomic prediction on gEPDs for industry animals:

The above genomic prediction reference data sets were used to explore appropriate methods to enhance genomic prediction accuracy on genetic merit, i.e. expected progeny difference (EPD) for the above traits in Canadian beef cattle, focusing on traits that have a lower heritability or lower genomic prediction accuracy. This activity was conducted by the co-funding BCRC Genomic Prediction Platform project.

The original phenotypic values were adjusted using a generalized linear model (GLM) function in R Statistical Software (v4.1.2; R Core Team 2021). The GLM function was used to quantify the effects or regression coefficients of breed fractions in the animal as well as other covariates or fixed effects. The breed fraction of each animal was predicted using Admixture software (Alexander et al. 2009. Genome Res, 19: 1655-1664.) and a SNP reference panel of 14 pure breeds developed by the team, which includes 4,721 cattle from Black Angus, Red Angus, Charolais, Simmental, Hereford, Limousin, Gelbvieh, Salers, Maine Anjou, Shorthorn, Holstein, Brown Swiss, Jersey, and Galloway. Other covariates or fixed effects included all identifiable contemporary groups that were defined based on herd, birth year, location, feedlot pen, animal age, levels of management system, etc. depending on the trait. The adjusted phenotype value of the trait for each animal was then obtained by subtracting the population mean and the sum of breed fraction, covariate and fixed factors multiplied by their respective effects or regression coefficients from the original phenotype value, leading to a phenotype data set that were accounted for identifiable non-genetic effects. Subsequently, adjusted phenotype values that were outside the range of 3 standard deviations of the mean were removed.

The SNP genotype data of the “Refined beef cattle genomic prediction reference data set for feed efficiency and carcass merit traits (Version 1, 2021)” (Table 1) were refined to 50K and imputed to High density SNP (428,895 SNPs) under a previous project (i.e. Genome Alberta/AAF, A3GP_37 (2018-2021)). The SNP genotype data from the various panels in Table 2 were called by this project on four different formats, namely top strand, forward strand, design strand, and AB formats based on the files from the genotyping labs. Furthermore, to ensure that SNP genotypes from different panels were merged correctly, SNPs were merged in the same format and alleles of each SNP were examined after merging to ensure that each SNP had only two alleles as expected. SNP quality control was then conducted where SNPs with a minor allele frequency less than 0.05, a missing rate larger than 0.10, or a call rate less than 90% were excluded. After merging, SNP genotypes were further imputed using Beagle 5.4 (Browning. et al. Am J Hum Genet, 108 (2021) 1880-1890.) with a multibreed reference of 3,093 cattle and 140K SNPs that were developed based on actual genotypes by the team. Quality control of the imputed SNP dataset was carried out using PLINK (v1.07) (Purcell et al., Am J Hum Genet, 81 (2007) 559-575), where SNPs that were significantly deviated from a Hardy-Weinberg equilibrium test (HWT) (P< $10^{-}6$) or on sex chromosomes were excluded from further analyses.

Genomic prediction for trait gEPD was conducted using a commonly used single trait genomic Best Linear Unbiased Prediction (GBLUP). The statistical model for GBLUP used in the project can be written as:

$$\mathbf{y}^* = \mathbf{1}\mu + \mathbf{Z}\mathbf{a} + \mathbf{e},$$

Where y* is the vector of adjusted phenotype values from animals in the reference population, while for candidate animals the phenotype values are not available, μ is the overall mean, a is the vector of breeding values with $N(0, G\ a_a^2)$, and **Z** is the incidence matrix relating a to y*. e is the vector of random residuals. The additive genomic relationship matrix ***G*** among animals was derived from the SNP markers using the second method of VanRaden (VanRaden, Journal of Dairy Science, 91 (2008) 4414-4423) based on the SNP genotypes, i.e.

$$\boldsymbol{G} = \frac{(\boldsymbol{M}-\boldsymbol{P})(\boldsymbol{M}-\boldsymbol{P})'}{2\sum \boldsymbol{p_i}(\boldsymbol{1}-\boldsymbol{p_i})}$$

where M means the (0, 1, 2)-encoded genotypic matrix for genotypes A1A1, A1A2, and A2A2, respectively, p_i is the minor allele frequency (MAF) of marker i, and $\boldsymbol{P}$ is a matrix with elements equal to $2p_i$. For genomic prediction, SNP genotypes of candidate animals were merged in the same format with SNP genotypes of the reference population, and the combined SNP genotypes were used to generate the G matrix, which leads to predict of molecular breeding values (MBV) or genomic estimated breeding values (GEBV). Subsequently, gEPD were calculated as MBV or GEBV divided by 0.5 for all cattle including the candidate animals. The accuracy of MBV/GEBV or gEPD was obtained using the variance of prediction errors on the GEBV (MBV) of the animal and its additive genetic variance from the GBLUP analysis, i.e., accuracy of GEBV=square root of (1- (variance of prediction errors divided by additive genetic variance)), where the additive genetic variance was corrected for inbreeding coefficient of the animal from the G matrix as proposed by Gorjanc et al. (Genetics Selection Evolution (2015) 47:65).

In addition to the above single trait GBLUP, the study also investigated the performance of multiple trait GBLUP, and machine learning methods including STKRR model, MAK_MTKRR model, STLinearSVR model, and MAK_MTLinearSVR model for traits that have low heritability or have a low accuracy from the single trait GBLUP. In the multiple GBLUP analysis, a trait that has the highest genetic correlation with the target trait was selected as an assistant trait to conduct a two-trait GBLUP. The machine learning STKRR model is single trait Kernel ridge regression (KRR) proposed by Saunders et al. 1998 (In Proceedings of the Fifteenth International Conference on Machine Learning (ICML'98). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 515-521.). In KRR, the kernel trick ($\phi(\mathrm{x_i})$) involves replacing the inner product in the original (genotype) space with an inner product in a more complex feature space and then builds the ridge regression model (RR) model in the latent feature space. The STKRR analysis was conducted based on the procedures described by Saunders et al. 1998. The MAK_MTKRR model is a MAK-based multiple traits framework using KRR algorithm. MAK is a machine learning framework that improves genomic prediction via multi-target ensemble of regressor chains (ERC) as proposed by Spyromitros-Xioufis et al. (Machine Learning, 104 (2016) 55-98), and automatic selection of assistant traits as proposed by Liang et al. (Brief Bioinform, 24 (2023)). The STLinearSVR model is a single trait linear support vector regression (SVR) that employs a deep learning mechanism as proposed by Cortes and Vapnik (Machine Learning, 20 (1995) 273-297). The MAK_MTLinearSVR model is a MAK-based multiple traits framework using LinearSVR algorithm and employs the LinearSVR model to perform the multiple trait genomic prediction. Under this project, the single and multiple trait GBLUP were performed using ASREML (v4.1) (Gilmour et al, Asreml User Guide Release 3.0, VSN International Ltd, Hemel Hemptead, HP1 1ES, UK, (2009)). For STKRR model, MAK_MTKRR model, STLinearSVR model, and MAK_MTLinearSVR, the genomic prediction was conducted using the scikit-learn and mak package (Liang et al. 2023, Brief Bioinform, 242) in Python 3.12. For comparison of various methods, the accuracy of genomic prediction was obtained using a 5-fold cross validation procedure, where individuals in the data set were randomly divided into 5 groups. One of the groups was taken as the validation group whose breeding values we wish to predict each time, assuming no phenotypic data, and the remaining 4 groups with known phenotypes were used as a reference dataset. The prediction accuracy is defined as the average Pearson correlation coefficient between the adjusted phenotypic values and genomic estimated breeding values (GEBVs) in the validation group divided by square root of trait heritability that was obtained by the GBLUP. Mathematically it can be expressed as prediction accuracy= $cor(y, gebv)/\sqrt{h^2}$, where y is the vector of the adjusted phenotypes, gebv is the vector of the GEBVs and h^2 is the heritability of the trait, as described in our previous studies (Chen et al. 2013, J. Anim. Sci. 2013.91:4669–4678; Chen et al. 2015, Animal Genetics, 46, 55–59). Furthermore, in collaboration with the complementary RDAR Feeder Profit Index validation project, a validation of trait gEPDs was performed using 4,336 industry animals with both genotypes and phenotypes that were collected by Dr. John Basarab under the previous "Development and deployment of MBVs/gEPDs for feed efficiency and carcass traits that perform in commercial beef cattle" funded by Genome Alberta and Alberta Meat and Livestock Agency (2015-2020; RES0028243) and under the RDAR Feeder Profit Index validation project. In the validation study, the trait GEBVs of the 4,336 industry cattle were predicted solely based on their SNP genotypes using the single trait GBLUP by this project, and the prediction accuracy was assessed through a regression analysis of trait GEBVs and the actually measured phenotype values of the trait.

Activity 3. Construction of flexible multiple trait selection index for producers to make a genetic selection to reduce GHG emissions

This activity was performed jointly by the co-funding the BCRC Genomic Prediction Platform project and co-funding BCRC Fertility DMI project and was in collaboration with the complementary RDAR Feeder Profit Index validation project.

Building on preliminary versions of two multiple trait selection indexes, namely feedlot profitability index (FPI) and replacement heifer profitability index (RHPI), the FPI was refined to put more economic weights for feed efficiency traits to improve carcass net returns and reduce greenhouse gas (GHG) emissions. Six traits were considered in the index:

FPI = b1*RFI_mbv + b2*ADG_mbv + b3*HCW_mbv + b4*AFAT_mbv + b5*LMY_mbv + b6*CMAR_mbv

For RHPI, as DMI and JulianDT had a high genetic correlation with RFIfat (0.81) and AFC (-0.97), respectively, these two traits were removed for the index. Therefore, the refined RHPI included 7 traits:

RHPI = b1* BRWT_mbv + b2* WT200D_mbv + b3* RFIfat_mbv + b4* DUR_mbv + b5* PBWT_mbv + b6* PBBF_mbv + b7* AFC_mbv

In the above FPI and RHPI, the bi values represent the weighting factor of the ith trait to obtain the single index value of an animal; mbvs represent molecular breeding values (MBV or GEBV) of the trait and the trait definition for the abbreviations can be found in Table 1 and Table 2.

In the traditional index theory, the b values are derived based on trait economic weights and (co) variance among the traits as proposed by Hazel (Genetics, 1943 Nov;28(6):476-90), i.e.

$$b = P^{-1}Gv$$

Where b vector represents the b values that are used to calculate the index value, P and G are phenotypic and genetic (co) variance among the traits, and *v* is the vector the economic weights of the traits. However, this method does not consider the accuracy of trait GEBV (or MBV) of each animal. Therefore, the b values are constant across animals.

Under this project, an important development of FPI and RHPI is derivation of the b values for each individual animal, considering the variation of genomic prediction accuracies of molecular breeding values (MBVs) or genomic estimated breeding values (GEBV) among animals due to their different genetic relationships to the reference population as proposed by Dekkers (J. Anim. Breed. Genet. 124 (2007) 331–341). As the calculation of the b values considers the variation of accuracy of GEBV among the animals, each animal has a unique set of b values for its index calculation. Subsequently, a pipeline was developed under this project to enable calculation of the above two indexes and a flexible multiple trait selection index that allows users to adjust trait economic weights in v and/or to remove one or more traits from the full list of traits in the multiple trait selection index depending on breeding objectives.

Activity 4. Assignment of mates in the herd to maximize both gEPDs and retained hybrid vigour

Activity 4. 1. Delineate molecular basis of retained hybrid vigour

The activity was conducted by the co-funded BCRC Genomic Prediction Platform project. The project used the above two reference data sets (Table 1 and Table 2) to delineate the molecular basis of retained hybrid vigour (RHV). The original SNP genotypes were combined in TOP format and each SNP's chromosome and position were determined based on bovine ARS-UCD1.2 genome assembly. To facilitate subsequent imputation analyses, the SNPs were converted to the PLUS format using a conversion table prepared by the team. Subsequently, imputation was performed on 29 autosomes using the software Beagle 5.4 (Browning. et al. Am J Hum Genet, 108 (2021) 1880-1890.) in a two-step procedure, (1) from the original SNP genotypes to the Affymetrix Axiom Genome-Wide BOS 1 High Density (HD, 444,558 SNPs) Array (Affymetrix, Inc., Santa Clara) using a reference population of 4,247 animals of beef mixed breeds; (2) from imputed HD to the full whole-genome sequence (WGS) variants in Run 8 of the 1000 Bull Genomes Project. The reference population of the 1000 Bull Genome project run 8 consists of 4,109 animals in the Taurus-only run and the average read fold coverage is 12.7x. The resultant imputed whole genome sequence (WGS) dataset of the reference populations contained a total of raw 69,980,033 (~70 million) DNA variants. Additional quality control (QC) procedures were applied to the imputed WGS variants, including removal of variants with minor allele frequency (MAF) less than 0.05, and removal of SNPs significantly deviated from the exact test of Hardy-Weinberg equilibrium (HWE) (P-value < 10^{-5}). After the QC, 22,489,695 WGS variants remained. For this study, we used SnpEff (Cingolani et al. 2012 Fly (Austin). 6(2):80-92) to annotate and predict the effects on the imputed 22.5 million WGS variants, and WGS variants that did not have predicted effects or functions from SnpEff were discarded. This yielded a refined WGS variant dataset of 13,134,057 and 16,641,519 DNA variants (SNV and INDEL) for the feed efficiency and carcass data set, and for the female feed intake and fertility data set, respectively. The DNA variants were further annotated to 9 functional classes, which included (1) intergenic region, fc1, (2) Downstream gene variant, fc2, (3) upstream gene variant, fc3, (4) synonymous variant, fc4, (5) intron variant, fc5, (6) missense variant, fc6, (7) 3'UTR variant, fc7, (8) 5'UTR variant, fc8, (9) Other regulatory fc9. Various indicators of heterosis or retained hybrid vigour were obtained for each animal and were evaluated for their association with the beef performance traits.

Activity 4.2. Develop mate selection strategies to maximize both gEPDs and gRHET

This activity was conducted jointly by this ERA project and the co-funding BCRC Genomic Prediction Platform project. Under the joint efforts, a mate selection tool has been developed through a virtual mating system, in which each of the bulls in the herd is virtually mated with each cow/heifer in the herd. For each virtual mating, the expected gRHET of the offspring was calculated as: offspring gRHET= 1- ∑ P_sire_i *P_dam_i; where P_sire_i and P_dam_i were the genomic breed composition fraction (gBC) of breed i of the sire and of the dam, respectively. The gBC was predicted as described above. Assuming each virtual mating produces one offspring, the average offspring gRHET of each bull in the herd, i.e. average Herd Progeny Heterosis (AHPH), was calculated by taking the average of offspring gRHET from the virtual mating of the bull with all the cows/heifers in the herd. A system has been developed under this ERA project to rank bulls based on their AHPH and gEPDs of the trait or indexes. Bulls that have both greater trait gEPDs or index values and greater AHPH are the top bulls that are recommended to the producers to maximize both EPDs and RHV in the herds.

Activity 5. Predict genetic gains for producers resulted from the application of the Whole Herd Genetic Management Platform

This activity was performed by the co-funding BCRC Genomic Prediction Platform project and co-funding BCRC Fertility DMI project and was in collaboration with the complementary RDAR Feeder Profit Index validation project. To evaluate economic benefits of genetic selection on FPI the feed intake, feed efficiency and carcass trait gEPDs and FPI were predicted by GBLUP. This was based solely on genotypes of 137 bulls with 1,894 crossbred feeder progenies with records for carcass quality from the RDAR Feeder Profit Index validation project. Female feed intake, feed efficiency, fertility trait gEPDs and replacement heifer profitability index (RHPI) and RHPI score were also predicted by GBLUP based on the genotypes of 1,544 commercial crossbred heifers from 10 Alberta herds collected by the BCRC Fertility DMI project. For validation, the industry data of the above 137 bulls and 1,544 commercial crossbred heifers were not included in the genomic prediction reference data sets. The economic benefits of both FPI and RHPI score selections were analyzed independently through a PROC MIXED in SAS (Version 9.4, SAS Institute Inc, Cary, NC) where the economic return was the dependant variable and FPI or RHPI score was included as independent variables.

Based on the economic benefits of the index selection obtained through the above validation, the cost-benefit analysis of using the developed Genomics-enhanced Whole Herd Genetic Management Platform was conducted by this project. The cost-benefit analysis considered two scenarios via simulation. The two scenarios included (1): a combined cow-calf and feedlot operation, which would incur all costs and revenue associated with genotyping; and (2): a separate cow-calf and feedlot operations, where the cow-calf operation would cover all genotyping costs and obtain a proportion of additional revenue due to genomic selection from the feedlot operation through marketing up their weaned calves. Two revenue splits were also considered, one in which the cow-calf operation received 25% of the additional revenue and another where the cow-calf operation received 50% of the additional revenue. The cost-benefit analysis considered inputs and outputs of a typical beef production system plus genotyping costs and additional economic benefits of genomic analyses and genomic selection. The cost-benefit analysis was also conducted based on observed genomic profiles of animals from each user whose cattle were analyzed through this ERA project.

Activity 6. Further refine the genomics-enhanced Whole Herd Genetic Management Platform

This activity was carried out by the co-funding BCRC Genomic Prediction Platform project. Based on the previous version of the Platform as shown in Figure 1, Figure 2, Figure 3, and Figure 4. Based on feedbacks from industry participants and collaborators. The Platform was further refined to include two interfaces. The first interface (End user interface) was designed for industry end users (e.g. beef producers) to create an account online, provide animal information, submit animals for genomic analysis, and to access their animal's genomic prediction results. It also helps users understand how to utilize these genomic prediction results to optimize herd genetics through genetic selection and breeding. The "End user interface" has six functions/tabs including account creation, animal submission, setting selection goals (i.e. desired traits to improve), viewing herd genomic profiles, viewing animal ranking, and accessing predicted genetic gain or their single $ index of each candidate animal. The End use interface has been integrated into the website (https://www.beefgenomicprediction.ca/) as a portal. The portal was developed with PHP, HTML, CSS and integrates with the MySQL database.

The second interface is for personnel (quantitative geneticist or data analyst) to conduct the genomic analysis, i.e. the genomic prediction pipeline. It has nine functions/tabs including intake of end user animal data, SNP data quality control, prediction of gEPD, multiple trait index, prediction of gBC and retained hybrid vigor, mate selection, prediction of genetic gain due to gEPD selectin, prediction of gain due to gRHET selection and report generation for each animal's genomic profile. The genomic prediction pipeline was developed with C++, Java, R and Python.

Activity 7. Deliver decision making tools to participating producers, analyze feedback and prepare for commercialization of the platform.

This activity was conducted by the current ERA project. Genotypes of all cattle submitted by producers were evaluated for quality control and merged with the genotypes of the refined genomic prediction reference populations for genomic analyses. All genomic analyses were performed using the Platform, and genomic profiles of each individual cattle included gBC, gRHET, gEPDs of RFI, ADG, DMI, HCW, AFAT, REA, LMY, CMAR, FPI, and genomic profiles of each heifer/cow included gBC, gRHET, gEPDs of BRWT, JulianDT, WT200d, DUR, on-test DMI, RFIfat, PBWT, PBBF, AFC, RHPI and RHPI score. Release of these genomic profiles to the respective producers was carried out via the Platform's End-user interface and through communications of technology transfer specialists from Livestock Gentec, University of Alberta.

5. Project Outcomes and Learnings (including, as appropriate)

5.1. Overall project achievements

Through collaborative efforts of this ERA project and the two co-funding projects and multiple commentary projects, all proposed activities have been completed as it proposed with deliverables/outputs achieved. Overall project achievements include:

(1). The project has successfully recruited 10,427 cattle from 88 beef producers. which are more than 10,000 cattle and 40-60 producers as originally planned. This reflects both the interests of producers in the genomic tools the team is providing and great efforts in communicating the tools and their values to producers by the team members.

(2). The project has achieved a very reliable genomic prediction on individual animal's breed composition (gBC) using the refined SNP reference data set of 14 pure breed beef cattle along with improved breed composition prediction methods/algorithms including the" small batch" method and "projection analyses" method. Major breeds predicted using the refined 14 breed SNP reference panel and the methods are highly correlated (r>0.90) with known animal breed composition or with herd breed composition information submitted by the producers, which indicates high confidence and reliability of the Platform from SNP genotyping, SNP data consolidation, genomic analyses, to output of the results. The high accuracy of the predicted gBC leads to reliable estimation of individual animal retained heterozygosity, i.e. gRHET, an indicator of retained hybrid vigor.

(3). With the refined genomic prediction reference data set of 11,292 animals for residual feed intake (RFI), post-weaning dry matter intake (DMI), post-weaning average daily gain (ADG), metabolic body weight (MWT), along with over 7,000 animals for hot carcass weight (HCW), average carcass backfat (AFAT), carcass rib-eye area (CREA), lean meat yield (LMY), and carcass marbling score (CMAR) and a greater density of 87,941 SNPs, the team achieved moderate to moderately high genomic prediction accuracies for the industry cattle submitted to this Emissions Reduction Alberta (ERA) project (ERA cattle). Average accuracy of gEPDs ranged from 0.42 for RFI to 0.49 for DMI for feed efficiency and from 0.41 for AFAT to 0.46 for HCW for carcass merit traits. Notably, three key feed efficiency traits (RFI, ADG, DMI) have even achieved gEPD with a maximal accuracy of 0.65 or greater on some industry animals.

(4). The team has developed and refined a genomic prediction reference data set for female fertility traits, ranging from 1,802 for AFC to 2,792 for DUR traits with a greater density of 81,689 SNPs. The refined reference population size for female fertility traits has doubled the size of ~N=1154 that was originally planned for most traits. The data set is able to predict nine female feed intake and fertility traits gEPDs with an average accuracy ranging from 0.26 for AFC to 0.38 for PBWT on the ERA cattle. The maximum accuracy of all female fertility trait gEPDs achieved 0.65 or greater, indicating the high reliability of genomic prediction tools and a very close genetic link of some ERA cattle to the animals in the reference data set.

(5). The team has successfully refined a feeder profitability index (FPI) with an accuracy from 0.22 to 0.57 with an average of 0.38, a replacement heifer profitability index (RHPI) with an accuracy from 0.22 to 0.78 with an average of 0.38 for the ERA cattle. Furthermore, the team has developed a RHPI score that combines RHPI and hybrid vigor score. In addition, the team has developed a flexible or customized multiple trait selection index calculation method based on the above feed efficiency, carcass, and female fertility traits and their economic weights.

(6). Using independent industry data sets, the team has validated the genetic gains and economic benefits of the selection indexes. The validation on the feeder profitability index (FPI) showed that one unit of the index value equaled $1.54 in feeder cattle economic net return, or selection of cattle with an index value of 100 units above the average herd index value could lead to an increase of $154 in feeder cattle economic net return per animal. The index also has a high and negative correlation (-0.74) with RFI. Therefore, selection of cattle with FPI above the herd average will result in improvement on feed efficiency, which will lead to production of more efficient cattle with less GHG emissions.

(7). The team has also successfully validated the genetic gain and economic benefits of the replacement heifer profitability index (RHPI) score. The validation showed that heifers with a greater than 48.4 RHPI score produced nine more calves per 100 replacement heifers exposed to natural breeding, which resulted in an increase of $22,725 net revenue from 100 heifers if selected based on their RHPI score. Moreover, as the RHPI is negatively correlated with RFI of heifers (-0.22) and the RHPI score puts more weight on the hybrid vigor score, a higher RHPI score will reduce GHG emission as demonstrated by Basarab e al. Animals 2012; 2(2):195-220.

(8). Further to the above genomic profiles including the FPI and RHPI score, the team has developed a virtual sire mating system that evaluates genetic potential of a bull to produce offspring with greater hybrid vigor score in the herd, i.e. average Herd Progeny Heterosis (AHPH). This tool allows producers to assign bulls based on genotypes of heifers/cows to achieve the maximized retained hybrid vigor in their herd.

(9). The team has successfully refined the Genomics-enhanced Whole Herd Genetic Management platform to optimize pipelines for genomic analyses including SNP quality control, SNP merging, gBC prediction, gEPD prediction, multiple trait index calculation, virtual sire mating system, as well as the website portal for producers to submit their animal information and samples and to download genomic profiles results. Using this refined Platform, the project has recruited, successfully genotyped and analyzed 10, 415 cattle for the ERA project and has predicted all genomic profiles as originally planned. All results have been released including the gBC, gRHET, gEPDs of the feed efficiency, carcass traits and FPI for all cattle. This also includes the gEPDs of the female fertility traits, RHPI, and RHPI score for all heifers/cows, and AHPH for bulls. All the genomic values have been released to the respective producers to assist with their genetic selection to improve feed efficiency, and female reproductivity, profitability, and reduce GHG emissions.

Detailed results and discussion are presented in the section of Results of experiments, genomic analyses, and economic model simulations below.

5.2. Technology development, installation and commissioning description

The project had been centered on development and refinement of genomic prediction methods to achieve the maximum prediction accuracy, This included consolidation and quality check on genomic prediction reference data sets, investigation of various statistical algorithms for genetic merit or gEPD prediction, exploring data analysis methods to accurately predict breed composition based on cattle DNA marker genotypes, development and refinement of multiple traits selection indexes, development of genomic analysis pipelines and genomics-enhanced whole herd genetic management platform, and analyzing genomic profiles of industry beef cattle using the Platform. All these technology development and installation were conducted by the research team and the analyses were completed using computers with assistance of Research Associates, postdoctoral research fellows, and students recruited by the project and by the co-funding projects.

To demonstrate the application of the Platform, recruiting beef cattle was conducted by the research team but genotyping of the beef cattle was completed by Neogen Canada, a service provider in livestock genomics. Once cattle genotypes were received, all genomic analyses were conduced by the team using the Platform, and genomic values was then released by the team to the respective producers through the End-use interface (i.e. website portal) and through communications of technology transfer specialists from Livestock Gentec, University of Alberta.

5.3. Experimental procedures/methodology

Consolidation and quality check on genomic prediction reference data sets followed a standard data management procedure to ensure that phenotype and genotype data from different beef cattle populations were merged correctly. Statistical algorithms for genetic merit or gEPD prediction included the commonly used genomic best linear unbiased prediction (GBLUP) and machine learning methods. Prediction of genomic breed composition was carried out based on comparison of genomic features or DNA genotypes of candidate animals with the DNA genotypes of the reference population that consisted of 14 beef cattle pure breeds. Development and refinement of multiple traits selection indexes started with the multiple trait selection index theory that balances the economic weights of the traits and considers correlations among the traits. Development of genomic analysis pipelines and genomics-enhanced whole herd genetic management platform were completed through computer scripts including C, C++, Python, and R that integrates computational processes of all genomic analysis. Detail data analyses and prediction of genomic profiles on industry cattle have been described in the “Work scope overview” above and in the “Results of experiments, genomic analyses, Platform deployment and model simulations” below.

5.4. Modelling details

5.4.1. Modelling genetic merit or gEPD and hybrid vigor of beef cattle traits

A linear model is assumed to predict genetic merit or gEPD, and hybrid vigor of beef cattle traits, which can be descried as:

$$\mathbf{P = \mu + A + D + E}$$

Where P represents phenotypic records of the animals, μ is the population mean, A is the additive genetic effect or breeding value, D is dominance effect, and E donates random environmental effects.

The breeding value is the genetic value that an animal can pass to its offspring. As a progeny inherits only half of genes from its sire or dam, expect progeny difference (EPD) is therefore half of the breeding value of the selection candidate. For genomic prediction based on DNA marker genomics, the predicted breeding values is also called as molecular breeding value of MBV, and as such the expect progeny difference (EPD) is called gEPD.

The dominance effect D is commonly considered as hybrid vigor or heterosis. However, the dominance effect D is not directly estimated. The commonly used indicator of hybrid vigor or heterosis of an animal is based on its breed composition using the formula of retained heterozygosity (RHET) =1- (sum(Pi x Pi)), where i represents the ith breed and Pi is the fraction of each of the n contributing breeds, which can be estimated based on pedigree (RHET) or based on DNA marker genotypes (gRHET). Alternatively, RHET can be obtained as 1-(sum (PSi x PDi), with PSi and PDi being the breed fraction of the ith breed in the sire and dam, respectively, if the pedigree of the animal is known. Animals with greater RHET or gRHET values are considered to have better heterosis and thus have better performance.

5.4.2. Modelling multiple trait selection indexes

A linear model is also assumed for a multiple trait selection index, which can be written as:

$$I=\sum bi*EBVi$$

Where I is a single value of multiple trait selection index, bi is the weighting factor of the molecular breeding value of the ith trait, EBVi is the estimated breeding value of the ith trait. If EBV is predicted using DNA marker genotypes, molecular breeding value (MBV) or genomic estimated breeding value (GEBV) of the trait will be used. Here i can range from 1 to n depending on the number of traits considered in the breeding objective.
Based on the traditional index theory, the b values are derived based on trait economic weights and (co) variance among the traits as proposed by Hazel (Genetics, 1943 Nov;28(6):476-90), i.e.

$$b = P^{-1}Gv$$

Where b vector represents the b values that are used to calculate the index value, P and G are phenotypic and genetic (co) variance among the traits, and *v* is the vector the economic weights of the traits. Under this project, the calculation of b values also considers the variations of accuracies of MBV among animals.

5.4.3. Modelling hybrid vigor

As hybrid vigor is assumed to be linearly correlated with its commonly used indicator genomic retained heterozygosity (gRHET), under this project gRHET is modeled based on its breed composition using the formula of genomic retained heterozygosity (gRHET) =1- (sum(Pi x Pi)), where i represents the ith breed and Pi is the fraction of each of the n contributing breeds, which is estimated based on genotypes (gRHET) (Akanno et al. 2017, Can. J. Anim. Sci. 97: 431–438).

The above modelling details on genetic merit or gEPD, hybrid vigor and multiple trait selection indexes provide theoretical bases for development of the genomic selection tools to improve performance of complex traits.

5.5. Results of experiments, genomic analyses, platform deployment and model simulations.

Detail project results and outcomes of each activity is described below.

Activity 1: Recruiting and training beef producers on the application of the genomics-enhanced Whole Herd Genetic Management platform

The objective was to recruit and train beef producers to use the Whole Herd Genetic Management Platform to improve genetics of their beef cattle herds with deliverables of recruitment of 5000 (+/- 10%) animals from beef producers in the first 1.5 years and 5000 (+/- 10%) animals in year 2. To recruit and train beef producers, the project website (https://www.beefgenomicprediction.ca/) has been developed and constantly refined to include introduction of the project, explanations on commonly used terminologies in animal genetics and genomics, breeding and selection methods, how to use the opportunities, and how to interpret the genomic profile results. Introduction of the project has also been distributed via newsletters from Livestock Gentec of the University of Alberta, Alberta Beef Producers, and Beef Cattle Research Council. In collaboration with the technology transfer team of Livestock Gentec, the project has been introduced through multiple events including presentations at the Grey Wooded Forage Association (June 22, 2022), the Kinsella field day (July 20, 2022), AgSmart (Olds College, August 9-10, 2022), Canadian Beef Conference in Penticton (August 15th-18th, 2022), and at the Lacombe field day (July 26th 2023). Furthermore, the team hosted a virtual information session for participating producers on March 29, 2022 and a virtual meeting with producers on October 27th 2022 to help interpret the genomic results and provide advices on how to use the genomic profile results to improve genetics of their herd. Under the project, 10,427 cattle from 88 individual producers and organizations were enrolled in the project and genetic materials of their cattle were received and genotyped on the GGP-100K single nucleotide polymorphisms (SNP) panel through the project, of which 10,415 genotypes or 99.9% passed the quality control and were used for subsequent genomic analyses (ERA cattle). Therefore, tasks of activity 1 have been completed and the goals have been achieved.

Activity 2: Characterization of whole herd genomic profiles for producers' herds with advanced genomic prediction tools

The objective of this activity was to further refine the genomic prediction reference data sets to conduct the most effective genomic predictions with improved genomic prediction accuracy. This included gEPDs for feed efficiency, carcass merit, female fertility, and GHG emission. For reliable genomic prediction, a genetically relevant and refined reference data set that includes both genotypes and phenotypes is essential. To achieve this goal, the project utilized three reference data sets including (1). A refined beef cattle pure breed SNP reference panel for breed composition prediction (Table 3); (2). A refined beef cattle genomic prediction reference data set for feed efficiency and carcass merit traits (Table 1); (3): A refined beef cattle genomic prediction reference data set for female fertility traits (Table 2). In addition, prediction of gEPDs for methane emissions was attempted using a beef cattle genomic prediction reference data set for greenhouse gas (GHG) emission that was obtained from multiple previous projects including "Methane emissions from beef cattle bred for low residual feed intake, 2014R073R, 2014-2017, led by John Basarab) and "gGreenBeefCow" project (2016R033R, 2016 to 2019 led by Carolyn Fitzsimmons).

The refined beef cattle pure breed SNP reference panel for breed composition prediction included genotypes of a 50K SNP panel of 4,719 cattle from 14 pure beef breeds that was refined under a previous Genome Alberta funded project (A3GP37, PI: Changxi Li, co-PI: John Basarab, 2019-2021). A list of the 14 breeds and numbers of reference animals within each breed were presented above in Table 3. This pure breed reference SNP panel was used to predict genomic breed composition (gBC) of the beef cattle submitted by producers. This prediction was based on 15,401 common SNPs between the pure breed reference SNP panel and the GGP_100K SNP panel which was used to genotype the industry cattle submitted to this ERA project. Prediction of gBC was conducted using both Admixture software (Alexander et al. Genome Res. 2009; 19(9):1655–64) and the customized pipeline of the Genomics-enhanced Whole Herd Genetic Management Platform developed in the current project. Results from both methods produced gBC results that were generally consistent. However, through the gBC analyses, the team identified that the Admixture software with default parameters produced unreasonable gBC results when crossbred animals were more genetically close to each other. The team did a trial by randomly selecting animals to form a smaller batch of cattle for gBC prediction, and it was found a batch size of 30 animals or less produced more reasonable gBC. Later, the team also used a method of "projection analyses" to replace default parameters, and the gBC results were similar to that obtained from the small batch analyses. As a result, the "projection analyses" or "small batch analyses" method had been used to predict individual animal gBC. With the refined reference SNP panel of the 14 pure breeds and the updated gBC prediction methods and algorithm, major breeds predicted from this project based on SNP genotypes have been all aligned with expected breeds submitted by the producers, indicating great reliability of the gBC prediction. The predicted gBC results have been used to calculate genomic retained heterozygosity (gRHET) for each animal, an indicator of retained hybrid vigor (RHV), for the industry beef cattle submitted by producers to this project in subsequent genomic analyses.

The Refined beef cattle genomic prediction reference data set for feed efficiency and carcass merit traits (Version 1, 2021) included 11,292 animals for residual feed intake (RFI), post-weaning dry matter intake (DMI), post-weaning average daily gain (ADG), metabolic body weight (MWT), and over 7,000 animals for hot carcass weight (HCW), average carcass backfat (AFAT), carcass rib-eye area (CREA), lean meat yield (LMY), and carcass marbling score (CMAR) (Table 1). These same animals all included 50K single nucleotide polymorphisms (SNPs) genotypes. In close collaboration with the co-funding BCRC Genomic Prediction Platform project, the 50K SNPs of the reference data set were imputed to 140K SNPs, resulting in 87,941 common SNPs with the GGP-100K SNP panel that is widely used to genotype industry cattle at present. The heritability captured by the 87,941 SNPs ranged from 0.22 for RFI to 0.455 for MWT, which are similar (RFI and HCW) or greater (DMI, ADG, MWT, AFAT, CREA, LMY and CMAR) than the heritability estimates obtained previously with the 50K SNPs (Zhang et al. BMC Genomics (2020) 21:36, Wang et al. BMC Genomics (2020) 21:38), indicating a high quality of the refined reference population for genomic prediction achieved from this project.

To further ensure a reliable genomic prediction, the reference data set was used to generate molecular breeding values (MBV=2 x gEPD) using the genomic Best Linear Unbiased Prediction (GBLUP) method with the G matrix constructed based on the 87,941 SNPs for a separate set of 4,336 industry animals. The industry cattle also had phenotypic data collected by Dr. John Basarab under the previous "Development and deployment of MBVs/gEPDs for feed efficiency and carcass traits that perform in commercial beef cattle" funded by Genome Alberta and Alberta Meat and Livestock Agency (2015-2020; RES0028243, referred to as the gEPDs for commercial cattle project) and the on-going RDAR Feeder Profit Index validation project. The molecular breeding values (MBV=2 x gEPD) of the 4,336 industry animals were predicted only based on their SNP genotypes by this project and the predicted MBVs were then compared to the actual phenotypes of the traits to assess the accuracy of genomic prediction by Dr. John Basarab. The validation results have showed an average accuracy ranging from 0.497 for RFI and 0.658 for CREA, indicating that the refined reference population is able to generate gEPDs with a moderately high accuracy for Canadian beef cattle on the feed efficiency and carcass merit traits.

Consolidation and refinement of genomic prediction reference data set for female fertility traits were conducted in close collaboration with the co-funding BCRC Fertility DMI project. Initially, phenotypic data including heifer birth weight (BRWT), heifer Julian date of birth (JulianDT), 200-days wean weight (WT200d), on-test feeding duration (DUR), on-test DMI (DMI), on-test residual feed intake adjusted for off-test backfat thickness (RFIfat), pre-breeding weight at first parity (PBWT), pre-breeding backfat at first parity (PBBF) and age at first calving (AFC) were compiled. The original fertility related data, which included from 1,003 animals for AFC to 1,305 animals for Julian date of birth, were primarily from replacement heifers born from 2004 to 2018 at Lacombe Research and Development Centre (LRDC) and Roy Berg Kinsella Research Ranch (KIN). The animals were genotyped on a 50K SNP panel and it has 35,438 SNPs in common with the GGP_100K SNP. This "version 1" genomic prediction reference data set yielded relatively low genomic prediction accuracy, ranging from 0.21 for AFC to 0.30 for DUR for industry beef cattle. To enhance genomic prediction accuracy, the team compiled fertility data of 1,523 commercial heifers under the BCRC Fertility DMI project, and added the commercial heifer data to the reference population. This created a larger genomic prediction reference data set (N=2,834), ranging from 1,802 animals for AFC to 2,792 animals for DUR. The original 50K or low-density SNP genotypes of the combined animals in the reference population set were imputed to 140K SNPs under this project, which has 81,689 SNPs in common with the GGP-100K SNP panel. As a process of data quality control, heritability for each of the 9 traits based on the newly combined data set was also estimated under this project. The results showed that the genomic heritability estimates ranged from 0.17 for AFC to 0.44 for PBWT. The genomic heritability estimates were greater than that obtained based on the reference data set without the commercial heifer data (i.e. version 1) except for AFC, which had the same heritability estimate, indicating that the newly combined and refined data set with imputed 140K SNPs (version 2) has added values and is more suitable for genomic prediction. As a comparison study on genomic prediction accuracy, the first version of the dataset without the commercial heifer data and with 50K SNPs (version 1) was compared to the newly refined version of dataset that combined LRDC, KIN and commercial heifers and with imputed 140K SNPs (version 2). Both were used to predict gEPDs of a subset of cattle (N=8,270) submitted by producers to this ERA project. The results showed that inclusion of more data set in the reference population enhanced the genomic prediction accuracy for all the traits, with an improvement ranging from 6.9% for DMI to 25.9% for PBBF.

Although the female fertility reference population size was almost doubled through the BCRC Fertility DMI project by including commercial cattle data and further refined by this project as described above, the genomic prediction accuracies were still lower than that for the refined reference data set for feed efficiency and carcass merit trait. This is expected as the reference data set of female fertility related traits was still smaller and the fertility traits are lowly heritable.

As the above genomic prediction was conducted using single trait GBLUP, it was hoped that more advanced genomic prediction methods may help improve the genomic prediction accuracy. Therefore, the co-funding BCRC Genomic Prediction Platform project evaluated the performance of multi-trait GBLUP and machine learning methods including single trait STKRR, STLinearSVR, and multiple trait MAK_MTKRR, MAK_MTLinearSVR using the reference data set of female fertility related traits.

The multi-trait methods evaluated all other traits in the data set for genetic correlations and used the most genetically correlated trait as 'Assistant Trait" to improve genomic prediction accuracy for the target trait. The results showed that target traits with relatively lower heritability benefited more from the multi-trait GBLUP when a genetically correlated trait with a greater heritability was included in the model. For example, the genomic prediction accuracy of AFC (heritability =0.17) increased by 10.8% when a genetically correlated trait JulianDT (genetic correlation with AFC=-0.83, heritability=0.26) was included in the model as the "Assistant Trait". However, if the target trait has relatively greater heritability, including a genetically correlated trait that has a lower heritability will not improve results and may even decrease its genomic prediction accuracy. This was shown by including AFC for genomic prediction of JulianDT, which reduced the accuracy. Therefore, careful review and selection of "Assistant Trait" is important to improve genomic prediction accuracy when multi-trait GBLUP is implemented.

As for the performance of machine learning methods under single and multiple trait scenarios, we found that the single trait STKRR model yielded 1.86%-26.86% greater prediction accuracies for feed intake and fertility traits including on-test residual feed intake adjusted for off-test backfat thickness (RFIfat), on-test daily dry matter intake (DMI), feeding event duration (DUR), age of first calving (AFC), pre-breeding backfat at first parity (PBBF), and pre-breeding weight at first parity (PBWT) in comparison to the

STGBLUP method that had an average accuracy of 0.39 for the traits investigated. For growth and development traits including weaning weight adjusted to 200 days (WT200d) and birth weight (BRWT) of heifers, STLinearSVR yielded 5.44%-6.76% greater accuracy than the STGBLUP. As for the multiple trait methods, MTGBLUP and MAK_MTLinearSVR yielded 0.42%-32.64% for AFC, RFIfat, WT200d and PBBF and 1.1%-20.42% greater accuracies for PBWT, AFC, WT200d, DUR, RFIfat, and BRWT than the single trait models of STGBLUP and STLinearSVR, respectively.

In an attempt to shed some light on the general trend and potential of the multiple trait GBLUP and machine learning methods in improving genomic prediction accuracy for complex traits, all the above comparisons and discussions were presented without considering the standard error (±SE) of the genomic prediction accuracy that was obtained through a 5-fold cross validation. However, it was noted that the improvements of genomic prediction were not significant if the standard error of accuracy was considered at $p<0.05$. Furthermore, the improvement of the multiple trait GBLUP and machine learning methods on accuracy was not consistent across traits. For example, the accuracies for BRWT and PBBF using SKRR and STLinearSVR were found to be relatively low in comparison to the single trait GBLUP method. This might be caused by many factors such as sub-optimizing algorithms for the multiple trait GBLUP and the complexity of converging machine learning methods in comparison to the simpler algorithms used in single trait GBLUP. Therefore, further research is needed to identify suitable multiple trait GBLUP, machine learning, or deep learning models based on trait data structure and genetic architecture. This will ensure these advanced genomic prediction models can be applied effectively in a routine genomic prediction. Given current considerations of reliability, feasibility, and computational demands, single trait GBLUP is still considered most appropriate for large scale genomic prediction in the beef cattle industry.

Activity 3. Construction of flexible multiple trait selection index for producers to make a genetic selection

With gEPDs predicted on each of the beef performance traits, producers can use gEPDs of a single trait to make a genetic selection as they like. However, this is not optimal due to some unfavorable genetic correlations between performance traits. To develop a single value tool for genetic selection, joint efforts from the co-funding the BCRC Fertility DMI project, co-funding BCRC Genomic Prediction Platform project, and complementary RDAR Feeder Profit Index validation project were made to construct and refine two multiple trait selection indexes that balance all economically important traits for a better profitability, namely feedlot profitability index (FPI), replacement heifer profitability index (RHPI) and its score.

The feedlot profitability index (FPI) was based on traits and economic weights from the previous Kinsella breeding project (2013-2018) and the "gEPDs for commercial cattle" project (2015-2020). Trait units were converted to metric, with more emphasis on the economic weighting of feed efficiency traits and carcass net returns to improve profitability. Six traits are considered in the index:

$$\mathbf{FPI = b1*RFI\_mbv + b2*ADG\_mbv + b3*HCW\_mbv + b4*AFAT\_mbv + b5*LMY\_mbv + b6*CMAR\_mbv}$$

Using the refined method in the estimation of the accuracy of molecular breeding values (mbv, mbv=2*gEPD) for each trait of individual cattle from the GBLUP analysis as described above and updated (co)variance among traits and their economic weights, the project has successfully derived the b values of the above FPI for each individual animal. Therefore, each animal has a unique set of b values for its index calculation.

Table 4. Summary of feedlot profitability index (FPI) relative emphasis and correlation with the traits to improve feed efficiency, growth and carcass values in a population of 4,336 industry cattle for validation

Trait	Relative emphasis	Correlation of FPI index with trait mbv*
RFI_mbv	-42	-0.74
ADG_mbv	20	0.59
HCW_mbv	16	0.51
AFAT_mbv	13	0.23
LMY_mbv	7	-0.14
CMAR_mbv	2	0.03

**mbv: molecular breeding value (MBV). RFI, ADG, HCW, AFAT and CMAR were defined previously in Table 1. b value is the weighting factor of the trait to obtain the single index value of an animal as defined in the generalized formula of FPI above.*

After the refinement, the method was applied to calculate index values on 4,336 industry animals with economic returns that were collected by Dr. John Basarab from the “gEPDs for commercial cattle” project (2015-2020). Considering the standard deviation of genetic variance for the traits, the relative emphasis of the index on the trait is 42% (absolute value of -42), 20%, 16%, 13%, 7%, and 2% for RFI, ADG, HCW, AFAT, LMY, and CMAR, respectively. The FPI showed a strong negative correlation with RFI (-0.74), and positive correlations with ADG (0.59) and HCW (0.51) (Table 4), indicating that animals selected with greater index values would lead to more feed efficient cattle with better growth and carcass weight. The index has weak to moderate but positive correlations with CMAR (0.03) and AFAT (0.23). However, due to a strong negative correlation between LMY and AFAT (-0.87) biologically, which was also accounted for in the index calculation, the index ended up having a weak negative correlation with LMY (-0.14), despite the positive economic weight assigned to LMY. To validate the economic benefit of the FPI, the index values were correlated with carcass net economic return on 137 commercial cattle collected under the complementary RDAR Feeder Profit Index validation project. These 137 beef cattle were not included in the reference population for gEPD prediction, thus representing an independent validation data set. This independent validation through the complementary BCRC Fertility DMI project showed that one unit of the index value equaled $1.54 in feeder cattle economic net return, or selection of cattle with an index value of 100 units above the average herd index value could lead to an increase of $154 in feeder cattle economic net return per animal. The validation results have been published in Basarab, Valente, Wang, Vinsky, Li, Plastow, 2023, Proceedings, 56th American Association of Bovine Practitioners Annual Conference (AABP proceedings, Vol: 56, No: 2, https://doi.org/10.21423/aabppro20238840.

Refinement of replacement heifer profitability index (RHPI) started with the economic weights generated by the complementary BCRC Fertility DMI project and was based on trait MBVs predicted by this project. Initially, all 9 female fertility related traits were considered. However, due to very strong genetic correlations between DMI and RFIfat (0.81) and between JulianDT and AFC (-0.97), DMI and JulianDT were dropped from the index calculation. Therefore, the RHPI has seven traits, which are expressed as:

RHPI = b1* BRWT_mbv + b2* WT200D_mbv + b3* RFIfat_mbv + b4* DUR_mbv + b5* PBWT_mbv + b6* PBBF_mbv + b7* AFC_mbv

Similar to FPI, the refinement method that considers the variation of genomic prediction accuracies of molecular breeding values among animals in addition to trait economic weights and (co)variance among traits has been successfully applied in deriving the b values for each individual animal for RHPI. Therefore, each animal has a unique set of b values for its index calculation. A summary of RHPI including economic weights with relative emphasis and correlation with the traits was presented in Table 5. Considering the standard deviation of trait genetic variation, the relative emphasis of the index on the trait are 25% (absolute value of -25), 8%, 10% (absolute value of -10), 11% (absolute value of -11), 19% (absolute value of -19), 14%, and 13% (absolute value of -13) for BRWT, WT200D, RFIfat, DUR, PBWT, PBBF, and AFC, respectively, i.e.

RHPI = (-25* BRWT_mbv) + 8* WT200D_mbv+ (-10* RFIfat_mbv)+ (-11* DUR_mbv)+ (-19* PBWT_mbv) +14* PBBF_mbv+ (-13)* AFC_mbv

Table 5. Summary of replacement heifer profitability index (RHPI) with relative emphasis and correlation with the traits

Trait	Relative emphasis	Correlation of RHPI with trait mbv*
BRWT	-25	-0.74
WT200D	8	-0.35
RFIfat	-10	-0.22
DUR	-11	-0.31
PBWT	-19	-0.55
PBBF	14	0.30
AFC	-13	-0.25
JulianDT	NA	-0.62
DMI	NA	-0.34

**mbv: molecular breeding value. BRWT, WT200D, RFIfat, DUR, PBWT, PBBF, AFC, JulianDT, and DMI were defined previously in Table 2. b value is the weighting factor of the trait to obtain the single index value of an animal as defined in the generalized formula of RHPI above. NA: Not applicable as the traits were not included in the index.*

The RHPI showed a strong negative correlation with BRWT (-0.74), moderate negative correlations with WT200D (-0.35), RFIfat (-0.22), DUR (-0.31), PBWT (-0.55), AFC (-0.25), and a moderate positive correlation with PBBF (0.30). The negative correlation (-0.22) between RHPI and RFIfat indicates that selection of RHPI with a greater value will result in a lower value of RFI or more feed efficient heifers, which will also lead to reduction of GHG emissions. The RHPI ended up having a negative correlation with WT200D (-0.35) although the economic weight on WT200D was positive, which is likely due to a relatively strong and positive correlation between BRWT and WT200D (0.58).

The RHPI provides a tool to select replacement heifers to improve female fertility related traits including feed efficiency. However, the genetic improvement may likely take generations to realize due to relatively low heritability of the traits. In general, traits with low heritability are more influenced by non-additive genetics including hybrid vigor or heterosis, therefore a Replacement Heifer Profit Index Score (RHPI Score™) was developed in close collaboration with the co-funding BCRC Fertility DMI project (FDE.06.17, led by Basarab and Li, completed in March 2023 with the final project report submitted in April 2023). The RHPI Score™ included both non-additive (gRHET) and additive genetic effects (RHPI), and was designed to select replacement heifers with an increased chance of becoming pregnant during their first normal breeding season, and then to produce a calf every year, while improving feed efficiency, longevity and controlling production costs. The RHPI score can be expressed as RHPI Score™ = gRHET *100 + RHPI *0.2, where gRHET is genomic retained heterozygosity (gRHET) with gRHET=1- (sum(Pi*Pi)) over n, where Pi is the fraction of each of the n contributing breeds of the heifer that is predicted based on the pure breed SNP reference of 14 breeds developed by the team. To evaluate usefulness of the RHPI score, a validation was conducted by the complementary BCRC Fertility DMI project. The results showed that heifers with an RHPI score greater than 48.4 produced 9 more calves per 100 replacement heifers exposed to natural breeding. This would result in an increase of $22,725 net revenue from 100 heifers if they were selected based on their RHPI score. The validation results have been published in Basarab, Valente, Wang, Vinsky, Li, Plastow, 2023, Proceedings, 56th American Association of Bovine Practitioners Annual Conference, Vol: 56, No: 2, https://doi.org/10.21423/aabppro20238840).

In summary, the above refined FPI™, RHPI™, and development of RHPI Score™ provide useful tools for producers to select bulls and replacement heifers, respectively, to improve feed efficiency, female reproductive performance, and thus to reduce the carbon footprint of beef production while improving profitability. These refined FPI, RHPI, and RHPI scores were also generated on cattle submitted by producers, which will be described below under Activity 7. In addition, a flexible multiple trait selection index calculation method has been developed based on the above feed efficiency, carcass, and female fertility traits, and their economic weights, (co)variances, and gEPD accuracies, which allows producers to select traits from a list of current traits in the dataset for calculation of a customized multiple trait selection index that fits their breeding objective. This customized or flexible index calculation has been integrated into the Genomics-enhanced Whole Herd Genetic Management Platform and can be generated for producers who provide a list of their own selected traits.

Therefore, tasks of activity 3 have been completed and the goals have been achieved.

Activity 4. Assignment of mates in the herd to maximize both gEPDs and retained hybrid vigour

Activity 4. 1. Delineate molecular basis of retained hybrid vigour

Although gRHET is commonly used as an indicator of hybrid vigor or heterosis, the team has attempted to delineate molecular basis of hybrid vigor of gRHET based on DNA markers and explore other heterosis indicators using regression analyses. This activity was carried out through the co-funding BCRC Genomic Prediction Platform project, and the investigation has made several interesting discoveries including: (1). For finishing cattle, effects of DNA genotype-based indicators are more evident for feed intake and carcass traits including DMI, ADG, MWT, HCW, AFAT, REA, LMY and CMAR but less evident on RFI. (2). Certain functional DNA genotype-based indicators on the growth-related traits were all positive, indicating that cattle with greater hybrid vigor tended to grow faster.

Activity 4.2. Develop mate selection strategies to maximize both gEPDs and gRHET

A virtual mating selection tool has been developed by this project in collaboration with the co-funding BCRC Genomic Prediction Platform project, in which each of the bulls in a herd is virtually mated with each cow/heifer in the same herd. For each virtual mating, the expected gRHET of the offspring is calculated as: offspring gRHET= 1- $\sum$ P_sire_i *P_dam_i; where P_sire_i and P_dam_i are the genomic breed composition fractions (gBC) of breed i of the sire and of the dam, respectively. Assuming each virtual mating produces one offspring, the average offspring gRHET of each bull in the herd or average Herd Progeny Heterosis (AHPH) is calculated by taking the average of offspring gRHET from the virtual mating of the bull with all the cows/heifers in the herd. Therefore, the tool is able to evaluate each of the candidate bulls through virtually mating with each cow/heifer in the herd for the potential to produce offspring with greater gRHET in the herd, i.e. average Herd Progeny Heterosis (AHPH).

Therefore, all tasks of activity 4 have been completed and the goals have been achieved.

Activity 5. Prediction on genetic gains for producers resulted from the application of the Whole Herd Genetic Management Platform

The genetic gains translated to profitability and cost-benefit analyses from the application of the Whole Herd Genetic Management Platform were conducted by the co-funding BCRC Genomic Prediction Platform project based on the validation results of economic benefits of the feedlot profitability index (FPI) and RHPI score from the complementary RDAR Feeder Profit Index validation project and from the co-founding BCRC Fertility DMI project, respectively. The analyses considered two scenarios via simulation. The two scenarios included (1): a combined cow-calf and feedlot operation, which would incur all costs and revenue associated with genotyping and genomic analysis; and (2): a separate cow-calf and feedlot operations, where the cow-calf operation would cover all genotyping and genomic analysis costs and obtain a proportion of additional revenue due to genomic selection from the feedlot operation through marketing up their weaned calves.

The results have showed that implementing genotyping and genomic analysis for bull selection on FPI can generate a lifetime additional or net revenue of approximately $64 (i.e. 852-788=64) for each cow it mates in comparison to not implementing genotyping and genomic analysis for bull selection (Control), assuming a cow crop rate of 85%. As the calf crop increases to 90% and 95%, the lifetime additional or net revenue can increase to $70 and $73, respectively, for each cow the selected bull mates.

For a separate cow-calf and feedlot operations, two revenue splits were considered, one in which the cow-calf operation received 25% of additional revenue and another where the cow-calf operation received 50% of additional revenue from the feedlot operation. The results showed that the lifetime net revenue observed in the cow-calf operation (who pays genotyping and genomic analysis cost for bull selection) was estimated at $21 for each cow that the selected bull mates, assuming a cow crop rate of 85% and the cow-calf operation received 25% of the feedlot net revenue (i.e. 1235-1214=21). Also, for each cow that the selected bull mates, assuming a cow crop rate of 85% and the cow-calf operation received 50% of the feedlot net revenue, the lifetime net revenue was $42 (i.e. 1256-1214=42). The lifetime net revenue observed in the cow-calf operation could increase to $23 for each cow that the selected bull mates, if the cow crop rate increased to 90% and the cow-calf operation received 25% of the feedlot net revenue. This could be further increased to $46 for each cow that the selected bull mates, if the cow crop rate increased to 90% and the cow-calf operation received 50% of the feedlot net revenue. Overall, the cost-benefit analysis results indicate that implementing the above multiple trait index selection under the Platform can lead to significant economic benefits in beef production, though the extent of these benefits varies among different users. This variation depends on several factors, including the top 20% FPI scores, overall FPI scores, calf crop percentages (85%, 90%, and 95%), the proportion of heifers above the average RHPI score, genotyping and genomic analysis costs, the genetic potential of the herd, existing management practices, and the baseline performance of the animals. The cost-benefit analysis demonstrates the economic viability of implementing genomic selection in beef production, particularly under conditions where traditional management practices are less effective. Producers can leverage this technology to not only enhance profitability but also achieve more consistent and predictable financial returns.

Therefore, tasks of activity 5 have been completed and all goals have been achieved.

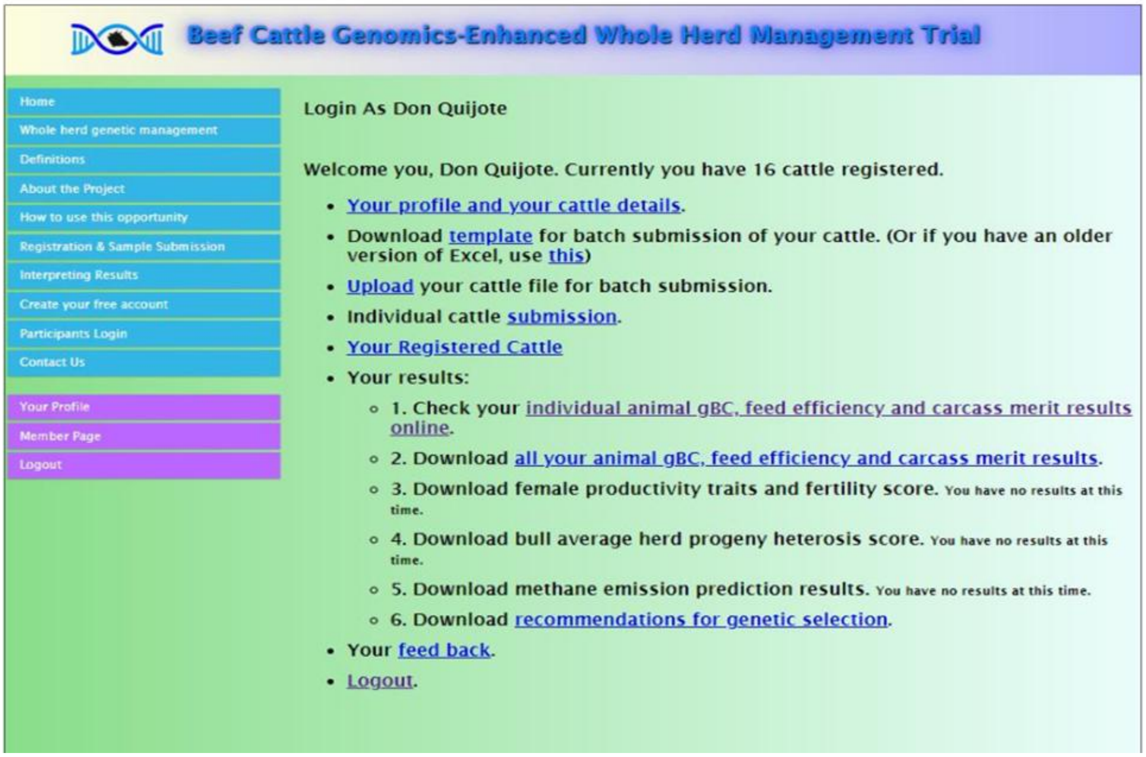
Activity 6. Further refine the genomics-enhanced Whole Herd Genetic Management platform

A web-based genomics enhanced whole herd genetic management platform has been successfully developed and updated according to feedbacks of collaborators and industry participants through this ERA project and through the co-funding BCRC Genomic Prediction Platform project. The Platform has two interfaces. The first interface is a website portal open to producers as shown in Figure 2 above. The website has multiple sections to introduce the genomic prediction tools and provide recommendations on how to submit cattle and use the genomic prediction tools to maximize and manage genetics of a cattle herd. The website also allows producers to create an account, submit information of their cattle, and login back to the account with a secure password to view and download genomic prediction results of their own cattle. The website is hosted at https://www.beefgenomicprediction.ca/, and a screenshot of added download functions is shown below in Figure 5.

The website also has functions to allow submission of cattle as a batch using the template provided, and download and view all cattle genomic prediction results as a single file. This function facilitates an easier animal submission process and also enables producers to view and compare cattle in a single file, which is illustrated by a screenshot of the page below. The developed web-site portal has been well received by the producers as one producer indicated “I really like being able to access the website from my phone and its very helpful”.

The second interface of the Platform has also been successfully developed. This interface provides a pipeline for a data analyst to analyze genomic profiles, which includes SNP data quality control, SNP merging, genomic breed composition (gBC) prediction, gEPD prediction through GBLUP, and multiple trait index calculation. The second interface of the Platform (i.e. genomic analysis pipeline) is only accessible to personnel who conduct the genomic analysis. However, the pipeline has also been expanded to take animal information submitted through the website by producers to construct customized or flexible indexes with different sets of traits to best fit in their breeding objectives. As an illustration of the genomic analysis pipeline, a screenshot of the program in windows is shown in Figure 6.

The two-interface Platform is ready for industry application. Therefore, all tasks of activity 6 have been completed and the goals have been achieved.

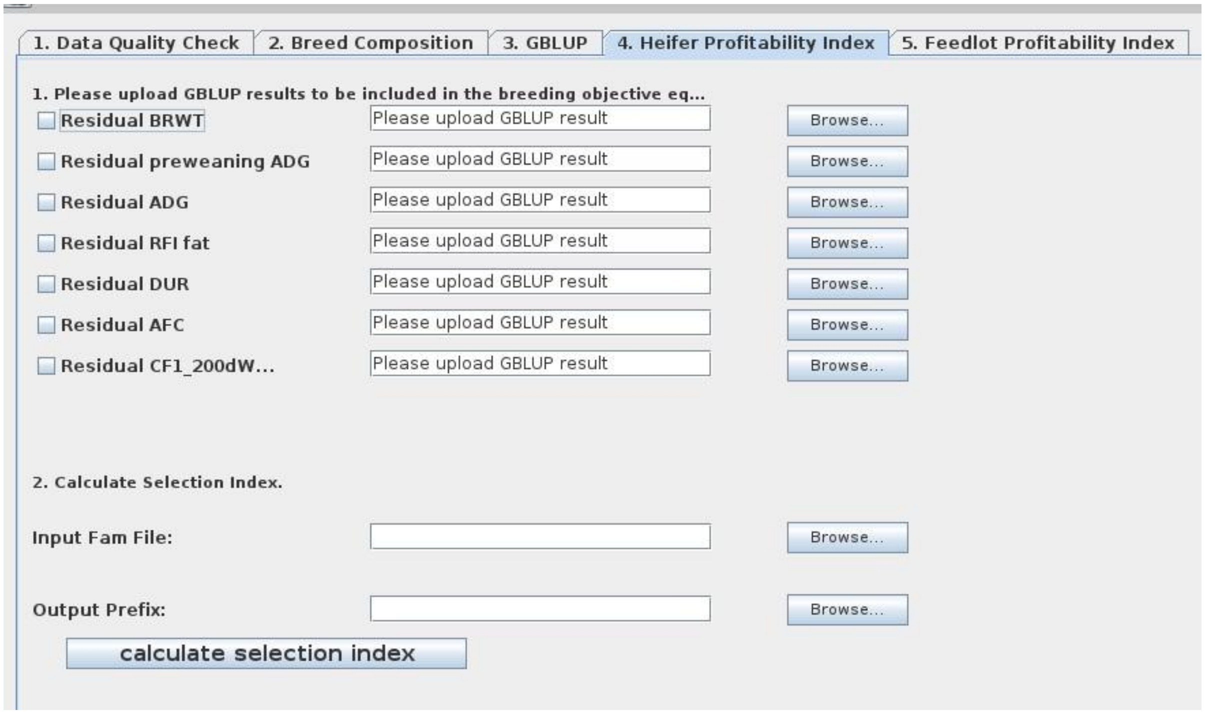


Login As Don Quijote

Welcome you, Don Quijote. Currently you have 16 cattle registered.

- Your profile and your cattle details.
- Download template for batch submission of your cattle. (Or if you have an older version of Excel, use this)
- Upload your cattle file for batch submission.
- Individual cattle submission.
- Your Registered Cattle
- Your results:
  - 1. Check your individual animal gBC, feed efficiency and carcass merit results online.
  - 2. Download all your animal gBC, feed efficiency and carcass merit results.
  - 3. Download female productivity traits and fertility score. You have no results at this time.
  - 4. Download bull average herd progeny heterosis score. You have no results at this time.
  - 5. Download methane emission prediction results. You have no results at this time.
  - 6. Download recommendations for genetic selection.
- Your feed back.
- Logout.

Figure 5. A screen shot of refined End-user interface with added download functions.



1. Please upload GBLUP results to be included in the breeding objective eq...

☐ Residual BRWT — Please upload GBLUP result
☐ Residual preweaning ADG — Please upload GBLUP result
☐ Residual ADG — Please upload GBLUP result
☐ Residual RFI fat — Please upload GBLUP result
☐ Residual DUR — Please upload GBLUP result
☐ Residual AFC — Please upload GBLUP result
☐ Residual CF1_200dW... — Please upload GBLUP result

2. Calculate Selection Index.

Input Fam File:

Output Prefix:

Figure 6. A screen shot of refined genomic prediction pipelines with program in windows.

Activity 7. Deliver decision making tools to participating producers, analyze feedback and prepare for commercialization of the platform.

This activity was conducted under the ERA project with collaborations of the co-funding BCRC Genomic Prediction Platform project and assistance of the Livestock Gentec technology transfer team. The 10,415 cattle that were successfully genotyped under Activity 1 includes 2,755 bulls and 7,660 heifers/cows (Alberta and outside) and they all have been analyzed for the genomic profiles using the genomic prediction pipeline refined under the Platform.

The genomic profiles of each of the 10,415 beef cattle (bulls, heifers/cows) included genomic breed composition (gBC), gRHET, gEPDs of RFI, ADG, DMI, HCW, AFAT, REA, LMY, CMAR, and FPI (Table 6). For the feed efficiency traits, the average accuracy of the trait gEPD ranged from 0.42 for RFI to 0.49 for DMI. For the carcass merit traits, the average accuracy of the trait gEPD ranged from 0.41 for AFAT to 0.46 for HCW and REA. For FPI, the accuracy ranged from 0.22 to 0.57 with an average of 0.38. It is noted that the maximum accuracy of gEPDs of all three key feed efficiency traits (RFI, ADG, DMI) have even reached a level of 0.65 or greater. As the economic benefit of FPI has been validated under Activity 3, the FPI value of each 10,415 cattle along with their gBC, gRHET, gEPDs of RFI, ADG, DMI, HCW, AFAT, REA, LMY, CMAR have been released to the respective producers through the project website to assist with their genetic selection.

Table 6. Summary statistics of genomic profiles of 10,415 beef cattle (bulls, heifers/cows) submitted by beef producers

Trait	Mean	Min.	Max.	Standard deviation (SD)	Average of prediction accuracy	Min. of prediction accuracy	Max. of prediction accuracy
gRHET*	0.52	0.00	0.89	0.26	---	---	---
RFI_gEPD	0.00	-0.19	0.28	0.05	0.42	0.24	0.65
ADG_gEPD	0.00	-0.11	0.11	0.03	0.45	0.26	0.68
DMI_gEPD	-0.01	-0.53	0.61	0.14	0.49	0.29	0.73
HCW_gEPD	-0.03	-9.33	6.46	1.86	0.46	0.27	0.61
AFAT_gEPD	0.01	-1.38	1.54	0.4	0.41	0.24	0.57
REA_ gEPD	-0.17	-5.68	4.82	1.31	0.46	0.27	0.62
LMY_ Gepd	-0.05	-1.68	1.66	0.43	0.44	0.25	0.59
CMAR_gEPD	0.68	-37.16	42.02	8.46	0.44	0.25	0.59
CH4_gEPD	0.01	-0.55	0.53	0.14	0.03	0.03	0.17
FPI	-1.05	-98.07	40.65	11.30	0.38	0.22	0.57

*gRHET: genomic retained heterozygosity, calculated for each individual animal based on its genomic prediction composition predicted using the refined SNP panel of 14 pure beef breeds. "---": accuracy is not available. gEPD: genomic expected progeny difference = 0.5*molecular breeding value (MBV). Residual feed intake (RFI), post-weaning dry matter intake (DMI), post-weaning average daily gain (ADG), hot carcass weight (HCW), average carcass backfat (AFAT), carcass rib-eye area (REA), Lean meat yield (LMY), and carcass marbling score (CMAR), CH4: enteric methane emissions in average daily CH4 in g/day, FPI: feedlot profitability index.*

For the 7,660 heifers/cows submitted by beef producers, additional genomic profiles on female fertility related traits including gEPDs of BRWT, WT200D, RFIfat, DUR, PBWT, PBBF, AFC, JulianDT, and DMI were predicted using the newly refined cow fertility and feed efficiency reference data set (Version 2), and RHPI values were also calculated using the RHPI formula as described above. The average accuracy of the trait gEPDs ranged from 0.26 for AFC to 0.38 for PBWT (Table 7). The average accuracy of RHPI was 0.37, ranging from 0.21 to 0.78 (Table 7). These gEPDs and RHPI represent additive genetic merit of the heifers/cows. As fertility related traits are also influenced by non-additive genetic component, e.g. gRHET, a RHPI score that includes both additive and non-additive genetic merit was also calculated for each of the 7,660 heifers/cows, i.e. RHPI score=gRHET *100 + RHPI *0.2. It is noted that the maximum accuracy of gEPDs of all female fertility trait have even reached a level of 0.65 or greater (Table 7), likely due to a closer genetic link of some ERA heifers/cows with the animals in the reference population. Furthermore, the economic benefit of RHPI score has been validated under Activity 3, therefore the RHPI score along with gRHET, gEPDs of the fertility traits, RHPI and have also been released to producers through the Livestock Gentec Technology team, where specific instructions of genetic selection along with individual animal genomic profile data were emailed to each producer. The specific instruction of replacement heifer selection included: (1): Commercial crossbred cows should have greater than 40% hybrid vigour (gRHET) and greater than 40 RHPI Score. (2): Ensure heifers are pre-selected for soundness of conformation and structure, are from an early, easy and single birth, and heifers can attain a body weight of 60% of mature cow weight at breeding (BCS >2.5 on a 1-5 system). These genomic profile prediction results and selection instructions will help the producers to select replacement heifers/cows to improve female reproductive performance and thus profitability with reduced GHG emissions

Table 7. Summary statistics of genomic profiles of 7,660 heifers/cows submitted by beef producers

Trait	Mean	Min.	Max.	Standard deviation (SD)	Average of prediction accuracy	Min. of prediction accuracy	Max. of prediction accuracy
gRHET*	0.49	0.00	0.89	0.25	---	---	---
BRWT_gEPD	0.01	-1.25	1.39	0.36	0.34	0.19	0.74
WT200D_gEPD	0.14	-6.37	7.58	1.65	0.34	0.19	0.85
RFIfat_gEPD	0.00	-0.15	0.12	0.03	0.30	0.16	0.79
DUR_gEPD	0.14	-9.29	11.56	2.14	0.36	0.21	0.86
PBWT_gEPD	0.20	-15.71	13.98	3.25	0.38	0.22	0.88
PBBF_ gEPD	0.00	-0.71	0.61	0.15	0.36	0.21	0.86
AFC_ gEPD	-0.03	-2.87	3.37	0.68	0.26	0.13	0.72
JulianDT _gEPD	0.03	-3.27	4.52	0.78	0.33	0.18	0.82
DMI _gEPD	0.00	-0.21	0.19	0.04	0.34	0.19	0.84
RHPI	-0.59	-60.33	79.79	17.83	0.37	0.21	0.78
RHPI score	48.41	-5.84	93.27	25.45	---	---	---

*gRHET: genomic retained heterozygosity, calculated for each individual animal based on its genomic prediction composition predicted using the refined SNP panel of 14 pure beef breeds. "---": accuracy is not available. gEPD: genomic expected progeny difference = 0.5*molecular breeding value (MBV). BRWT, WT200D, RFIfat, DUR, PBWT, PBBF, AFC, JulianDT, and DMI were defined previously in Table 1. RHPI: replacement heifer profitability index. RHPI score=gRHET *100 + RHPI *0.2.*

The potential of all the bulls (N=2,634) along with heifers/cows in the same herd submitted by beef producers has been analyzed using the virtual mating selection tool, and a summary statistics of average Herd Progeny Heterosis (AHPH) is shown below (Table 8).

Table 8. Summary statistics of average Herd Progeny Heterosis (AHPH) on 2,634 bulls with heifers/cows in the same herd submitted by beef producers

Trait	Mean	Min.	Max.	Standard deviation (SD)	Average of prediction accuracy	Min. of prediction accuracy	Max. of prediction accuracy
AHPH	0.71	0.11	0.98	0.21	---	---	---

AHPH: Average Herd Progeny Heterosis. This value is obtained through a virtual mating analysis where the candidate bull is randomly mated to all heifers/cows in the herd.

The results have showed that there is great variation of heterosis potential of bulls (i.e. AHPH) submitted by the producers, ranging from 0.11 to 0.98. Based on the validation of heterosis results by John Basarab from the above BCRC co-funding projects, a greater value (>0.40) of AHPH is preferred for bull selection for the herd. The genomic profiles of AHPH for the bulls have also been released to their respective producers with selection recommendations, which state as "For cow-calf producers that would like to maintain a hybrid vigour score of their crossbred replacement heifers greater than 40%, it is recommended to use a bull of a different breed or of the minor breed in the herd average breed composition. However, care must be taken when selecting a complementary bull breed or crossbred to ensure it also meets the breeding objectives and does not increase calving difficulty or mature cow weight. The use of AHPH in combination with Feeder Profitability Index (FPI) and individual trait gEPDs can be conducted depending on the breeding objective of the herd".

The project has proposed to reduce GHG emissions through using genomic selection on feed efficiency, i.e. RFI and gRHET. The released genomic profiles as described above have enabled producers to achieve the goal. As the team has compiled a reference population (N=520) for methane emission, an attempt was made to predict genetic merit (i.e. gEPD) directly on methene emissions (CH4). The genomic prediction was completed on all the 10,415 cattle submitted. However, the accuracy of gEPDs were low from 0.03 to 0.17 (Table 6). The low accuracy is likely due to the very small sample size of the reference population (N=520) and relatively low heritability of methane emissions (i.e. ~0.18). These results are not released to producers because of the lower accuracy. Instead, we still recommend to use genomic profiles on feed efficiency, selection indexes, and hybrid vigor of heifers/cows/bulls that have already been released to the producers, to reduce GHG emissions. Nevertheless, the compiled genomic prediction refined population for methane emission provides a valuable resource for future development of genomic prediction tools on direct genetic selection to reduce methane emissions.

In summary, all the 10,415 cattle (2,755 bulls, 7,660 heifers/cows) submitted by 88 producers and organizations have been successfully genotyped and analyzed for genomic profiles. The genomic profiles of each animal including gBC, gRHET, gEPDs of RFI, ADG, DMI, HCW, AFAT, REA, LMY, CMAR, FPI and genomic profiles of each heifer/cow including gBC, gRHET, gEPDs of BRWT, JulianDT, WT200d, DUR, on-test DMI, RFIfat, PBWT, PBBF, AFC, RHPI, RHPI score along with bull's potential to produce offspring with greater hybrid vigor in the herd (i.e. AHPH as in Table 8) have been released to the respective producers through the End-user interface of the Platform (i.e. the website portal) and through communications of technology transfer specialists from Livestock Gentec, University of Alberta.

Therefore, all tasks of activity 7 have been completed and the goals have been achieved.

5.6. Status of the technology risks at the end of the project (both retired risks and risks to be retired)

5.6.1. Technical/scientific risks

As GHG emissions in beef cattle are complex traits, and they are affected by many genes with relatively small effects, the underlying genetic influence and causative mutations for the traits in beef cattle are largely unknown, especially when different beef breeds are examined. In addition, direct measure of GHG emissions in cattle still remains an expensive approach for industry application. Therefore, the project used feed efficiency as an indicator trait to select more genetically efficient beef cattle to reduce GHG emissions, included a hybrid vigor score in the replacement heifer selection index, and implemented a genomic prediction method that only requires DNA genotypes of beef cattle. However, the prediction accuracy on genetic merit (gEPDs), or genomic prediction, is limited by multiple factors including heritability and the reference population size. To reduce the risk, the project used over 11,000 beef cattle for genomic prediction for feed efficiency, over 2,700 female feed efficiency measures, and obtained 520 cattle with GHG emission data. The increase of reference data sets was able to generate genetic merit (gEPDs) for feed efficiency traits (RFI and DMI) with a moderately high accuracy (0.42 to 0.49) for feeder cattle on average (Table 6), and with a moderate accuracy (0.30 to 0.34) for replacement heifers on average (Table 7). For GHG emissions, the current reference data set was able to generate gEPDs for methane emissions but with a very low accuracy (0.03 to 0.17). The project also explored various genomic prediction methods including multiple trait genomic best linear unbiased prediction (GBLUP) and machine learning. However, these machine learning methods did not lead to tangible improvement on genomic prediction accuracy compared to traditional GBLUP. Therefore, single trait BLUP is still considered as an appropriate approach for industry genomic prediction for gEPDs. Although singe trait gEPDs are available from the genomic analyses, the project recommended beef producers to use the multiple trait feedlot profitability index (FPI) and replacement heifer profitability index (RHPI) score to select bulls and replacement heifers as the benefits of the index were validated using independent data sets. However, low genomic prediction accuracy is still a major technology risk associated with genomic prediction, and continuous improvement on genomic prediction accuracy still be an important research focus in the future. To mitigate the technology risks, the project recommends the following research and development areas:

1. Increasing the size and quality of genomic prediction reference data sets still remains the most important strategy to further improve the genomic prediction accuracy, in particular for traits with low heritability. While advanced genomic prediction methods like multi-trait GBLUP and machine learning warrant further investigation.

2. The current Genomics-enhanced Whole Herd Genetic Management Platform provides a valuable tool to assist the Canadian beef industry with their genetic selection and breeding decisions. However, due to the complexity of beef cattle performance traits and their multigenerational genetic influences, the reference data sets require constant recalibration to maintain or improve the accuracy. Therefore, ongoing data collection on representative herds should remain a top priority to ensure viability of the genomic tools.

3. Further exploration of other genomic tools including selection indexes incorporating novel traits and molecular indicators of heterosis based on DNA marker genotype heterozygosity are worthwhile to enhance the benefits of the genomic tools.

5.6.2. Commercialization and market adoption risks

The project recruited 10,427 cattle from 88 producers and organizations within 2 years, which were more than the 10,000 cattle and 40-60 producers as originally planned. As the project had reached the objective of genotyping and analyzing 10,000 (+10/-10%) beef cattle, the team had to decline requests from producers and recommended the producers to contact Livestock Gentec for genomic service. This has indicated that the beef industry is interested in the genomic tools. However, the ERA project provided a subsidy to cover most of the genotyping cost ($30/$45=66.67% per animal) and genomic analysis was free. For commercialization and market adoption, there could be risked associated the relatively high costs of genotyping and genomic analysis. Therefore, market research including a survey on how much the beef industry is willing to pay for the genomic tools may be necessary for the service producer to develop strategies for wider adoption of the genomic tools.

5.6.3. Project plan and timelines risks/uncertainties

The project had progressed well, and had achieved all the objectives and deliverables as originally planned. All the millstones (Milestone 1, Milestone 2, and Milestone 3) had been successfully achieved within the timelines. Although the senior postal doctoral fellow (PDF) left the team for a faculty position before the project was completed, the team hired a junior PDF and multiple summer students to complete all research activities for the co-funding BCRC projects. Furthermore, although AAFC had some restrictions on conference travels and industry outreach travels during and post Covid, the team had completed industry outreach activities through the website portal and through more remote online meetings.

5.6.4. Budget uncertainties and cost escalation risks, including foreign exchange risks

The ERA project and the co-funding BCRC Genomic Prediction Platform project and co-funding BCRC Fertility DMI project had completed all proposed activities within the budgets.

5.6.5. Funding risks

Co-funding of the BCRC Fertility DMI project was already secured before the start of the ERA project. Co-funding of the BCRC Genomic Prediction Platform project was secured as soon as the ERA project was approved. The ERA funds along with co-funding were managed well by the AAFC finance and its collaborative institute University of Alberta.

5.7. Discussion of any changes to the corporate structure of the company or project consortium since commencement of the project

The project has been conducted by a research team of Agriculture and Agri-Food Canada and the University of Alberta. No changes that would affect the execution of the project had occurred to the corporate structure of the company or project consortium since commencement of the project. Furthermore, no original objectives and deliverables for the project had been modified. The project has successfully completed all proposed activities and all objectives and deliverables have been achieved.

5.8. Discussion of any advancements made toward commercialization, commercial deployment or market adoption

The Platform has also been used to predict genomic profiles for multiple other projects including the Gentec Feeder Profit Index validation and deployment project (PI: John Basarab, co-PI: Changxi Li, 2022-2024). This collaboration has further promoted the current Genomics-enhanced Whole Herd Genetic Management Platform so more and more producers have become familiar with the Platform and are more interested in participating in the ERA project. As there is an increased demand from producers after the ERA project, Livestock Gentec at the University of Alberta will carry on the approach of the ERA project to genotype 40,000 heifers/bulls/feeder cattle over 4 years under a recent RDAR/Sustainable CAPS funded "Adoption of Genomic Tools" project "Using genomics to improve the profitability and sustainability of Alberta's beef industry" (PI: John Basarab, 2024-2028). https://rdar.ca/latest/press-releases/rdar-supports-beef-genomics-technology-adoption-with-the-delivery-of-a-2-2m-investment-into-livestock-gentec. Supported by the ERA project, the team has finalized and made the Platform ready to be used by a service provider and will transfer the Platform to Livestock Gentec as requested by the University of Alberta and AAFC. A genomic prediction data sharing agreement has been signed between the University of Alberta and AAFC, and all the Refined genomic prediction reference data sets have been transferred to Livestock Gentec, an Alberta Innovates Centre that brings the commercial benefits of genomics to the Canadian livestock industry. Transfer of genomic prediction pipeline to Livestock Gentec will occur as soon as all related agreements are signed. Dr. Changxi, the principal investigator (PI) of this ERA project, will train members of Livestock Gentec on how to use the Refined Reference Data Sets and genomic prediction pipeline so Livestock Gentec is able to run the genomic analyses independent on a routine basis. This illustrates that the genomic tools or Platform has evolved from the "Demonstration" stage from the ERA project to the "Adoption" stage.

Commercialization of the genomic tools/Platform is expected to happen during or after the "Adoption of Genomic Tools" project, which will be led by Livestock Gentec and will be discussed by the commercialization committee under the AAFC-University of Alberta co-location agreement.

5.9 Description of technology advancement over the course of the project

Through joint efforts of this ERA project and co-funding BCRC Genomic Prediction Platform project and the BCRC Fertility DMI project, major technology advancements have been made, which can be summarized as below:

(1). The utilization of a large genomic prediction reference dataset (i.e. 11,292 for feed efficiency, and 7,299 to 8,081 for carcass traits) and imputed high density of 140K SNPs have led to greater heritability (genetic variance) captured by the DNA markers. As a result, the genomic prediction accuracies for the traits had been improved for Canadian crossbred beef cattle in comparison to previous versions of genomic prediction (Mujibi et al. 2011, J. Anim. Sci. 89:3353–3361; Lu et al 216, J. Anim. Sci. 94(4), 1342-1353), with a maximum genomic prediction accuracy on gEPDs reaching >0.65 for all three key feed efficiency related traits (i.e. RFI_gEPD, ADG_gEPD, DMI_gEPD) in some industry animals.

(2). In close collaboration with the BCRC Fertility DMI project, the genomic prediction reference data set was increased by including commercial beef cow data, leading to a reasonable genomic prediction reference data set for Canadian beef cow feed intake and fertility related traits (1,802 for AFC to 2,792 for DUR). Furthermore, genotypes of these cattle had been consolidated from various SNP panels and had been imputed to from their original 50K SNPs or low-density SNPs to a high density of 140K SNPs resulting in more SNP genotypes in common with the commonly used GGP-100K SNP panel (i.e. 81,689). This refined genomic prediction reference data set represents the first of its kind for Canadian beef cattle female feed intake and fertility traits. Although the overall average genomic prediction accuracy was moderate (i.e. 0.34), this refined genomic prediction reference data set was able to generated gEPDs with a maximum genomic prediction accuracy of >0.65 for all the traits in some industry animals, indicting high relevance of the reference data set to the industry cattle.

(3). Unlike the traditional construction method of multiple trait selection indexes, the calculation of two selection indexes, i.e. feedlot profitability index (FPI) and replacement heifer profitability index (RHPI) considered the variation of molecular breeding values among animals due to their different genetic links to the reference population, in addition to the trait economic weights and (co)variance among the traits. As such, each animal had a unique set of weighing factors to calculate the index value. This represents an advanced approach for multiple trait selection index calculation.

(4). For individual animal's breed composition (gBC) prediction, the project developed and implemented a "small batch" method (i.e. randomly select 30 cattle from industry cattle) and "projection analyses" method for crossbred beef cattle that had a very close genetic relation among themselves. These advanced analyses had led to improved genomic prediction on individual animal's breed composition (gBC). As such, major breeds predicted using the refined 14 breed SNP reference panel and the above methods are highly correlated (r>0.90) with known animal breed composition or with herd breed composition information submitted by the producers, which also led to more reliable estimation of individual animal retained heterozygosity, i.e. gRHET, an indicator of retained hybrid vigor.

(5). Through collaboration with co-funding BCRC Genomic Prediction Platform project, the team evaluated new genomic prediction methods including multi-trait genomic best linear unbiased prediction (GBLUP), single trait STKRR, STLinearSVR and multi-trait MAK_MTKRR and MAK_MTLinearSVR. It was found that traits with relatively low heritability benefited more from the multi-trait GBLUP when a genetically correlated trait with greater heritability was included in the model. In addition, single trait machine learning methods STKRR and STLinearSVR had improved genomic prediction accuracy. However, the performance of multi-trait GBLUP, single trait STKRR, STLinearSVR and multi-trait MAK_MTKRR and MAK_MTLinearSVR was not consistent across the traits, therefore, in consideration of feasibility and requirement of computational power of implementing genomic prediction, the single trait GBLUP is still considered appropriate in beef cattle for large scale industry application. The study has improved our understanding on genetic complexity of beef performance traits and genomic prediction methods and has laid a foundation to further explore new genomic prediction methods including machine learning and deep learning.

(6). The project developed a virtual mating selection tool based on breed composition of sires and dams. This genomic tool allows evaluation of influence of a candidate bull on offspring's hybrid vigor and genetic merit through in silico crossing with cows in the herd. Bulls that have greater trait EPDs or FPI and also produce offspring with greater gRHET (an indicator of hybrid vigor) on average in the herd are the top bulls that should be used to maximize both EPDs and RHV.

(7). The most important technology advancement is the development and refinement of the Genomics-enhanced Whole Herd Genetic Management Platform. The Platform integrates the web-site portal for industry access and the optimized genomic prediction pipelines for routine genomic prediction services including uptake of cattle information or/and genotypes submitted by end-users including beef producers, SNP quality control, SNP merging with reference data sets, gBC and hybrid vigor prediction, gEPD prediction and multiple trait index calculation, virtual mating selection, and release of genomic values to the end-users.

5.10. Analysis of results

With 10,415 cattle recruited from 88 producers and organizations, the project has exceeded the project target number of industry animals analyzed via the Whole Herd Genetic Management Platform (i.e. 10,000 (+/-10%) of beef cattle from Alberta).

For prediction of gBC and gRHET on the industry cattle, a high accuracy of predicting (>0.90) was achieved. For the genomic prediction accuracy on gEPDs for industry animals, the maximal accuracies of all the three key feed efficiency traits and female feed intake and fertility related traits had reached to a level of >0.65. However, the average accuracies of the traits were still below 0.65, and some industry cattle had an accuracy below 0.20, in particular for female fertility related traits. Although the size of genomic prediction reference populations and quality of both phenotype and genotype data are two important factors to achieve reliable accuracy of predictions on genetic merit for industry application in beef cattle, it is expected that the accuracy of gEPDs of the traits will vary depending on genetic links of industry cattle with the reference population. This is even more evident in beef cattle where distinct breeds or crossbreds with different breed composition are used in commercial herds. Therefore, inclusion of representative industry cattle in the reference data set, identification of causative DNA variants for genomic prediction, and exploring new genomic prediction statistical methods still remains to be the important areas for future research to further improve genomic prediction for the performance traits in beef cattle.

Due to combinations of multiple traits, the accuracy of the flexible multiple trait selection indexes was moderate (0.37 for RHPI and 0.38 for FPI). To improve the effectiveness and genetic gain of the index selection, a RHPI score was developed under the co-funding BCRC Fertility DMI project. The RHPI score included both non-additive (gRHET) and additive genetic effects (RHPI), and was designed to select replacement heifers with an increased chance of becoming pregnant during their first normal breeding season, and then to produce a calf every year, while improving feed efficiency, longevity and controlling production costs. The RHPI score can be expressed as RHPI Score™ = gRHET *100 + RHPI *0.2, where gRHET is genomic retained heterozygosity (gRHET) with gRHET=1- (sum(Pi*Pi)) over n, where Pi is the fraction of each of the n contributing breeds of the heifer that is predicted based on the pure breed SNP reference of 14 breeds developed by the team and industry cattle DNA genotypes. Furthermore, the economic benefits of index selection on FPI and RHPI score were validated using independent industry cattle data through the complementary RDAR Feeder Profit Index validation project and through the co-funding the BCRC Fertility DMI project. The validation results showed that 100-point change in the sire's FPI equaled $154 CDN per feeder in net income if the bull was selected based on FPI, and an increase of $22,725 net revenue from 100 heifers if they were selected based on their RHPI score. These validation results support the application of the FPI and RHPI score that the project has offered to the producers. Moreover, these economic benefits are equivalent to a six-to-one return on investment for selection of RHPI score and FPI, which exceeds the 2.5-fold economic return on their investment on genotyping their animals as proposed in the project target.

5.11. Discussion of any challenges, delays or obstacles encountered during the project

The project has successfully completed all proposed activities, and all objectives and deliverables have been achieved. There have been no delays or obstacles encountered during the project.

6. Greenhouse Gas Benefits

6.1. Description of how the completed project and the advanced technology will result in GHG reductions in Alberta, and whether the reductions are direct or indirect.

Beef cattle production is one of the major contributors of GHG in agriculture, accounting for 31% of agricultural emissions. Production of beef cattle with improved genetics on feed efficiency and hybrid vigor will lead to direct reduction of GHG emissions. Although an integrated approach including genetic selection and breeding, diet manipulation or supplementation, rumen microbiota measurement and modification, plus advanced beef production and grazing management is recommended to reduce GHG emissions in beef cattle, genetic selection and breeding offers some unique advantages over other methods, which incudes:

(1). Unlike other methods such as diet manipulation or supplementation, the genetic progress made through genetic selection and breeding on GHG emission reduction is accumulative, although it could be slow and may take multiple generations to reach full potential. Furthermore, genetic selection and breeding including increasing hybrid score through crossbreeding have also ready been part of beef production, there will be no concerns of its negative impacts on animal health and palatability, which however may be an issue with diet manipulation or supplementation.

(2). The developed and refined the Genomics-enhanced Whole Herd Genetic Management Platform that integrates the web-site portal and the optimized genomic prediction pipeline along with the genomic prediction reference data sets offers an important and valuable suite of genomic tools for the beef cattle industry to improve beef production efficiency and quality. Mostly importantly, the economic benefits of adopting the genomic tools have been validated to be a six-to-one return on investment on genomic selection, which represents a win-win situation for the beef industry to use the genomic tools to not only enhance net profitability but also to reduce GHG emissions.

(3). The genomic tools including the Genomics-enhanced Whole Herd Genetic Management Platform were developed by researchers from Alberta (i.e. AAFC Lacombe and University of Alberta) and these tools have been demonstrated on over 10,000 beef cattle in Alberta. Through the demonstration and efforts of the technology transfer team at the Livestock Gentec, the genomic tools have gained popularity among Alberta beef producers, which has paved the way for more Alberta beef producers to use the genomic tools to select and breed cattle with improved genetics on feed efficiency and hybrid vigor in years to come. The Platform can also be applied to beef cattle across Canada. This will ultimately lead to direct reductions of GHG emissions from the beef production.

6.2. Discussion regarding how the completed project will help facilitate a low-carbon economy and secure Alberta's success in a GHG-constrained future.

Commercial beef producers represent over 90% of beef production in Alberta and most of them do not have access to breeding tools from breed associations, especially for feed efficiency and GHG emission traits. Moreover, commercial beef producers lack scientific tools to make the best use of retained heterozygosity via crossbreeding as this requires sophisticated record keeping and pedigree information which is unavailable for these herds which use natural mating except via genomics. Therefore, our Whole

Herd Genetic Management Platform is particularly useful to commercial beef producers, and it will benefit the whole beef industry as well. Since the Platform is able to predict gEPDs on expensive and difficult to measure traits such as feed efficiency, and female fertility related traits, and feedlot profitability index and replacement heifer profitability index score, beef producers can use the genomic values to select and breed beef cattle with improved feed efficiency and productivity. Breed associations or breeders can use the Platform to screen their bulls for further genetic evaluation. Feedlot operators can use the platform to sort their feeder cattle into more uniform finishing groups or to feed the cattle based on their genetic potential to maximize profits. All of these will lead to increased beef production efficiency and thus reduce enteric methane and GHG emissions.

6.3. Quantification of the annual GHG reductions resulting directly from implementation of the completed project

The project aimed to demonstrate the impact of Genomics-Enhanced Whole Herd Genetic Management Platform on reducing beef greenhouse gas emissions. Quantification of specified greenhouse gas emission reduction is based on basic or average GHG emissions for different beef cattle types (bulls, cows, calves), and GHG emissions reduction rates of different cattle types (bulls, cows, calves) due to selection of better genetic merit, i.e. gEPDs, and due to selection for greater hybrid vigor, i.e. gRHET. However, actual measuring GHG emissions on selected cattle in the field are not practical at this stage. Therefore, quantification of GHG emissions reduction is predicted based on parameters and methods or quantification protocols from multiple research publications (Basarab et al. 2012, Animals 2012, 2, 195-220; Plastow et al. 2018, Final project Final report to Beef Cattle Research Council (BCRC) and Alberta Agriculture and Forestry). The quantification of GHG emissions reduction protocol has been described in the MMV document that has been approved by the ERA GHG team.

In summary: (1): the base line or average of bull GHG emissions is 4,506 kgCO2e/head/year, and the base line of GHG emissions for cows is 3,636 kgCO2e/head/year. (2). 20% replacement rate or selection rate was assumed for analyzed bulls and cows/heifers, i. e. 20% of the analyzed bulls and cows will be selected as breeding stock; (3). The rate of reduction on GHG emissions due to selection on genetic merit (i.e. gEPDs) of improved feed efficiency is estimated as 0.7% per year per animal; (4): The rate of reduction on GHG emissions due to selection on improved retained hybrid vigor (i.e. gRHET) is estimated at 30% per year per animal.

In the first year of the project, the team analyzed and released gEPDs of feed efficiency and gRHET on 6,991 cattle. Of the 6,991 cattle, the team was sure that 1,768 were bulls and 4,013 were cows/heifers based on the animal information submitted by the producers. Of the 1,768 bulls and 4,013 heifers/cows, 1,761 bulls and 4,005 heifers/cows were from Alberta and were used in the calculation of GHG emissions reduction in Alberta. Based on these numbers, the estimated reductions of GHG emissions in year 1 due to genetic selection are presented in Table 9, which was also reported for milestone 1.

In the second year, the team analyzed additional 3,424 cattle, bringing the total number of cattle to 10,415. Of the 10,415 cattle, 10,217 cattle including 2,744 bulls and 7,473 heifers/cows from Alberta. Releasing the gEPDs of feed efficiency, gRHET of these Alberta bulls and heifers/cows and the selection indexes would lead to a further reduction in GHG emissions in year 2 for Alberta due to genetic selection, which is estimated below in Table 9.

Table 9. Prediction of GHG emissions reduction due to genetic selection on 1,761 bulls and 4,005 heifers/cows from Alberta analyzed in the first year and a total of 2,744 bulls and 7,473 heifers/cows from Alberta as of the second year from the project

Year	Cattle from Alberta	20% analyzed cattle selected as breeding stock	Average GHG/hd in tone(t) per animal per year before selection	Total GHG per cattle type before selection	GHG reduction 0.7% per hd due to selection on gEPDs of feed efficiency in tone(t)	GHG reduction 30% due to selection on gRHET selection in tone(t)	total GHG reduction of gEPD and gRHET selection in tone(t)
Year 1	1761 bulls analyzed	352	4.506	1587.0	11.1	476.1	487.2
Year 1	4005 heifers/cows analyzed	801	3.636	2912.4	20.4	873.7	894.1
Year 1	Total	1153		4499.4	31.5	1349.8	1381.3
Year 2	Total 2744 bulls analyzed	549	4.506	2473.8	17.3	741.9	759.2
Year 2	Total 7473 heifers/cows analyzed	1495	3.636	5435.8	38.0	1630.3	1668.3
Year 2	New born calves from heifers/cows selected in year 1	681	1.02	694.6	4.9	208.4	213.2
Year 2	Total	2044		8604.2	60.2	2580.6	2640.8
*Sum of GHG reducti					91.7	3930.4	4022.1

on Year 1 and Year 2							

** Sum of GHG reduction Year 1 and Year 2 represents an accumulated reduction as of year 2 of the project as it is assumed that the selected bulls and heifers/cows in year 1 will remain in the herd for reproduction in year 2, i.e. 2024.*

As shown in Table 9 above, the anticipated GHG emissions reduction of the first year is 1,381.3 tones. This is 254.4 tones more than the 1,126.9 tons of GHG emission reduction as originally expected in the proposal (Table 10 below). The greater amount of anticipated GHG emissions is due to the total number of cattle analyzed in the first year is more than 5,000 cattle that was originally planned.

Table 10. GHG emissions reduction in Alberta originally expected from the first year of the project

Cattle from Alberta	**20% analyzed cattle selected as breeding stock**	**Average GHG/hd in tone(t) per animal per year before selection**	**Total GHG per cattle type before selection**	**GHG reduction 0.7% per hd due to selection on gEPDs of feed efficiency in tone(t)**	**GHG reduction 30% due to selection on gRHET selection in tone(t)**	**total GHG reduction of gEPD and gRHET selection in tone(t)**
200 bulls analyzed	40	4.506	180.2	1.3	54.1	55.3
4800 heifers/cows analyzed	960	3.636	3490.6	24.4	1047.2	1071.6
Total	1000		3670.8	25.7	1101.2	1126.9

The anticipated GHG emissions reduction of the second year is 4,022.1 tones. This is about 1,513 tones more than the 2,509.3 tons of GHG emission reduction as originally expected in the proposal (Table 11 below). The greater amount of anticipated GHG emissions in year 2 is due to two reasons: (1), significantly more cattle analyzed and results released in year 1 than it was originally planned (i.e. (1761+4005)-5000=766 more cattle), as a results, greater amount of GHG emission reduction started in year 1 and this reduction is accumulated to year 2; (2). The team analyzed and released results on 217 more cattle in Alberta than it was originally planned (i.e. 10,217-10,000=217 more cattle), which leads to additional reduction of GHG emissions as a result of genetic selection based on the predicted genomic profiles.

Table 11. GHG emissions reduction in Alberta originally expected from the second year of the project

Cattle from Alberta	20% analyzed cattle selected as breeding stock	Average GHG/hd in tone(t) per animal per year before selection	Total GHG per cattle type before selection	GHG reduction 0.7% per hd due to selection on gEPDs of feed efficiency in tone(t)	GHG reduction 30% due to selection on gRHET selection in tone(t)	total GHG reduction of gEPD and gRHET selection in tone(t)
200 bulls analyzed	40	4.506	180.2	1.3	54.1	55.3
4800 heifers/cows analyzed	960	3.636	3490.6	24.4	1047.2	1071.6
New born from cows selected in year 1	816	1.02	832.3	5.8	249.7	255.5
Cumulative reduction of GHG e from year 1				25.7	1101.2	1126.9
Sum of GHG reduction from Year 1 and Year 2				57.2	2452.2	2509.3

6.4. Forecast of the estimated annual GHG reductions in Alberta from commercial roll out of the technology in Alberta by the years 2025, 2030, and thereafter.

With an approximate 1,805,660 beef bulls, cows and new born calves in Alberta, and assuming a 5% adoption rate starting in 2025, about 90,283 cattle from Alberta could be analyzed via the Platform, which could lead to a reduction of GHG emission of 21,871 tCO2e/yr in Alberta (Table 12). Reduction of GHG emissions will continue and will increase every year in Alberta (Table 12), assuming an increase of 5% adoption per year. As such, the reduction of GHG emission from Alberta beef cattle will reach 498,083 tCO2e/yr by 2030 even at an adoption rate of 30%. This represents a 6% reduction on GHG emissions in comparison to the current baseline of 8,272,721 tCO2e/yr GHG emissions in Alberta (i.e.

498,083/8,272,721=6%). The reduction of GHG emission from Alberta beef cattle will continue as the adoption rate increases, and by 2040 (i.e. at an adoption rate of 80%) the GHG emission reduction will be 1,975,947 tCO2e/yr, which indicates that GHG emissions from beef cattle will be reduced by 24% in Alberta (i.e. 1,975,947/ 8,272,721=24%). Reduction of GHG emissions will continue in Alberta as a result of the commercial roll-out plan as shown in Table 12. The pace of GHG reduction per year will slow down as the effect of increasing gRHET is limited by its theoretical values of 100% (i.e. animals can only have 100% retained heterozygosity at maximum). However, the selection on gEPD will continue to reduce GHG emissions from beef cattle and the reduction as a result of improved gEPD and high gHRET maintained in the herd will reach 2,470,192 tCO2e/yr in Alberta or a 29.8% reduction (i.e. 2,470,192 /8,272,721=30%) by 2050 in comparison to the baseline GHG emissions (Table 12).

Beef cattle outside Alberta account about 59% of the total. The genomic tool/Platform can also be applied to beef cattle outside Alberta. During the project, a total of 198 cattle (11 bulls and 187 cows/heifers) were analyzed on cattle from outside Alberta in year 2, which led to a GHG reduction of 44.8 (tone) (Table 12). As for future GHG benefits outside Alberta, similar reduction rates on GHG emissions are observed from beef cattle of out of province (Table 12). By 2030 at the projected adoption rate of 30%, the reduction of GHG emissions from beef cattle out of province will reach 586,841 tCO2e/yr, which is 5.7% GHG emission reduction (i.e. 586,841/10,321,201=5.7%). By 2040 when the projected adoption rate is assumed 80%, GHG emissions will be 2,140,609 tCO2e/yr out of province or a 20.7% cut on GHG emissions from beef cattle out of Alberta (i.e. 2,140,609/10,321,201=20.7%). As the improvement of gEPD on RFI and management of high gRHET continue via the platform, the reduction of GHG emissions from out of province beef cattle will reach 2,676,041 tCO2e/yr by 2050, representing a 25.9 % cut on the GHG emissions in comparison to the baseline GHG emissions at the start of the proposed project (i.e. 2,676,041/10,321,201=25.9%) (Table 12).

Table 12. Summary Annual Project Level GHG Reductions In Alberta and in Out of Province using emission factors in Table 13 and 5% increase of adoption rate of the proposed genomics-enhanced Whole Herd Genetic Management Platform

Year	Adoption rate	Number of Alberta beef cattle analyzed via the platform	Annual Project Level Emission Reductions - Alberta [tCO2e/yr]	No. of Out-of-Province beef cattle analyzed via the platform	Annual Project Level Emission Reduction - Out of Province [tCO2e/yr]	Total Annual Project Level Emission Reductions [tCO2e/yr]
2022	0	0	0	0	0	0
2023	Project	5,766	1,381.3	0	0	1,381.3
2024	Project	10,217	4,022.1	198	44.8	4,066.9
2025	5%	90,283	21,871	97,807	23,694	45,565
2026	10%	180,566	67,143	195,614	72,738	139,881
2027	15%	270,850	137,585	293,420	149,050	286,635
2028	20%	361,133	232,886	391,227	252,293	485,179

2029	25%	451,416	353,047	489,034	382,468	735,515
2030	30%	541,699	498,083	586,841	539,590	1,037,672
2031	35%	631,982	643,775	684,648	697,423	1,341,197
2032	40%	722,266	789,932	782,454	855,760	1,645,692
2033	45%	812,549	936,555	880,261	1,014,601	1,951,156
2034	50%	902,832	1,083,643	978,068	1,173,947	2,257,590
2035	55%	993,115	1,231,197	1,075,875	1,333,797	2,564,993
2036	60%	1,083,398	1,379,216	1,173,682	1,494,151	2,873,366
2037	65%	1,173,682	1,527,700	1,271,488	1,655,009	3,182,709
2038	70%	1,263,965	1,676,650	1,369,295	1,816,371	3,493,022
2039	75%	1,354,248	1,826,066	1,467,102	1,978,238	3,804,304
2040	80%	1,444,531	1,975,947	1,564,909	2,140,609	4,116,556
2041	85%	1,453,051	2,142,481	1,574,139	2,321,021	4,463,503
2042	90%	1,453,051	2,250,393	1,574,139	2,437,926	4,688,318
2043	95%	1,453,051	2,333,807	1,574,139	2,528,291	4,862,098
2044	100%	1,453,051	2,392,724	1,574,139	2,592,118	4,984,842
2045	100%	1,453,051	2,427,144	1,574,139	2,629,406	5,056,549
2046	100%	1,453,051	2,437,066	1,574,139	2,640,155	5,077,221
2047	100%	1,453,051	2,445,347	1,574,139	2,649,126	5,094,474
2048	100%	1,453,051	2,453,629	1,574,139	2,658,098	5,111,727
2049	100%	1,453,051	2,461,910	1,574,139	2,667,070	5,128,980
2050	100%	1,453,051	2,470,192	1,574,139	2,676,041	5,146,233

6.5 Description of the baseline scenario for all GHG reduction calculations, including justification for why that is the most appropriate choice of baseline.

Quantification of GHG emissions from Alberta and out of Alberta was based on beef cattle numbers from Statistics Canada, 2020, and average amount of GHG emissions per animal per year. GHG emissions from cattle differ among cattle with different body weight and they ranged from 1,020 kg per head per year (kgCO2e/hd/yr) for young calves to 4,506 kgCO2e/hd/yr for bulls, with a weighted average of 2,029 kg/year (Basarab et al. 2012. Animals 2, 195-220). Therefore, cattle were grouped into Beef Bulls, Beef Cows, and Calves produced from beef cows. In Alberta, there are 81,200 Bulls, 1,464,200 Cows, and 2,532,355 Calves based an inventory of Statistics Canada, 2020 (Table 13). Out of Alberta, there are 119,200 Bulls, 2,097,200, Cows, and 2,116,340 Calves (Table 13). Therefore, the baseline GHG emissions in Alberta from beef cattle is estimated at 8,272,721 tCO2e/yr, while out of province, the baseline GHG emissions from beef cattle is 10,321,201 tCO2e/yr (Table 14).

As a direct measure on GHG emissions and on reduction of GHG emissions in cattle is not feasible on a large number of cattle at this stage, the baseline for quantification on GHG emissions reduction was based on previous studies including Basarab et al. 2012, Animals 2012, 2, 195-220; Plastow et al. 2018, Final project Final report to Beef Cattle Research Council (BCRC) and Alberta Agriculture and Forestry). Specifically, for genetic selection via gEPDs, a GHG reduction rate of 0.7% per year (yr) per head (hd) is assumed based on the realized genetic progress from the Kinsella breeding project (Plastow et al. 2018, Final project Final report to Beef Cattle Research Council (BCRC) and Alberta Agriculture and Forestry). As the genetic effects of gEPDs on GHG emissions will pass to the next generations so they are therefore cumulative in the calculation. The gHRET effect on GHG emission reduction is assumed 639 kgCO2e/hd/yr as a result of 30% increase of gRHET due to genetic selection based on a study by Basarab et al. Animals 2012, 2, 195-220, which is 31% per year reduction of GHG emissions (i.e. 639 kg/average GHG emission 2,029 kg per year per head=31%). In this project, it is assumed that only 20% of bulls and cows analyzed via the platform will benefit from gEPD and gRHET selection. The 20% was assumed because the average gRHET of top 20% animals is 27% greater than the average of all animals in a sample of crossbred population (n=997), which is very close to the 30% increase of gRHET used by Basarab et al. Animals 2012, 2, 195-220 in determining the effect of increased gRHET on reducing GHG emissions (i.e. the 639 kgCO2e/hd/yr reduction). The 20% selection rate is also applied to gEPD selection as it is similar to what the team used in the Kinsella breeding project. Therefore, of all cattle analyzed via the platform only 20% will benefit from the effects of increased gRHET and improved gEPD on RFI, and the reduction rate of gRHET and gEPD on GHG emissions is set at 30% and 0.07% per year per head, respectively. The effect of gRHET on reduction of GHG emission on bulls and cows will last for 5 years as it assumed that they will stay in the herd for 5 years on average. The effects of gEPD and gRHET on calves were calculated on all new born calves from only the selected bulls and selected cows, assuming 85% calf weaning rate, and the effects of gEPD and gRHET on reduction of GHG emissions for calves will last only for 1 year assuming most of them will be finished a year after weaning. Details of these Emission Factors are provided in Table 13.

Table 13. Emission Factors and Activity Data used in Greenhouse Gas Quantification under the proposed project "Demonstrating the impact of genomics-enhanced whole herd genetic management platform on reducing beef greenhouse gas emissions" (B0160957)

Emission Factor	Total number of cattle in Alberta (head)	Total number of cattle out of province in Canada	Activity Data units	Emission Factor Value	EF Units*
Bulls before selection	81,200	119,200	per head (hd) per year (yr)	4,506.00	kgCO2e/hd/yr
Cows before selectin	1,464,200	2,097,200	per head (hd) per year (yr)	3,636.00	kgCO2e/hd/yr
All calves before selection	2,532,355	2,116,340	per head (hd) per year (yr)	1,020.00	kgCO2e/hd/yr
Total	4,077,755	4,332,740			

Emission Factor (continued)	**CO2e reduction rate via genetic merit gEPD selection, 20% cattle will be selected as parents and 100% of their new born calves beneft**	**Number of animals analyzed each year**	**CO2e reduction rate via genome retained heterozygosity gRHET) selection, 20% cattle will benefit**	**Model of calculation**	**EF Units**
Selected Bulls via running animals by the platform	0.07% per year per head	200 bulls each year in the first two years, 2023, 2024, and 5% adoption in 2025 and 5% increase of adoption/yr after 2025 on numbers of bulls and cows on the farms.	30% per year per head	gEPD effects pass to next generations. gHRET effects last for for 5 years in the herd.	kgCO2e/hd/yr
Selected Cows via running animals by the platform	0.07% per year per head	4800 young cows/heifers for replacemet each year in the first two years, 2023, 2024, and 5% adoption/yr in 2025 and 5% increase of adoption after 2025 on numbers of	30% per year per head	gEPD effects pass to next generations. gHRET effects last for for 5 years in the herd.	kgCO2e/hd/yr

		bulls and cows on the farms.			

** EF Reference/Source: (1). Basarab et al. 2012. Greenhouse Gas Emissions from Calf- and Yearling-Fed Beef Production Systems, With and Without the Use of Growth Promotants. Animals 2012, 2, 195-220; doi:10.3390/ani2020195; (2). Statistics Canada, 2020, Alberta shares 41% bulls and cows, i.e. (81,200+1,464,200)/(81,200+1,464,200+119,200+2,097,200)=0.41. (3). Plastow, G. Fitzsimmons C., Li C., et al. Improvement of cow feed efficiency and the production of consistent quality beef using molecular breeding values for RFI and carcass traits (i.e. the Kinsella breeding project) (2018), Final report to Beef Cattle Research Council (BCRC) and Alberta Agriculture and Forestry.*

Table 14. Baseline and Project Emissions for Relevant Sources and Sinks for the proposed project "Demonstrating the impact of genomics-enhanced whole herd genetic management platform on reducing beef greenhouse gas emissions" (B0160957)

Project start: Sept 1, 2021. **Project ends: August 31, 2024**	**Number of cattle in Alberta in 2021, 2022**	**Number of cattle in Alberta in 2023**	**Number of cattle in Alberta in 2024**	**Emission Factor Value, GHG emission in kg per animal per year**	**2021**	**2022**	**2023**	**2024**
Source, Sink, Reservoir (SSR) - Alberta - Emissions					**[tCO2e/yr]**	**[tCO2e/yr]**	**[tCO2e/yr]**	**[tCO2e/yr]**
B1 - Bulls [tCO2e/yr]	81,200	81,200	81,200	4,506	365,887	365,887	365,887	365,887
B2 - Cows [tCO2e/yr]	1,464,200	1,464,200	1,464,200	3,636	5,323,831	5,323,831	5,323,831	5,323,831

B3- Calves, [tCO2e/yr]	2,532,355	2,532,355	2,532,355	1,020	2,583,002	2,583,002	2,583,002	2,583,002
Total	4,077,755	4,077,755	4,077,755		8,272,721	8,272,721	8,272,721	8,272,721
Project start: Sept 1, 2021. **Project ends: August 31, 2024**	**Number of cattle with improved genetics via the platform**	**Number of cattle with improved genetics via the platform**	**Number of cattle with improved genetics via the platform**	**Emission Factor Value, GHG emission in kg per animal per year**	**2021**	**2022**	**2023**	**2024**
Source, Sink, Reservoir (SSR) - Alberta - Emissions					**[tCO2e/yr]**	**[tCO2e/yr]**	**[tCO2e/yr]**	**[tCO2e/yr]**
P1 - Bulls [tCO2e/yr], replaced with selected bulls	0	40	40	4,506	365,887	365,887	365,832	365,777
P2 - Cows [tCO2e/yr], replaced with selected cows	0	960	960	3,636	5,323,831	5,323,831	5,322,760	5,321,688
P3- Calves [tCO2e/yr], with new born calves with selected parents	0	0	816	1,020	2,583,002	2,583,002	2,583,002	2,582,747

Total	0	1,000	1,000		8,272,721	8,272,721	8,271,594	8,270,211
Emission Reductions [tCO2e/yr]					0	0	1,127	2,509
Project start: Sept 1, 2021. **Project ends: August 31, 2024**	**Number of catttle out of Aberta in 2021, 2022**	**Number of catttle out of Aberta in 2023**	**Number of catttle out of Aberta in 2024**	**Emission Factor Value, GHG emission in kg per animal per year**	**2021**	**2022**	**2023**	**2024**
Source, Sink, Reservoir (SSR) - out of Alberta - Emissions					**[tCO2e/yr]**	**[tCO2e/yr]**	**[tCO2e/yr]**	**[tCO2e/yr]**
B1 - Bulls [tCO2e/yr]	119,200	119,200	119,200	4,506	537,115	537,115	537,115	537,115
B2 - Cows [tCO2e/yr]	2,097,200	2,097,200	2,097,200	3,636	7,625,419	7,625,419	7,625,419	7,625,419
B3- Calves [tCO2e/yr]	2,116,340	2,116,340	2,116,340	1,020	2,158,667	2,158,667	2,158,667	2,158,667
Total	4,332,740	4,332,740	4,332,740		10,321,201	10,321,201	10,321,201	10,321,201
Source, Sink, Reservoir (SSR) - out of Alberta - Emissions	**Number of cattle out of Aberta with improved genetics**	**Number of cattle out of Aberta with improved genetics**	**Number of cattle out of Aberta with improved genetics**	**Emission Factor Value, GHG emission in kg for**	**[tCO2e/yr]**	**[tCO2e/yr]**	**[tCO2e/yr]**	**[tCO2e/yr]**

	via the platform	via the platform	via the platform	animal per year				
B1 - Bulls [tCO2e/yr]	0	0	0	4,506	537,115	537,115	537,115	537,115
B2 - Cows [tCO2e/yr]	0	0	0	3,636	7,625,419	7,625,419	7,625,419	7,625,419
B3- Calves [tCO2e/yr]	0	0	0	1,020	2,158,667	2,158,667	2,158,667	2,158,667
Total	0	0	0		10,321,201	10,321,201	10,321,201	10,321,201
Emission Reductions [tCO2e/yr]					0	0	0	0

7. Economic and Environmental Impacts

7.1. Description of the projected economic impacts in Alberta, including revenues, cost savings, job creation, investment attraction, economic diversification, tax revenue, etc., based on the outcomes of the project.

Economic benefits of implementing the genomic tools, in particular on index selection on FPI and RHPI score were validated using independent industry cattle data through the complementary RDAR Feeder Profit Index validation project and through the co-funding the BCRC Fertility DMI project. The validation results showed that 100-point change in the sire's FPI equaled $154 CDN per feeder in net income if the bull was selected based on the FPI generated by the Platform, and an increase of $22,725 net revenue from 100 heifers if they were selected based on their RHPI score, which was about $227.25 per selected heifer. Given a total 2.5 million cattle to be finished to the market weight in Alberta each year (Statistics Canada 2020), and with even a 5% rate of adoption of the Platform, i.e. 5% of the bulls were selected based on the FPI, it could translate into a saving of $19.2 million alone on feed per year for the Alberta beef industry.

With over 1.46 million beef cows in Alberta (Statistics Canada 2020), selection of replacement heifers and maintaining cows with a RHPI score greater than 48.4 would lead to an increase in profit of over $16.4 million per year even at a conservative 5% adoption rate. With a 30% adoption rate, the increase in profit could be $99.6 million per year.

These saving on production costs and increase in profit per year will lead to more profitable beef production in Alberta, which will attract more investment and result in more economic diversification, job creation and tax revenue related to the value chain of Alberta beef production.

7.2. Discussion about the immediate and potential future environmental benefits (including criteria air contaminants, land use, soil, water consumption, etc.) resulting from the completed project and advanced technology.

The developed the Genomics-enhanced Whole Herd Genetic Management Platform represents the most advanced suite of genomic tools in Canada that allows commercial beef producers to select and breed cattle with superior genetics to improve feed efficiency and to enhance female cattle reproductive performance. Improved feed efficiency and better female cattle reproductive performance will maintain the same amount of beef production but with reduced land use, reduced water consumption, and reduced environmental footprints including GHG emissions. This also contributes to a more profitable and sustainable beef production of future generations.

7.3. Description of how the project has resulted in increased innovation capacity in the province through training of highly-skilled personnel, knowledge development, postsecondary partnerships, research organizations, startup companies, etc.

Beef performance traits are complex and they are controlled by many genes with smaller effects. Development of genomic prediction tools requires in-depth understanding on animal genetics, quantitative genomics, animal breeding and production. The optimization or refinement of genomic prediction pipelines also require statistical and computer programing skills. The cost-benefit analyses on implementing the genomic tools need knowledge of beef production system and skills in economics. This ERA project and the two co-funding projects have trained highly-skilled personnel including two post-doctoral researchers in quantitative genetics/genomics, one post-doctoral researcher in quantitative genomics, bioinformatics and computer programming, one research assistant in running the genomic prediction pipeline, two biologists in website design, database management, and DNA sequence analysis, three summer students on economic analyses, Furthermore, two highly-skilled personnel have also been trained on technology transfer of the genomic tools to the industry through Livestock Gentec, University of Alberta during the project. All of these trainings have led to increased innovation capacity in development and implementation of genomic tools in beef cattle.

8. Overall Conclusions

Through consolidating and further refining large historical beef cattle data sets from multiple previous projects and exploring various genomic analysis methods, a genomic prediction pipeline has been developed and optimized to analyze genomic profiles of Canadian beef cattle. Subsequently, a Genomics-enhanced Whole Herd Genetic Management Platform has been successfully developed for industry application. The Platform integrates a website portal for producer access and the genomic prediction pipeline utilizing the refined reference beef cattle data sets. This enables the generation of genomic profiles for individual cattle including genomic breed composition and genomic retained hybrid vigor, genetic merit for growth, feed intake, feed efficiency, carcass, fertility traits, multiple trait selection indexes including feeder profitability index and replacement heifer profitability index and score, and potential of bulls to produce hybrid vigor in offspring. These genomic profiles allow producers to make informed selection on bulls and replacement heifers to maximize genetics of beef cattle herds. In collaboration with the co-funding BCRC Fertility DMI project, the co-funding BCRC Genomics Platform project, and with the complementary RDAR Feeder Profit Index validation project, the benefits of implementing the above genomic tools have been validated in industry beef cattle and the genomic tools have been successfully deployed to over 10,415 beef cattle submitted by more than 88 Canadian beef producers and organizations. Although further research efforts are required to continue calibrating the genomic prediction with additional data sets and more advanced genomic analysis methods, the Platform represents an advanced and valuable suite of genomic tools for the Canadian beef industry to select and breed beef cattle with improved profitability and a reduced environmental footprint.

9. Scientific Achievements

9.1. List of all applied for or obtained patents, published books, journal articles, conference presentations, student theses, etc., based on work conducted during the project.

A list of publications in scientific journals:

1. John Basarab, Tiago Valente, Yining Wang, Michael Vinsky, Changxi Li, Graham Plastow, 2023. Validation of genomic tools to improve production efficiency, health resilience and the carbon footprint of beef production. Proceedings, 56th American Association of Bovine Practitioners Annual Conference, Vol: 56, No: 2, https://doi.org/10.21423/aabppro20238840. (The MBVs, FPI and RHPI were predicted using the reference populations and genomic prediction pipeline of the Platform).

2. Changxi Li, John A Basarab, Tiago Valente, Graham Plastow, Paul Stothard, Yining Wang, Michael Vinsky. 2024. Development, deployment, and calibration of genomic selection tools to improve performance traits for Canadian beef cattle, conference abstract, Journal of Animal Science, Volume 102, Issue Supplement_3, September 2024, Pages 179–180, https://doi.org/10.1093/jas/skae234.210. (The current project contributed to development and deployment of genomic selection tools used in the study).

3. Hongwei Li, Yining Wang, Michael Vinsky, Tiago Valente, John A Basarab, Changxi Li, 2024. Accuracy of genomic predictions using single and multiple-trait machine learning methods in Canadian beef cattle population, conference abstract, Journal of Animal Science, Volume 102, Issue Supplement_3, September 2024, Pages 33–34, https://doi.org/10.1093/jas/skae234.037 (The current project contributed to the analyses of genomic prediction using single and multiple-trait machine learning methods in Canadian beef cattle populations).

A list of conference presentations:

1. 2024 ASAS-CSAS-WSASAS Annual Meeting, July 21-25 at the Calgary Convention Centre in Calgary, Alberta, Canada: Hongwei Li et al. presented "Accuracy of genomic predictions using single and multiple-trait machine learning methods in Canadian beef cattle population" as an oral presentation under the session of animal genetics and breeding (travel expenses to the conference was covered by the current ERA project).

2. 2024 ASAS-CSAS-WSASAS Annual Meeting, July 21-25 at the Calgary Convention Centre in Calgary, Alberta, Canada: Changxi Li et al. presented "Development, deployment, and calibration of genomic

selection tools to improve performance traits for Canadian beef cattle" as an oral presentation during the animal genetics symposium (travel expenses to the conference was covered by the current ERA project).

10. Next Steps

10.1. Discussion about the next steps for the technology/process/innovation, including potential follow-up projects.

The next steps will be to deploy genomic tools to the Alberta and Canadian beef industry through Livestock Gentec, an Alberta Innovates Centre that brings the commercial benefits of genomics to the Canadian livestock industry. This involves signing agreements between Agriculture and Agri-Food Canada (AAFC) and the University of Alberta for genomic prediction reference data transfer and for genomic analysis pipeline license transfer. Some progress was made with the transfer of two Refined Reference Data Sets from AAFC to Livestock Gentec, University of Alberta on October 9, 2024. The two Reference Data Sets are:

1) Refined beef cattle genomic prediction reference data set for feed efficiency and carcass merit traits (Version 1, 2021)

2) Refined beef cattle pure breed SNP reference panel for breed composition prediction (Version 1, 2021).

However, as of October 15, 2024, the AAFC researchers are still waiting for the license transfer agreement to be signed to grant the University of Alberta a license to use the pipeline (e.g., procedures and algorithms) that generates genomic values. Therefore, Gentec researchers still cannot generate genomic values for Alberta and Canadian beef producers. As a temporary solution, genomic values are still being generated by the AAFC staff. As soon as the agreement is signed, the AAFC team will transfer the genomic prediction pipeline to Livestock Gentec to conduct "serve to industry" work, which is the stated objective of Livestock Gentec, and the RDAR Professor for Beef Genetics and Genomics (JA Basarab). Once the license of genomic prediction pipeline is granted, Gentec researchers will continue to deploy genomic breed composition, hybrid vigour score, molecular breeding values (MBVs) for 14 traits, and DNA-based selection indexes for replacement heifer selection and bulls that sire profitable feeder cattle on 40,000 cattle over the next three years under project RES0067061. In addition, MBVs for methane emission reduction traits will be developed through a NSERC Alliance project which will also develop a carbon offset protocol for greenhouse gas (GHG) reduction for the cow-calf sector. This project will collaborate and share resources with the novel vaccine project cited below. Future projects include evaluating of feed conversion ratio as a more effective trait than residual feed intake as a way to select terminal bulls for feed efficiency and to validate molecular heterozygosity at the SNP and chromosomal level as more effective measurement of the benefits of genomic diversity and heterosis in crossbred and purebred beef cattle. As of October 15, 2024, three projects related to the next steps for technology deployment have been approved. They are:

1. RES0067061, Apr 1, 2024 to Dec 21, 2027, PI = JA Basarab, co-PI = Plastow: 2024G2876R - Using genomics to improve the profitability and sustainability of Alberta's beef industry. Total project cost $4,518,338; Sustainable CAP funded under the administration of RDAR = $2,199,950.

2. RES0065472, May 15, 2024 to May 14, 2029, PI = JA Basarab: Reducing Greenhouse Gas Emissions from the Canadian Beef Industry through the Development and Adoption of Genomic Tools. Total project budget = $5,647,832; received from NSERC Alliance - SSHRC = $3,803,801.

3. RES0064895, Apr 1, 2024 to Mar 31, 2027, PI = JA Basarab: Sire-based Feeder Profit Index for Beef on Dairy production and marketing systems- 2024N3297R. Total project budget = $6,456,380; Sustainable CAP funded under the administration of RDAR = $736,006.

10.2. Long-term plan for commercialization of the project technology/learnings.

Commercialization will occur through Livestock Gentec in partnership with AgSights, Telus Agriculture, Cow-calf Health Management Solutions (CCHMS), and GK Jim Farms, and commercialization activities will be managed by the commercialization committee under the Bovine Genomics and Forage Co-location Agreement between AAFC and the University of Alberta.

10.3. Commercialization-related actions to be undertaken within two years of project completion.

The action includes acquiring a license from AAFC to use their pipeline to generate genomic values for the Alberta and Canadian beef industry and deploying genomic values on 20,000 beef cattle to over 100 beef producers within the next two years. This will occur in collaboration with AgSights, Telus Agriculture, CCHMS, Dalhousie University, Maritimes Bull Test Station, Olds College, and Lakeland College. Custom indexes will be explored through projects with GK Jim Farms. Discussions on commercialization activities will start through the commercialization committee within two years under the Bovine Genomics and Forage Co-location Agreement between AAFC and the University of Alberta.

10.4. Potential partnerships under development with technology integrators, adopters, etc.

Partners include AgSights, Telus Agriculture, Olds College, Lakeland College, Canadian Galloway Association, Maritimes Bull Test Station and GK Jim Farms.

11. Communications plan

11.1. Summary of key knowledge-sharing or communications activities undertaken during the project.

Communications activities, key knowledge-sharing and technology transfer activities were carried out through this project and collaborative efforts with co-funding and complementary projects including co-funding BCRC Fertility DMI project and BCRC Genomics Platform Development project, and through collaborations with Livestock Gentec, the University of Alberta. A list of technology transfer activities includes:

1. September 2021 to August 31, 2024 and beyond: The project website (https://www.beefgenomicprediction.ca/) has served as a main channel of technology transfer. The website is open to the public and it contains comprehensive information about genomics including explanations on commonly used terminologies in animal genetics, illustration of breeding and selection methods, interpretation of the genomic profile results and recommendations on how to use the genomic prediction results to select and breed cattle based on breeding objectives.

2. September 2021 to August 31, 2024: A major technology transfer activity was the deployment of genomic tools via demonstration on 10,415 cattle submitted by 88 beef producers/organizations to this Emissions Reduction Alberta (ERA) Genomics Platform Demonstration project. Genomic profiles of the 10,415 cattle including gBC, gRHET, gEPDs of RFI, ADG, DMI, HCW, AFAT, CREA, LMY, CMAR, FPI, and genomic profiles of each heifer/cow including gBC, gRHET, gEPDs of BRWT, JulianDT, WT200d, DUR, on-test DMI, RFIfat, PBWT, PBBF, AFC, RHPI, RHPI score along with bull's potential to produce offspring with greater hybrid vigor in the herd (i.e. AHPH) have been released to the producers to assist with their genetic selection and breeding decisions.

3. September 2021 to August 31, 2024: In addition to analyzing and delivering genomic profiles of the 10,427 cattle submitted to this Emissions Reduction Alberta (ERA) Genomics Platform Demonstration project, the developed and refined Platform was also used to analyze and share genomic prediction results with multiple industry partners and researchers of other projects, which included (1). Analyzing and sharing individual animal genomic breed composition (gBC) of 3,477 cattle (2,971 for Beefbooster and 506 and WA Ranch) for the herd DNA pooling project (PI: John Basarab, funded by Results Driven Agriculture Research (RDAR), 2021-2022). (2). Analyzing and sharing individual animal genomic breed composition (gBC), gRHET, gEPDs of RFI, ADG, DMI, HCW, AFAT, CREA, LMY, CMAR, and FPI values on 6,015 steers/bulls submitted by the industry partners for the complementary RDAR Feeder Profit Index validation and deployment project (PI: John Basarab, co-PI: Changxi Li, 2022-2024); (3): Analyzing and sharing gBC, gRHET, gEPDs of BRWT, WT200D, RFIfat, DUR, PBWT, PBBF, AFC, JulianDT, DMI, RHPI and RHPI score on 189 heifers/cows for project "Replacement Heifer Selection – Demonstration project" (PI: Susan Markus, co-PI: Obioha Durunna, funded by RDAR, 2022-2023), allowing the project to showcase the genomic tools for replacement heifer selection to enhance cow reproductive performance; (4): Predicting and sharing individual gBC and gRHET on 176 cows from the Roy Berg Kinsella Ranch, University of Albert for project "Can cow/calf production efficiency be explained by maternal habitat selection and dietary composition in diverse pasture?" (PI: Edward Bork, funded by BCRC/ Results Driven Agriculture Research (RDAR), 2021-2024). These individual animal gBC and gRHET allowed the above project to relate gBC and gRHET to cow's behaviors in habitant selection under different pastures.

4. Virtual information session (March 29, 2022): Changxi Li et al. presented in an industry webinar on "Using genomics-enhanced tools to improve performance of whole beef cattle herds".

5. The Kinsella field day (July 20, 2022): Changxi Li et al. presented "Introduction to beef cattle genomic prediction" to approximately 150 attendances.

6. Virtual meeting with producers (October 27th 2022): Changxi Li, John Basarab, Michael Vinsky hosted an industry information session online to help interpret the genomic results and provide advices on how to use the genomic profile results to improve genetics of their herd.

7. The Lacombe field day (July 26th 2023): Changxi Li et al. presented “Genomic prediction tools and their application to improve beef production efficiency, carcass quality, and female production traits - gEPD prediction and multi-trait selection indexes” to approximately 200 attendances.

8. The Lacombe field day (July 26th 2023): Changxi Li et al. presented “Genomic prediction tools and their application to monitor and enhance hybrid vigor in crossbred beef cattle” to approximately 200 attendances.

9. 2024 ASAS-CSAS-WSASAS Annual Meeting, July 21-25 at the Calgary Convention Centre in Calgary, Alberta, Canada: Changxi Li et al. presented “Development, deployment, and calibration of genomic selection tools to improve performance traits for Canadian beef cattle” as an oral presentation during the animal genetics symposium.

10. 2024 ASAS-CSAS-WSASAS Annual Meeting, July 21-25 at the Calgary Convention Centre in Calgary, Alberta, Canada: Hongwei Li et al. presented “Accuracy of genomic predictions using single and multiple-trait machine learning methods in Canadian beef cattle population” as an oral presentation under the session of animal genetics and breeding.

11.2. Plans for communicating information about the project, project findings, and results or the underlying technology with third parties, including a description of communication tools that will be used.

The project website (https://www.beefgenomicprediction.ca/) will be transferred to Livestock Gentec and it will continue to be used as a channel to communicate this project to the industry and public. As the genomic tools will be deployed to 40,000 heifers/bulls/feeder cattle over 4 years under a recent funded “Adoption of Genomic Tools” project by RDAR/Sustainable CAPS (https://rdar.ca/latest/press-releases/rdar-supports-beef-genomics-technology-adoption-with-the-delivery-of-a-2-2m-investment-into-livestock-gentec), the project lead institute Livestock Gentec will continue communicating information about the genomic tools and Platform through newsletters, media interviews, field days, and industry meetings.

Scientific results and advancements on multiple trait genomic prediction and machine learning genomic prediction of genetic merit, on multiple trait selection indexes, on cost-benefit analyses of implementing the Platform, and on molecular indictors of heterosis will be communicated to the community through presentations at scientific conferences and through publications in peer-review scientific journals in the next few years.